Contents

Part A

Plenary Lectures

Keynote Lectures

Contributed Papers

Modeling, Supporting Methods and Tools

Oral

Modeling, Supporting Methods and Tools

Poster

Product and Process Design

Oral

Product and Process Design

Poster

Part B

Operations and Control

Oral

Operations and Control

Poster

Non-traditional Applications of Process Systems Engineering

Oral

Non-traditional Applications of Process Systems Engineering

Poster

Business Decision Support

Oral

Business Decision Support

Poster

Education in Process Systems Engineering

Oral

Education in Process Systems Engineering

Poster

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Technology Advances for Dynamic Real-Time Optimization

L. T. Biegler

Chemical Engineering Department, Carnegie Mellon University, Pittsburgh, PA 15213 USA

Abstract

Integration of real-time optimization and control is an essential task for profitable process operation in a highly competitive environment. While integrated large-scale optimization models have been formulated for this task, their size and complexity is often a challenge to many available optimization solvers. On the other hand, recent development of powerful, large-scale solvers leads to a reconsideration of these formulations, in particular, through application of efficient large-scale barrier methods for nonlinear programming (NLP). Moreover, in-built NLP sensitivity capabilities quickly compute approximate solutions of perturbed NLPs. This allows on-line computations to be drastically reduced, even when large nonlinear optimization models are considered. Moreover, extensions of these approaches to consider dynamic complementarity systems allow the treatment of a class of discrete decisions. These recent results motivate the development of dynamic real-time optimization (D-RTO) strategies that can be used to merge and replace the tasks of (steady-state) real-time optimization (RTO) and (linear) model predictive control (MPC). This is enabled by continuing advances in nonlinear model predictive control and its generalization to deal with economic objective functions. Finally, successful implementation of D-RTO suggests a bottom up strategy starting from MPC to build a single D-RTO layer with consistent dynamic models, rather than modifying the existing RTO layer.

Keywords: Model Predictive Control, Real-time Optimization, Dynamic Real-time Optimization

1. Introduction

With the development of efficient strategies for the optimization of differential-algebraic equation (DAE) systems, recent research has focused on the development of on-line optimization of dynamic systems. Currently, RTO systems consist of two layers, as shown in Figure 1. The MPC layer typically consists of linear dynamic models while the RTO layer consists of nonlinear steady state models. While this approach can work well when there is an adequate separation of time scales, difficulties have been reported in the two-layer structure due to inconsistency of different optimization models in each layer, inadequate treatment of unmeasured disturbances, and the corresponding impact on stability of the RTO loop. A number of these issues are discussed and analyzed in [16]. Many of these difficulties can be addressed through the adoption of a single layer that combines both MPC and RTO, as shown in Figure 1 and initially suggested in [13]. With provisions to deal with a consistent optimization model that is made compatible with the plant, better performance of the process can be expected and further extensions can be made that would better integrate on-line operations with planning, scheduling and other off-line decision-making tasks [3, 10].

Moreover, recent advances in the development of fast, large-scale NLP methods for dynamic optimization are key enablers to realizing D-RTO loops. This study explores these advances in dynamic optimization strategies, particularly for time-critical applications. In the next section, we briefly discuss fast nonlinear MPC (NMPC) strategies that include state/parameter estimation with nonlinear, dynamic models. Two case studies that demonstrate the effectiveness of these approaches are briefly summarized in Section 3. In Section 4 we describe additional elements that needed to incorporated within D-RTO. These include analysis of NMPC strategies with economic objective functions, treatment of uncertainty through robust problem formulations and embedded observers, and the incorporation of hybrid elements through the use of complementarity conditions that allow switching in the dynamic system. Finally, Section 5 concludes with a few remarks on the realization of D-RTO through a bottom-up approach that exploits these advances.

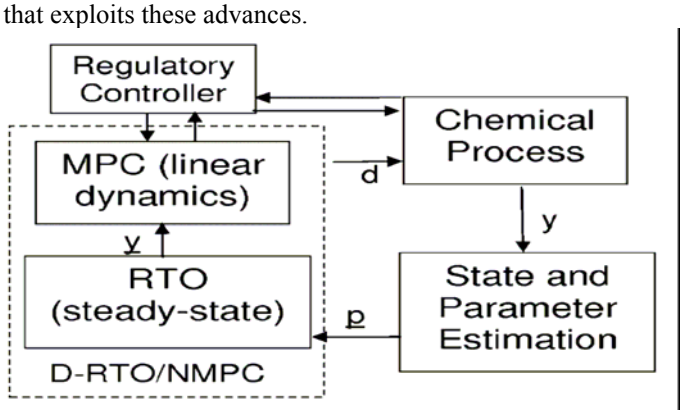


Figure 1. Emerging Structure for Dynamic Real-time Optimization

2. Advances in Model Predictive Control

Linear Model Predictive Control (MPC) has been widely applied for advanced control for over 30 years. Most recent MPC implementations rely on linear dynamic models in state-space form, which enable the treatment of feedforward disturbances, offset correction for model mismatch and robustness to noise. Moreover, as summarized in [6], numerous extensions to hybrid systems have been developed, including the use of complementarity conditions to model discrete decisions. The evolution of MPC to NMPC requires the application of nonlinear DAE models, particularly first principle models that describe the process over a wide range. This extension is essential for nonlinear processes that exhibit wide swings in operation, require frequent product transitions, and consist of multi-stage and non-standard operations. For over ten years, significant advances have been made in the solution of dynamic optimization problems, and they have been applied in practice for time-critical NMPC applications, see e.g., [1, 4, 5]. In particular, the simultaneous collocation approach discretizes the DAEs along with the state and control profiles to form a large-scale NLP, which can be solved very efficiently with full-space barrier NLP solvers such as IPOPT [14]. For a fixed set of parameters, p_0, the resulting NLP can be written as:

$$x^*(p_0) = \arg\{\min f(x, p_0),\ s.t.,\ c(x, p_0) = 0,\ x_l \le x \le x_u\} \quad (1)$$

Moreover, as IPOPT is a Newton method, which exploits exact first and second derivatives, we can easily compute the *sensitivity of the optimal solution* with negligible

computational cost. Under mild conditions, this scheme leads to a first order estimate of $x^*(p)$ where p is any small perturbation from p_0. For NMPC we solve (1) over a predictive time horizon at time t_k with an initial state estimate ($p_0 = x(k)$). A serious drawback of NMPC (and D-RTO) strategies has been computational feedback delay (i.e., the time between receipt of the measured output and injection of the input into the plant) induced by solution of a potentially large NLP. A large delay leads to deterioration of NMPC performance and possible loss of stability.

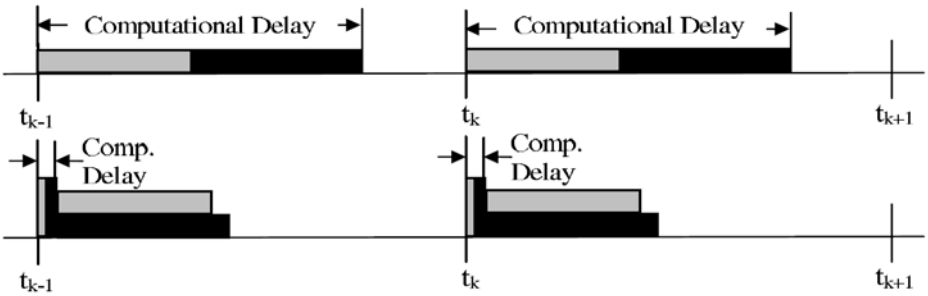


Figure 2: Top graphic depicts computational delay for MHE (gray) and NMPC (black). Bottom graphic indicates background asMHE (gray) and asNMPC (black) calculations (horizontal bars); vertical bars indicate sensitivity calculations.

Recent NMPC advances avoid this delay by restructuring the NLP calculations based on fast solution and sensitivity strategies. Note that the optimization model (1) remains the same for all time horizons, except for the initial state estimate in the horizon. As a result, a real-time iteration strategy (see [4, 15, 17]) allows decomposition of the NMPC into background solutions of (1), between the sampling times, updated with fast on-line *sensitivity* components, once the measured output is received. This approach, called advanced-step NMPC (asNMPC), was developed in [17, 19] and a corresponding nonlinear moving horizon estimator, called asMHE, was developed in [18]. The partitioning of background and on-line calculations is sketched in Figure 2. Even though the injected input is only a first order approximation to the optimum, asNMPC has surprisingly strong stability properties [17]:

- Assuming that NLP (1) can be solved within one sampling time ($t_k - t_{k-1}$) and ideal NMPC stability assumptions hold, then the nominal stability properties of asNMPC are identical to the ideal NMPC controller (i.e., without computational delay).
- Assuming that NLP (1) can be solved within one sampling time and that robust stability assumptions hold for ideal NMPC, then there exist bounds on output noise and model mismatch so that the resulting closed-loop system is input-to-state stable.

The local robustness property is encouraging, but to evolve to handle greater model mismatch and unmeasured disturbances, the optimization framework can be extended in two ways. First, we can embed robust observers and estimators (e.g., MHE, Extended Luenberger Observers, and Extended Kalman Filters) within (1) in order to update the NMPC model with output measurements and thereby reject uncertainty quickly. Recently, recent robust stability results [9] that that this embedding leads to input-to-state practical stability of the combined NMPC system. Second, we can extend (1) to perform robust optimization to determine a control sequence that is feasible for a broad instance of process models. This task can be captured by a multi-scenario formulation [8], with manipulated variables common to all uncertainty scenarios. Fortunately, due to intrinsic local robustness of NMPC, the multi-scenario formulation does not require a dense population of uncertainties.

3. Case Studies of asNMPC and asMHE

Two challenging process case studies are summarized in Table 1 to demonstrate the performance of these advanced step strategies. The first is a low density polyethylene (LDPE) process, which consists of the polymerization reactor and ancillary equipment (see [19]). In these reactors, high-pressure ethylene polymerizes over multiple reaction zones. Heat produced by polymerization is removed in cooling jackets and careful treatment of fouling layers is needed to prevent loss of production and eventual thermal runaway. The LDPE optimization model (1) consists of very large sets of partial DAEs that describe the evolution of the reactor mixture. After axial discretization, a typical LDPE reactor model can easily contain more than 10,000 DAEs. An initial study for a smaller model [18] demonstrated the application of asMHE, while a much larger model was used in [19] for NMPC. The fast performance for the large model was facilitated through a nested dissection of the linearized DAEs in each optimization iteration.

The second case study is an air separation unit (ASU) consisting of coupled low and high pressure columns for the production of high purity oxygen and nitrogen. The ASU distillation model contains MESH equations for 80 trays leading to 800 DAEs [7]. This process is subjected to 30-40% feed ramps along with 5% process noise. Its performance is significantly better than linear MPC and much more robust to noise and model mismatch. The results in Table 1 also show the multiscenario case where three instances of the ASU model are coupled to develop a robust NMPC strategy over a broad uncertainty range [8]. All of these cases consider very large NLPs and on-line sensitivity computation is at least two orders of magnitude less than background NLP.

Table 1. Case study results for asNMPC and asMHE

Case Study	# DAEs	#Vars./#Eqns.	Background CPU s.	On-line CPU s.
LDPE–asMHE [18]	294	27416/27121	202	0.9
LDPE–asNMPC [19]	~10000	243960/242850	120	0.33
ASU–asNMPC [7]	800	117140/116900	120 - 240	1.0
ASU–robust MS [8]	2400	350940/350700	1200-1800	2.0

4. Evolving to Dynamic Real-time Optimization

The previous section describes the enabling tools for the extension of NMPC to D-RTO. In particular, the advanced-step strategies allow us to now consider fast state estimations as well as control with large models. Nevertheless, widespread realization of D-RTO requires additional research in the following areas.

First, NMPC approaches need to be extended to *include economic objectives for D-RTO*. Industrial NMPC with economic objectives has already been realized by Odloak and co-workers at Petrobras [11, 13], and by Bartusiak and coworkers [1] at Exxon Chemicals. These indicate considerable improvement over tracking objectives. For instance, a recent D-RTO study [19] to maximize production in the LDPE process led to 3% LDPE production increase over the (optimal) tracking case. However, questions still remain in assessing stability and robustness with economic objectives. In particular, economic functions do not have properties of Lyapunov functions [5, 11] so additional analysis and modification of the NMPC strategy is needed to ensure stability and

robustness properties of D-RTO. A promising approach is the use of state regularization, similar to the schemes used in the IPOPT solver.

Second, the D-RTO systems require *efficient and robust state estimation schemes*. This can take the form of MHE, solved in parallel with the NMPC step or through embedding of fast observers or state estimators directly within the NMPC model. In either case, these formulations need to extended to reject measurement noise and process disturbances along with treatment of model mismatch. Analysis of stability and robustness properties of such combined schemes has seen recent activity, but still has many open questions (see [9]). In addition, reliable performance requires the incorporation of suitable noise models and extensions to consider fault detection and data reconciliation.

Finally, large integrated D-RTO formulations require the incorporation of *discrete decisions and hybrid elements.* A class of these elements has recently been explored through the application of complementarity conditions [2, 6], i.e.,

$$\min f(x, p_0),\ s.t.,\ c(x, p_0) = 0,\ x_l \le x \le x_u,\ \ 0 \le u \perp v \ge 0 \qquad (2)$$

where $x = [u^T, v^T, w^T]^T$ and (2) is a mathematical program with complementarity constraints (MPEC). Under mild assumptions, the (2) can be reformulated as an NLP and solved using the same NLP tools as in the NMPC strategies described above [2]. This allows many discrete decisions found in RTO models and scheduling tasks to be incorporated within the D-RTO through MPEC formulations.

5. Synthesis of Concepts and Future Directions

Based on the continuing advances described above, this concluding section explores a gentle evolution of current two-layer RTO strategies to a one-layer D-RTO strategy. Industrial successes with the one-layer strategy [1, 11, 13] have shown such a strategy to be successful. Moreover, as MPC loops are more widely applied and better understood than RTO implementations, it appears that a bottom-up D-RTO evolution should be quite suitable for many process systems, using the following guidelines.

- A natural starting point would consider large MIMO MPC applications (up to 50 inputs and outputs); many have already been implemented. These could be retrofitted to replace current linear state-space models by nonlinear DAEs that capture process behavior over a wider range. In addition, the tracking objective then evolves to an economic objective. This evolution leads to an approach analogous to [1].
- Additional robustness can be gained through the addition of embedded state estimators within the NMPC formulation, or through MHE models solved in parallel. The application of advanced-step strategies will allow these tasks to be calculated even for large models with negligible on-line computation.
- Ready extensions to multi-stage operations can be enabled through multiple models and complementarity constraints that handle model transitions and other discrete decisions. These models then evolve to handle additional dynamic features including the management of inventories and external demands, and dynamic integration with other processes. Many of these features cannot be handled by current two-layer RTO.

As a result, implementation o the above advances suggests that D-RTO can be within reach and can lead to significant benefits over current RTO strategies.

How far can D-RTO implementations go? Clearly this is motivated by the need to capture nonlinear dynamic behavior of challenging processes, and also to provide consistent models of process performance. In addition, there is a need to integrate control and RTO tasks with higher level planning and scheduling that can directly exploit the (optimized) performance capabilities of the plant [3, 10]. Finally, the integration of D-RTO over multiple plants requires further development of decentralized nonlinear control [12]. These remain as important and exciting challenges for the work ahead.

References

1. Bartusiak, R. D. (2007). NLMPC: A platform for optimal control of feed- or product flexible manufacturing. In Assessment and future directions of NMPC, Berlin: Springer.
2. Baumrucker, B., L.T. Biegler, (in press) MPEC Strategies for Optimization of Hybrid Dynamic Systems, J. Proc. Control
3. Biegler, L. T. and V. M. Zavala, (2009) " Large-scale Nonlinear Programming using IPOPT: An Integrating Framework for Enterprise-Wide Dynamic Optimization," Computers and Chemical Engineering, 33 , 575-582
4. Diehl, M., H. G. Bock, J. P. Schlöder (2005) Real-time iteration scheme for nonlinear optimization in optimal feedback control. SIAM J. Control Opt., 43:1714–1736.
5. Engell, S. (2007) Feedback control for optimal process operation. J. Proc. Cont., 17:203-219
6. Heemels, W., B. Brogliato (2003) The complementarity class of hybrid dynamical systems, European J. of Control, 9(2-3), 322-360
7. Huang, R., V. M. Zavala, L. T. Biegler, (in press) Advanced Step Nonlinear Model Predictive Control for Air Separation Units," Journal of Process Control
8. Huang, R., S. C. Patwardhan, L. T. Biegler, (2009) Multi-scenario-based robust nonlinear model predictive control with first principle dynamic models, this conference
9. Huang, R., S. C. Patwardhan, L. T. Biegler, (2009) Robust output-feedback nonlinear model predictive control using high gain observers, this conference
10. Kadam, J., W. Marquardt: Integration of economical optimization and control for intentionally transient process operation, In Assessment and Future Directions of Nonlinear Model Predictive Control, Springer Berlin, 419-434
11. Odloak, D., Zanin, A. C., Tvrzska de Gouvea, M. (2002). Integrating real-time optimization into the MPC of FCC system. Control Engineering Practice, 10, 819–831.
12. Rawlings, J. B., R. Amrit, (in press) Optimizing process economic performance using model predictive control, In Assessment and future directions of NMPC
13. Tvrzska de Gouvea, M., D. Odloak (1998). One-layer RTO of LPG production in the FCC unit, Comp. Chem. Eng., S22, 191-198
14. Wächter, A., L. T. Biegler (2006) On the Implementation of an Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming, Math. Prog., 106, 1, 25-57
15. Würth, L., R. Hannemann, W. Marquardt: An Efficient Strategy for Dynamic Real-Time Optimization based on Parametric Sensitivities In: IFAC World Congress, Seoul, Korea, 06-11.07.2008
16. Yip, W. S., Marlin, T. (2004). The effect of model fidelity on real-time optimization performance. Computers and Chemical Engineering, 28, 267.
17. Zavala, V., Biegler, L. T. (2009). The advanced-step NMPC controller: Optimality, stability and robustness. *Automatica*, 45, pp. 86-93
18. Zavala, V. M., Laird, C. D., Biegler, L. T. (2008). A fast moving horizon estimation algorithm based on nonlinear programming sensitivity. J. Proc. Cont., 18(9), 876–884.
19. Zavala, V. M., L.T.Biegler (in press) Optimization-Based Strategies for the Operation of Low-Density Polyethylene Tubular Reactors: NMPC, Comp. Chem. Engr.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modelling for PSE and Product-Process Design

Rafiqul Gani

CAPEC, Department of Chemical & Biochemical Engineering, Technical University of Denmark, Soltofts Plads, Building 229, 2800 Lyngby, Denmark

Abstract

This paper gives a perspective on modelling and the important role it has within PSE and product-process design. Different modelling issues related to development and application of systematic model-based solution approaches for product-process design is discussed and the need for a hybrid model-based framework is highlighted. This framework should be able to manage knowledge-data, models, and associated methods and tools integrated with design work-flows and data-flows for specific product-process design problems. In particular, the framework needs to manage models of different types, forms and complexity, together with their associated parameters. An example of a model-based system for design of chemicals based formulated products is also given.

Keywords: Modelling, Product-Process Design, Process Systems Engineering,

1. Introduction

Process systems engineering promotes the solution of problems in a systematic manner [1]. In a changing world, the topics covered within chemical engineering are also changing, influencing thereby, the scope and significance of process systems engineering [2, 3 & 4]. Therefore, although PSE has traditionally been applied by the chemical engineering community to solve problems for the oil, chemical and petrochemical industries, its potential application range is much wider. This is because the word “process” also implies, among others, the process of solving a problem; design of biochemical / biological processes for conversion of biomaterial to specific chemicals; and, the process of finding/designing chemicals with desired properties. That is, through the application of a systems approach, a wide range of problems associated with product-process engineering can be analyzed and solved.

In the area of product-process design, the problems differ in terms of the type of chemical(s) being produced. The products (and the processes that make them) from petrochemical and chemical industries are usually commodity chemicals, which could be classified as small and/or structurally simple molecules, produced in large amounts. In this case, process optimization in terms of operational efficiency and cost is usually a defining factor for a candidate product-process. The products (and the processes that make them) from life sciences, pharmaceutical, food and related industries, on the other hand, are usually large and/or complex molecules, produced in small amounts. Here, process optimization in terms of operational reliability and time of operation is usually a defining factor for a candidate product-process. This means that although the steps in the systematic solution of product-process design problems could be the same, the models and data, and the methods and tools that employ them in the various solution steps may be very different.

As illustrated in Fig. 1, models play a very important role in the systematic solution of product-process design problems. Various types of product-process design problems are listed on the right hand side of the figure. These problems are solved through corresponding methods of solution (algorithms) when enough knowledge and/or data are available. Since in most cases, the necessary knowledge and/or data may not be available, models are needed to supplement the available information. For example, models are needed to predict the behaviour of the product-process, to evaluate the performance of the product-process, to monitor and/or control the product-process, and many more. These models may be of different type (different types of equations are used to represent the system); scales (may involve sub-systems requiring different size and time scales); complexity (number of equations, degree of non-linearity, dimension, *etc.*) and simulation mode (steady state, dynamic, batch, identification, *etc.*).

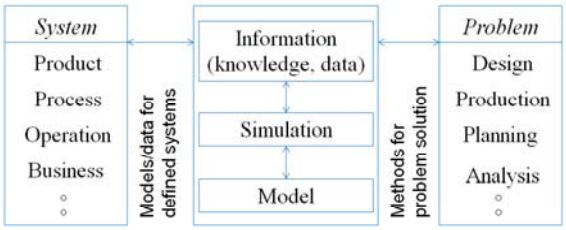


Figure 1: Need for models/data in problem solution

Product-process design in the life sciences, pharmaceutical, food and related industries, as opposed to the oil and petrochemical industries, is principally dependent on experiment-based trial & error approaches, providing thereby, opportunities for PSE to develop model-based solution approaches to reduce some of the experiments. Also, concerns related to increasing industrial activities accompanied by rapid depletion of resources and increase in pollution / waste, have provided opportunities for PSE to develop and implement issues such as energy management, supply chain and sustainability within an enlarged system boundary for product-process design problems. This means that solution approaches for product-process design need to address many inter-related issues, three of which are discussed below:

- Multi-scale: important data related to the chemicals may come from different sources, at different scales of time and size; for example, the properties that define the product characteristics could be based on the microstructure of the molecule or material, while the process behaviour that needs to be monitored and controlled during operation may be defined by the macroscopic (end-use) properties of the chemical system; the supply chain and sustainability issues need to be addressed at the mega-scale.
- Multidiscipline: the conversion, for example, of the biomaterial through biocatalysis requires knowledge of organic synthesis, enzymes, reaction catalysis, bioreactor design and operation – information about these topics come from different disciplines; sustainability analysis requires data, methods and tools from different disciplines.
- Integrated computer-aided techniques: lack of models to predict the behaviour of chemicals at different scales, of enzymes during organic synthesis, of reaction kinetics, *etc*., means that appropriate model-based computer aided techniques have not been developed and use of experiment-based techniques is the only option for these problems. Also, the multi-scalar and multi-

disciplinary nature of problems point to the need for integration of appropriate methods and tools.

In principle, product-process design (and development) needs to consider the demand for improved chemical-based products, made from more sustainable raw material resources and employing more efficient processes to make them. Therefore, methods and tools suitable for current and future product-process design need to manage a collection of sub-problems that require efficient and consistent handling of data and knowledge from different sources and at different time and size scales. A systems approach that can efficiently "manage the complexity" becomes therefore a very desirable option.

This paper provides a perspective on the opportunities for the development and use of hybrid model-based frameworks for systematic solution of chemical product-process design problems.

2. Managing the Complexity

The generic mathematical definition of the product-process design problem is given by,

$$F_{obj} = \{\boldsymbol{S}^T \underline{y} + f(\underline{x})\} \quad (1)$$

Subject to

$$\underline{D_1} = \underline{h_1}(\underline{x}, \underline{z}, \underline{p}) \quad (2)$$

$$\underline{D_2} = \underline{h_2}(\underline{x}, \underline{z}, \underline{p}) \quad (3)$$

$$0 \geq \underline{g_1}(\underline{x}, \underline{z}, \underline{p}, \underline{\theta}) \quad (4)$$

$$0 \geq g_2(\underline{x}, \underline{z}, \underline{p}, \underline{\theta}) \quad (5)$$

$$\boldsymbol{B}\underline{y} + \boldsymbol{C}\underline{x} \leq \underline{d} \quad (6)$$

In the above generic problem definition, the performance index (Eq. 1) needs to be minimized or maximized and may include multiple terms; the process model (Eq. 2) satisfies the conservation of mass and/or energy as a function of the product-process variables ($\underline{x}$), design variables ($\underline{z}$) and model parameters ($\underline{p}$); the product performance model (Eq. 3) predicts the behaviour of the product during its application; the product-process structure equations (Eq. 6) generate the feasible molecular-flowsheet structures as a function of decision (integer) variables ($\underline{y}$) and $\underline{x}$; and finally, the product-process constraints (Eqs. 4-5) define operational and/or chemical functional constraints ($\underline{\theta}$). Considering that the models maybe multi-scalar, non-linear and the variables involved may be integer as well as real, the generic problem defined above could represent a complex multi-dimensional problem of the MINLP type. Several variations of the above problem have also been formulated and solved in process optimization [3], heat-mass exchange networks [5], and in product-process design [6].

Before the generic product-process design problem defined above can be solved, several issues need to be addressed. For example, how to generate (or create) the product-process model objects for a large range of chemical systems? How to collect the necessary data for $\underline{z}$ and $\underline{p}$ for a wide range of models and systems? How to find a good initial estimate for $\underline{x}$ and $\underline{y}$ for different types of problems? One way to manage this complexity would be to provide a hybrid model-based framework for handling a diverse set of design work-flows corresponding to a wide range of problems, through an integrated computer-aided system. Such systems need to have a knowledge base of data, a library of models, a collection of algorithms (the work-flow and data-flow guiding the engineer/scientist through the solution steps), and, other associated methods-tools (such

as a tool to analyze data; a tool to create the missing model; a tool to screen feasible alternatives). An important question here is what is the role of models?

2.1. Model Role

The role of models within the context of product-process design is briefly discussed here. Should the model be used to simply replace the experiment, or, should the model play a more active role in finding truly innovative solutions? That is, should the model(s) be used mainly to verify a design (for given $\underline{y}$, $\underline{z}$, $\underline{p}$, $\underline{\theta}$, verify if the optimal $\underline{x}$ is obtained and if the value of the corresponding F_{obj} is really optimal?); or should they be used as part of a trial & error solution approach (assume values for $\underline{z}$, calculate $\underline{x}$ and $\underline{y}$, and, check if $\underline{\theta}$ is satisfied, check if F_{obj} is optimal - if not - repeat the procedure for another value of $\underline{z}$); or should the problem be decomposed into sub-problems and solved according to a derived hierarchy?

For the verification role, a validated model (in the form of a process simulator) performs the same function as experiments. The design decisions, are however, obtained separately (using the same models in a different role). For the trial & error solution approach, a model-based simulator simply replaces the experiment. While experiment based trial and error is sometimes the only alternative to solve a specific product-design problem, if the appropriate models are available, the search can be conducted over a larger space and the time and resources used to obtain a design can be significantly reduced. In this case, the experimental resources are reserved for the verification step in the final stages of the design process, rather than using them in the early stages of the design process. Here, the process model is used to define design targets ($\underline{x}^*$, $\underline{\theta}^*$) for the chemical product, while a collection of property models are used in the search for a match of the target. The method is called "reverse design" because first the optimal solution is located and from it, the design targets are defined. Then, the set of values for the design variables ($\underline{z}$) matching the targets (the optimal solution) is determined [7].

2.2. Model Generation

The application range of the hybrid model-based solution approach depends on the application range of the available models. Therefore, to achieve a wide search space during the early stages of the design process, the corresponding models need to be predictive by nature. In the later stages of the design process when quantitative values of the design variables are determined, the models need to be quantitatively correct but the application range does not need to be wide. In this respect, a multi-scale modelling scheme that can generate the necessary model(s) would be an interesting option, especially if the necessary model parameters ($\underline{p}$) can be predicted on-line without the need for additional experimental data. The main idea here is to use the same set of experimental data to regress model parameters at different scales. The models at the lower scales need less parameters to represent the same dataset and the model descriptors for the lower scales can be used to estimate the missing parameters in the adjacent higher scale. In this way, the predictive power of the model-based framework for product-process design is extended without the need for new data.

2.3. Problem Decomposition

The principal idea here is to decompose a complex problem into a set of sub-problems that are easier to solve and identify those that can be solved through model-based solution approaches. Solving these sub-problems according to a pre-determined sequence helps to reduce the search space through each subsequent sub-problem solution, until a sub-problem cannot be solved with models anymore. At this point, the

experiment-based trial and error approach takes over to determine the final solution. The advantage of this combined hybrid (systems) approach is that during the early stages, where enough data and models are available (or could be easily generated), the search space is rapidly reduced. In the later stages, where quantitative values become important and data/models become more unreliable, the experimental resources are employed, sometimes only to evaluate a few feasible alternatives to identify the truly innovative and best solution.

3. Example: The Virtual Product-Process Design Lab

Here, only one example of a model-based system, the *virtual* product-process design lab for design and analysis of various chemicals based formulated products, is briefly highlighted. Examples of other integrated systems managing the complexity can be found in [3, 4 & 9].

The idea behind the *virtual* product-process design lab is the following: instead of doing the experiments needed to search for a product and its process to manufacture it, the engineer / scientist performs *virtual* experiments, through the *v*PPD-lab software. The *forward* and the *reverse* solution approaches are available. The software contains a large knowledge base of data (of chemicals, of solvents, of plants, of microcapsule devices, *etc*.); a large collection of models (models for property prediction, models for controlled release, models for mixing, *etc*.); of design algorithms (methods for formulation design, methods for molecule design, methods for polymer design, methods for process flowsheet synthesis, *etc*); other tools (property prediction software; model generation software; equipment design software; design of experiments software, etc.). These are organized through a framework for efficient management of the complexity. Figure 2 gives an overview of the main features of *v*PPD-lab, which has been used in the design and evaluation of the controlled release of a drug active ingredient through a polymeric microcapsule, pesticide formulation, paint formulation, fuel-cells, and many more.

In the first step the problem is defined (for example, identity of the active ingredient; the desired controlled release parameters, *etc.*, are given in the "documentation" box of *v*PPD-lab). In the second step the selection of the application source (for example, codeine released into the body) and the primary properties of solvent and the polymer (needed by the controlled release model) are made. If the user is unable to provide this information, methods for solvent design and polymer design are used to generate a list of candidates to select from. In the next step the selection and calculation of the functional properties needed to evaluate the controlled release design is made (if models are not available, the modelling software helps to generate new models). In the next steps, the product performance model is used to predict the product behaviour. If the desired (target) performance is matched, then the last step of verifying the product performance through experiments is performed. If the target is not matched, it is possible to repeat from any of the earlier steps with a new design alternative.

Important issues to note from this example are that multi-scale models have been used, data and knowledge from different disciplines have been used and, design / evaluation problem has been effectively used by solving a collection of sub-problems according to a pre-determined sequence. The final step (not shown) would be to select a few of the alternatives and perform the necessary experiments to validate the selection. Therefore, the experiments are done not to design the product but to verify the product. This approach has the potential to save time and money in bringing a chemical based product to the market. Obviously, the accuracy and range of application of the *v*PPD-lab software depends on the available data and models in the software.

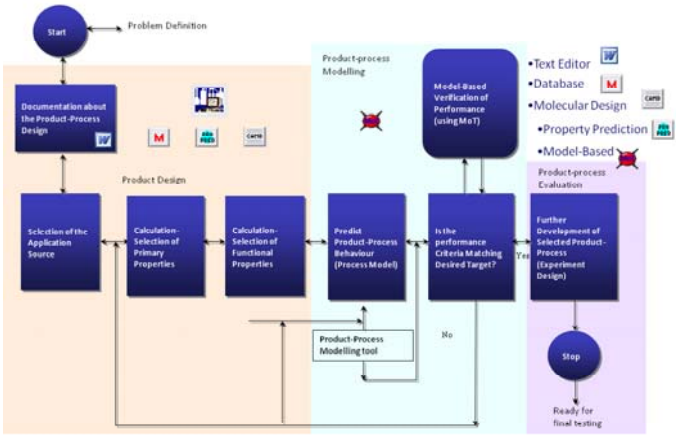


Figure 2: Virtual Product-Process Design Lab [10]

4. Conclusions and future work

The products of PSE are ideally suited to manage the complexity related to the analysis and solution of various types of product-process design problems. For the methods and tools to be able to solve a wide range of problems, models valid for a wide range of problems also need to be available. A systematic computer-aided modelling framework can play a major role in advancing the state of the art in PSE. The role of the model, however, needs to be clearly defined. This paper highlighted some of the opportunities and needs related to the development and use of systematic model-based solution approaches for chemical product-process design.

References

1. Takamatsu, T. The nature and role of process systems engineering. *Computers and Chemical Engineering*, 7, 1983, 203-218.
2. Gani, R. & Grossmann, I. E. Process Systems Engineering and CAPE – What Next? Computer Aided Chemical Engineering, 24, 2007, 1-5.
3. Grossmann, I. E. Enterprise-wide optimization: A new frontier in process systems engineering. *AIChE Journal*, 51, 2005, 1846-1857.
4. Klatt, K.-U. & Marquardt, W. Perspectives for process systems engineering – Personal views from academia and industry. *Computers and Chemical Engineering*, 33, 2009, 536-550.
5. Bagajewicz, M. J. A review of recent design procedures for water networks in refineries and process plants. *Computers and Chemical Engineering*, 24, 2000, 2113.
6. Karunanithi, A. T., Achenie, L. E. K. & Gani, R. A new decomposition based CAMD methodology for the design of optimal solvents and solvent mixtures. *Industrial and Engineering Chemistry Research*, 44, 2005, 4785-4797
7. Gani, R., Chemical product design: Challenges and opportunities. *Computers and Chemical Engineering*, 28, 2004, 2441-2457.
8. D'Anterroches, L. & Gani, R. Group contribution based process flowsheet synthesis, design and analysis. *Fluid Phase Equilibria*, 228-229, 2005, 141-146.
9. Singh, R., Gernaey, K. V. & Gani, R. Model-based computer-aided framework for design of process monitoring and analysis systems. *Computers and Chemical Engineering*, 33, 2009, 22-42.
10. Morales-Rodriguez, R. Computer-aided multiscale modelling for chemical product-process design. PhD-thesis, 2009, DTU Chemical Engineering, Lyngby, Denmark.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

WebLabs Dilemma in Chemical Engineering Teaching: Hindering or Promoting Creativity?

Roberto de Campos Giordano

Department of Chemical Engineering, UFSCar, Brazil
roberto@ufscar.br

Abstract

The universal spreading of links to the world net, in exponential growth, is an irreversible trend. Simultaneously, there is a continuous expansion of the accessibility to Internet 2 in Universities and Colleges of different countries, including Brazil. Ever-decreasing latency, bandwidth enhancement, a growing number of available tools and computational resources, all these trends are expected to change several aspects of the teaching/learning interaction.

The field of Chemical/Biological and Process/Bioprocess Engineering is certainly part of this process. And one important aspect in this area is the access to experimental set-ups, which are essential for development and consolidation of knowledge, assimilation of theoretical concepts, stimulation of critical reasoning and induction of cooperative work.

In this context, one important tool may be the use of WebLabs, remote-access laboratories available through the commercial or advanced Internet. Here we define a WebLab as a real experiment, properly automated and linked to the net, which may me operated from distance. Of course, virtual labs are important, too, and usually are part of the WebLab – allowing training, repetition when the apparatus is off-line, and so on. But one first point that we put forth is that the real experiment remains essential, at least in our field of knowledge.

The idea of using remote-access real experiments, connected to the internet, was brought about by MIT's iLab program, in mid-90's, focusing mainly on electronic, structural and chemical engineering experiments that would be available 24x7. Of course, within this conception, the experiment must be designed for stand-alone operation. Despite the breakthrough concept of opening didactic laboratories for www access, this approach provides limited space for a creative intervention from the students themselves.

Here, we discuss the feasibility of building a different methodology. The question is: can WebLabs be used as a tool that promotes creativity, and stimulates active or cooperative learning (both understood within the scope defined by the works of Felder, www4.ncsu.edu/unity/lockers/users/f/felder/public/, accessed in March 2009)?

Some real cases will be presented, involving cooperative Weblabs, that is, experiments performed collaboratively between two groups of undergraduate students, from two different Institutions, remotely joined in one team.

That was the result of a joint project, gathering four Chemical Engineering schools (USP, UFSCar, USP-RP and UNICAMP), funded by the São Paulo research agency FAPESP, within the Kyatera-TIDIA program (from the Portuguese acronym for Information Technology for the Development of the Advanced Internet).

Particularly, Universities in Brazil and France, and Universities from different regions in Brazil, have taken part in this pedagogical experiment. We believe this was a unique approach for the application of WebLabs in undergraduate Engineering education. Besides the improvement of cooperation abilities for hands-on problem solving, the students were exposed to different cultures, enhanced their communication skills and, last but not least, had a deeper insight of the fundamentals involved in the problem under study. Hence, this approach aimed to partially fulfill the lack of interaction in this level of engineering education, within an active learning environment.

To go further, in the direction of cooperative learning, the students should be involved in long-term projects and, under supervision, would have to understand the fundamentals of the problem on focus, to get acquainted with the experimental facilities that are available, design (all the group) and set-up (local fellows) the experiment, run assays, analyze results, go back to the lab... and write reports. The continuous contact with the remote fellows would emulate the real-world professional environment of the modern global economy, where remote interaction between professionals in different countries, working on the same project, becomes evermore frequent.

In conclusion, the point we make here is that it is possible to involve mixed groups of students of different Institutions, even from different countries, in open experiments, demanding initiative and teamwork. But this will only happen if professors and lecturers of both Institutions are committed, also as a team, to establishing new teaching methodologies. However, one might ask: are they available for this kind of work nowadays?

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Research Challenges in Planning and Scheduling for Enterprise-wide Optimization of Process Industries

Ignacio E. Grossmann

Center for Advanced Process Decision-making,Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

Enterprise-wide optimization (EWO) is a new emerging area that lies at the interface of chemical engineering and operations research, and has become a major goal in the process industries due to the increasing pressures for remaining competitive in the global marketplace. EWO involves optimizing the operations of supply, production and distribution activities of a company to reduce costs and inventories. A major focus in EWO is the optimization of manufacturing plants, which often requires the use of nonlinear process models. Major operational items include production planning, scheduling, and control. This paper provides an overview of major challenges in the development of deterministic and stochastic linear/nonlinear optimization models for planning and scheduling for the optimization of entire supply chains that are involved in EWO problems.

Keywords: supply chain optimization, planning, scheduling, mixed-integer programming, stochastic programming.

1. Introduction

The goal of this paper is to provide a general overview of the role of planning and scheduling in Enterprise-wide Optimization as applied to the process industry. This paper is largely a follow-up of the ideas, concepts of our previous paper in the area (Grossmann, 2005). It has also been motivated by our experiences in the Enterprise-wide Optimization project of the Center of Advanced Process Decision-making at Carnegie Mellon University (http://egon.cheme.cmu.edu/ewocp/) which involves faculty from Carnegie Mellon, Lehigh and University of Pittsburgh in collaboration with ABB, Air Products, BP, Dow Chemical, ExxonMobil, Nova Chemicals, PPG, Praxair and Total.

Enterprise-wide optimization (EWO) has become a major goal in this industry due to the increasing pressure for remaining competitive in the global marketplace (Grossmann, 2005; Varma et al, 2007). EWO involves optimizing the operations of R&D, supply, production and distribution to reduce costs and inventories. A major focus in EWO is the scheduling of production facilities, as well as their modeling at the proper level of detail, often requiring nonlinear process models. Major operational items include planning, scheduling, real-time optimization and control. One of the key features in EWO is integration of the information and decision-making among the various functions that comprise the supply chain of the company. This is being achieved with modern IT tools. To fully realize the potential of transactional IT tools, the development of deterministic and stochastic linear and nonlinear optimization models and algorithms (analytical IT tools) is needed to explore and analyze alternatives of the

supply chain to yield overall optimum economic performance, as well as high level of customer satisfaction. An additional challenge is the integrated and coordinated decision-making across the various functions in a company (purchasing, production, distribution, sales), across various geographically distributed organizations (vendors, facilities and markets), and across various levels of decision-making (strategic, tactical and operational).

As an example of a large-scale EWO problem consider in Fig. 1 the supply chain in the petroleum industry which comprises many intermediate steps starting from the exploration phase at the wellhead, going through trading and transportation, before reaching the refinery, and finally the distribution and delivery of its products at the retail level (e.g. gasoline). In this case it is clear that the effective coordination of the various stages is essential to accomplish the goal of EWO. As another example, Fig. 2 shows the supply chain in the pharmaceutical industry that starts with the R&D phase for testing of new drugs, a major bottleneck that must be eventually coupled with batch manufacturing and global distribution for drugs that achieve certification.

From the two examples, we can define Enterprise-wide Optimization as optimizing the operations of R&D, material supply, production, distribution and financial activities of a company to reduce costs and inventories, and to maximize profits, asset utilization, responsiveness and customer satisfaction. The goal of achieving Enterprise-wide Optimization in the two examples is clearly still elusive and motivates the research challenges outlined in the paper.

Fig. 1. Supply chain in the petroleum industry (courtesy ExxonMobil).

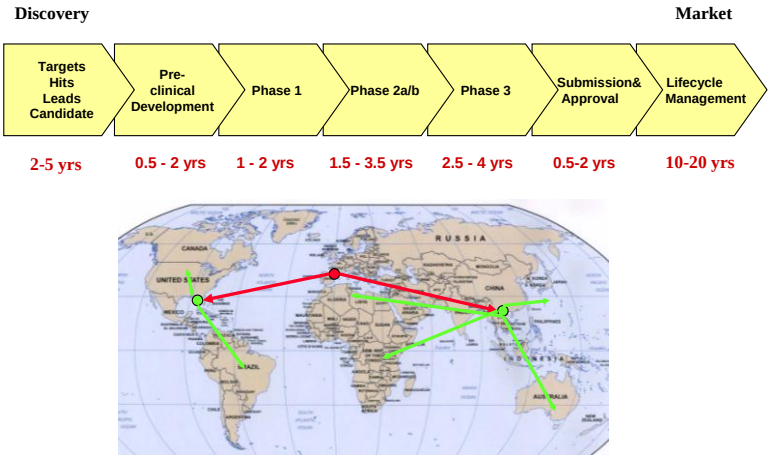


Fig. 2. R&D coupled with manufacturing and distribution in the supply chain of the pharmaceutical industry

One of the key features in EWO is integration of the information and decision making among the various functions that comprise the supply chain of the company. Integration of information is being achieved with modern IT tools such as SAP and Oracle that allow the sharing and instantaneous flow of information along the various organizations in a company (Fig. 3). While these systems have become quite powerful, they do not provide comprehensive decision making capabilities that account for complex trade-offs and interactions across the various functions, subsystems and levels of decision making. Therefore, companies are faced with the problem of deciding as to whether to develop their own in-house tools for integration, or else make use of commercial software from vendors. While progress has been made in developing generic commercial tools, the great diversity of application areas as well as the integration that is required across geographical and time scales for EWO continues to be a major barrier for the effective deployment of these commercial tools. Therefore, currently there is still need for significant in-house development in order to exploit the particular features of the problem at hand so as to produce effective, practical useful tools.

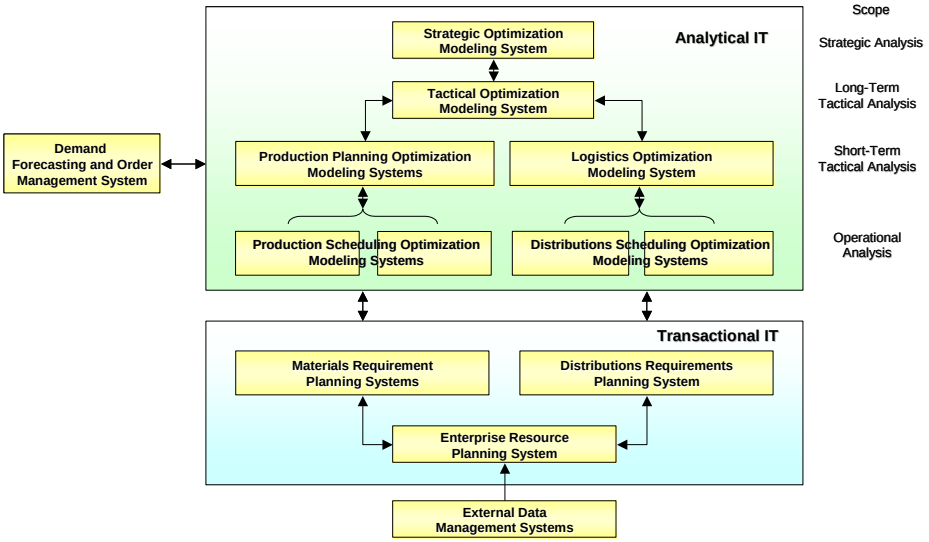


Fig. 3. Transactional and Analytical IT (Tayur et al., 1999).

2. Challenges in Enterprise-wide Optimization

A major challenge that is involved in EWO of process industries is the integrated and coordinated decision-making across the various functions in a company (purchasing, production, distribution, sales), across various geographically distributed organizations (vendors, facilities and markets), and across various levels of decision-making (strategic, tactical and operational) as seen in Figure 3 (Shapiro, 2001). The first two items conceptually deal with issues related to *spatial integration* in that they involve coordinating the activities of the various subsystems of an enterprise. The third item deals with issues related to *temporal integration* in that they involve coordinating decisions across different time scales. Addressing these spatial and temporal integration problems is important because they provide a basis to optimize the decision-making in an enterprise through the IT infrastructure.

In order to achieve EWO throughout the process industry, this goal will require a new generation of computational tools for which three major challenges must be addressed:

a) *The modeling challenge:* What type of production planning and scheduling models should be developed for the various components of the supply chain, including *nonlinear* production processes, that through integration can ultimately achieve enterprise-wide optimization? Major issues here are the development of novel mathematical programming and logic-based models that can be effectively integrated to capture the complexity of the various operations.

b) *The multi-scale optimization challenge:* How to coordinate the optimization of these models over a given time horizon (from weeks to years), and how to coordinate the long-term strategic decisions (years) related to sourcing and investment, with the medium-term decisions (months) related to tactical decisions of production planning and material flow, and with the short-term operational decisions (weeks, days) related to scheduling and control? Major issues involve decomposition procedures that can effectively work across large spatial and temporal scales.

c) *The uncertainty challenge*: How to account for stochastic variations in order to effectively handle the effect of uncertainties (e.g. demands, equipment breakdown)? Major issues here are the development of novel, meaningful and effective stochastic programming tools.

As described in Grossmann (2005), there are also many algorithmic challenges, particularly for solving mixed-integer linear and nonlinear programming problems. For a recent review on MILP models and algorithms that are relevant to this area see Pochet and Wolsey (2007).

The modeling challenge. While the area of planning and scheduling has seen the development of many models in Operations Research (OR) (e.g. Pinedo, 2001), over the last decade a significant number of planning and scheduling models have been proposed in process systems engineering specifically for process applications (Kallrath, 2002). In contrast to general OR scheduling models, the process-oriented models may require the use of material flows and balances, and very often network topologies that are quite different from the more traditional serial and multistage systems. Furthermore, they address both batch and continuous processes, and may require the use of detailed nonlinear process models. One of the most important issues in the development of planning and scheduling models is time representation, namely discrete and continuous time representations. Discrete time models, which often have tighter relaxations, have become more competitive since it is increasingly possible to solve problems with smaller time intervals due to great advances in MILP optimization. Nevertheless, continuous time models are still required, especially when sequence dependent changeovers are involved. Recent reviews on discrete and continuous time scheduling models can be found in Floudas and Lin (2004) and Mendez et al, 2006.

The multi-scale optimization challenge. Planning and scheduling of process systems are closely linked activities since they deal with the allocation of available resources over time to perform a collection of tasks required to manufacture one or several products (Bodington, 1995; Kallrath, 2002). The aim in planning is to determine high

level decisions such as production levels and product inventories for given marketing forecasts and demands over a long time horizon (e.g. months to years). Scheduling, on the other hand, is defined over a short time horizon (e.g. days to weeks) and involves lower level decisions such as the sequence and detailed timing in which various products should be processed at each equipment in order to meet the production goals set by the planning problem. Integration and coordination of planning and scheduling are key components in EWO. Conceptually, the simplest alternative for solving planning and scheduling problems is to formulate a single simultaneous planning and scheduling model that spans the entire planning horizon of interest. However, the limitation of this approach is that when typical planning horizons are considered, the size of this detailed model becomes intractable due to the potential exponential increase in the computation. The traditional strategy for solving planning and scheduling problems is to follow a hierarchical approach in which the planning problem is solved first to define the production targets. The scheduling problem is solved next to meet these targets (Bodington, 1995; Shapiro, 2001). The problem of this approach, however, is that a solution determined at the planning level does not necessarily lead to feasible schedules. These infeasibilities may arise because the effects of changeovers are neglected at the planning level, thereby producing optimistic targets that cannot be met at the scheduling level. Therefore, methods and approaches are being developed that can more effectively integrate planning and scheduling (e.g. Dimitriadis et al, 1997; Erdirik-Dogan and Grossmann, 2006; Sung and Marvelias, 2007). On the other hand, supply chain optimization (Neiro and Pinto, 2004; Shah, 2005) is another multiscale optimization problem where the dominant issue is the handling the spatial and temporal scales, which require the use of decomposition (e.g. Jackson and Grossmann, 2003).

The Uncertainty Challenge. Uncertainty is a critical issue in EWO. Furthermore, it is complicated by the fact that the nature of the uncertainties can be quite different in the various levels of the decision making (e.g. strategic planning vs. short term scheduling vs. control). Research in the process industry has largely focused on strategic uncertainty (e.g. Ahmed and Sahinidis, 1998; Goel et al., 2006). Handling uncertainty in inventory management (e.g. Zipkin, 2000) is only beginning to be considered in process systems engineering (e.g. You and Grossmann, 2008). The same applies at the tactical level for production planning with uncertain demand (Gupta and Maranas, 2003; Sand and Engell, 2004). The reason for this situation is that the corresponding deterministic optimization problems are difficult to solve, and hence the resulting optimization problems under uncertainty with probability distribution functions become extremely complex. A comprehensive review on methods for addressing optimization under uncertainty can be found in Sahinidis (2004).

2. Concluding Remarks

This paper has provided a brief overview of three major research challenges in the emerging area of Enterprise-wide Optimization: modeling of planning and scheduling, multiscale optimization and handling of uncertainties. From this overview it is clear that further research is required to expand the scope and size of production planning and scheduling models that can be solved in order to achieve the goal of EWO. Major challenges remain in the development of novel models for integrating, planning, scheduling and control. Of particular significance is the development of effective decomposition schemes that have the capability of handling large-scale problems over

geographically distributed sites and over wide temporal scales. There is also an increasing need for the effective global optimization of nonconvex planning and scheduling problems that incorporate process models. The development of effective and meaningful solution methods for stochastic programming problems for handling uncertainties continues to be of paramount importance. It is worth noting that all these developments will require a synergy of advances in basic algorithms and advances in large-scale computation. These efforts should help to expand the scope and nature of EWO models that can be effectively solved for real-world industrial problems.

References

Ahmed, S. and N. V. Sahinidis, Robust Process Planning under Uncertainty, *Industrial &Engineering Chemistry Research*, **37**(5), 1883-1892, 1998.

Bodington, E. C., Planning, Scheduling and Control: Integration in the Process Industries. McGrawHill, 1995.

Dimitriadis, A.D., Shah, N. and Pantelides, C.C., "RTN-Based Rolling Horizon Algorithms for Medium Term Scheduling of Multipurpose Plants," *Computers & Chemical Engineering*, **21**, S1061 (1997).

Erdirik-Dogan, M..; Grossmann I.E., "A Decomposition Method for the Simultaneous Planning and Scheduling of Single-Stage Continuous Multiproduct Plants," *Ind. Eng.Chem. Res*., **45**, 299-315, (2006).

Floudas, C.A.; Lin, X. "Continuous-time Versus Discrete-time Approaches for Scheduling of Chemical Processes: A Review." *Comp. and Chem. Eng.*, **28**, 2109 – 2129 (2004).

Goel, V., I. E. Grossmann, A.S. El-Bakry and E.L. Mulkay, "A Novel Branch and Bound Algorithm for Optimal Development of Gas Fields under Uncertainty in Reserves," *Computers and Chemical Engineering*, **30**, 1076-1092 (2006).

Grossmann, I.E., "Enterprise-wide Optimization: A New Frontier in Process Systems Engineering," *AIChE J*., 2005, **51**, 1846-1857.

Gupta, A. and Maranas, C.D. "Managing Demand Uncertainty in Supply Chain Planning," *Computers and Chemical Engineering*, **27**, 1219-1227 (2003).

Jackson, J. and Grossmann, I.E., "A Temporal Decomposition Scheme for Nonlinear Multisite Production Planning and Distribution Models," *I&EC Research*, **42**, 3045-3055 (2003).

Kallrath, J. (2002). Planning and Scheduling in the Process Industry. *OR Spectrum*, **24**, 219 – 250.

Mendez, C.A., J. Cerdá , I. E. Grossmann, I. Harjunkoski, and M. Fahl, "State-Of-The-Art Review of Optimization Methods for Short-Term Scheduling of Batch Processes," *Computers & Chemical Engineering* **30**, 913-946 (2006).

Neiro, S.M.S. and Pinto, J.M., "Supply Chain Optimization of Petroleum Refinery Complexes," *Computers and Chemical Engineering* **28**, 871-896 (2004).

Pinedo, M., Scheduling: Theory, Algorithms, and Systems. Prentice Hall (2001).

Pochet, Y. and L. Wolsey, "Production Planning by Mixed Integer Programming," Springer (2007).

Sand, G. and S. Engell: Modelling and Solving Real-time Scheduling Problems by Stochastic Integer Programming. *Computers and Chemical Engineering* **28**, 1087-1103 (2004).

Shah, N. Process Industry Supply Chains: Advances and Challenges. *Comp Chemical Engineering*, **29**, 1225-1235 (2005).

Shapiro, J.F., *Modeling the Supply Chain*, Duxbury, Pacific Grove (2001).

Sung, C.; Maravelias, C. T. An Attainable Region Approach for Effective Production Planning of Multi-product Processes. *AIChE J*., **53** (5), 1298-1315, 2007.

Tayur, S., Ganeshan, R. and Magazine, M. (Eds.), *Quantitative Models for Supply Chain Management*, Kluwer Academic, Norwell, Massachusetts (1999).

Varma, V. A.; Reklaitis, G. V.; Blau, G. E.; Pekny, J. F., Enterprise-wide Modeling & Optimization – An Overview of Emerging Research Challenges and Oopportunities. *Computers & Chemical Engineering* **31**, 692 (2007).

You, F. and I. E. Grossmann, "Mixed-Integer Nonlinear Programming Models and Algorithms for Large-Scale Supply Chain Design with Stochastic Inventory Management", *I&EC Research* **47**, 7802-7817 (2008).

Zipkin, P. H., *Foundations of Inventory Management*, McGraw-Hill (2000).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Production Scheduling in the Process Industries: Current trends, emerging challenges and opportunities

Gabriela P. Henning

INTEC (UNL-CONICET), Güemes 3450, Santa Fe, CP 3000, Argentina

Abstract

This paper discusses the current trends and defies associated with production scheduling, both from an industrial and academic perspective. First, the new challenges that appear in the context of globalized and competitive economies are addressed. They stem from the need of considering scheduling as a building block of the Advanced Planning Systems (APSs) that nowadays participate in Supply Chain Management (SCM) functions. They are primarily associated with business process coordination and information integration requirements. Then, main features, strengths and limitations of current academic proposals are briefly addressed. Finally, some of the reasons for the slow acceptance and modest penetration of these research results are highlighted. Thus, challenges and opportunities to be faced in order to alleviate the miscommunication of the academic and industrial worlds are pointed out.

Keywords: Process Scheduling and Operations, Advanced Planning Systems, Supply Chain Management, Business Process Coordination and Integration

1. Introduction

Nowadays, most companies experience a growing international competition and recognize the needs of supply chain (SC) efficiency and responsiveness. These requirements originate from the pressures of shortened product life-cycles and increasing customer demands on a great variety of low-price, high quality products. On top of these needs, there are also higher expectations on accurate deliveries, with short lead-times and always rising customer service levels. The growth in globalization, and the additional management challenges it brings, has motivated interest in global Supply Chain Management (SCM). Successful SCM requires a change from managing individual functions to integrating activities into key SC business processes.

In order to compete, companies must be more efficient, transparent, and agile. This requires real-time visibility into their business operations and the ability to react if needed. To achieve this goal, Operational Intelligence (OI) focuses on providing on-line monitoring of business processes and activities as they are executed, in assisting them by identifying and detecting situations which correspond to interruptions, inefficiencies, threats, bottlenecks, etc., and communicating them to relevant stakeholders.

Since many process industries try to adopt SCM and OI philosophies, this introduces new pressures into the way they perform their scheduling activities. The trend is to adopt scheduling systems (SSs) which no longer are employed as isolated support applications. On the contrary, they are perceived as integrated building blocks of these advanced management tools. Thus, new tendencies and problems have appeared. Modern SSs are nowadays part of the Advanced Planning Systems (APSs) that industry

is quickly adopting (Stadtler and Kilger, 2005), and new functional and non-functional requirements have arisen for them. These new requisites are discussed in Section 2, along with more traditional industrial demands not covered yet in most SSs.
On the other hand, from the academic side, there has been enormous progress on methodologies to tackle scheduling problems in the last two decades. A brief overview of their strengths and limitations, as well as current trends, is presented in Section 3. Despite the flourish of approaches, a dichotomy between the industrial and academic worlds is still evident. Requirements that SSs need to face in a competitive and globalized context, as well as capabilities and characteristics that academic proposals need to incorporate to be accepted by the industrial community, are discussed in Section 4, which summarizes emerging challenges and opportunities.

2. Industrial practice: Activities in the context of Advanced Planning, Commercial Scheduling Systems and OI applications

As organizations strive to focus on core competencies, they have increased the manufacturing outsourcing of some components and intermediate products and have reduced the ownership of raw materials sources and distribution channels. Thus, the number of different partners along extended SCs has progressively grown. This led in most cases to complex planning situations having multiple products, manufactured at multiple work centers in multi-site production systems, which are transported by third or fourth party logistics. Under these circumstances, scheduling is crucial for achieving the timely and cost-effective execution of industrial production processes. However, the pressures being faced now are much higher than those identified by Shah (1998).
In the context of the burdens that globalization and networked economies pose, advanced concepts on supply chain management (SCM) have emerged. According to Stadtler and Kilger (2005), the goals of SCM can be achieved by resting on two pillars: the integration of the organizational network and the coordination of material, information, and financial flows. In turn, the coordination pillar comprises several building blocks: proper use of information and communication technology, process orientation, and advanced planning. All of them are quite relevant to define the context in which a scheduling system must operate and to specify the requirements it must fulfill.
Advanced Planning includes long, mid and short-term planning levels. Software products, referred as Advanced Planning Systems, are now available to support these tasks. APSs aim at complementing existing ERP systems and overcoming some of their weaknesses. ERPs have traditionally addressed planning tasks poorly and have focused on a single company. On the contrary, modern APSs are being designed for inter-organizational SCs. APSs try to find feasible, near optimal plans across the SC as a whole, while potential bottlenecks are considered explicitly (Stadtler and Kilger, 2005). In terms of software components, an APS means a broad group of integrated software applications developed by various software vendors, such as i2 Technologies, Manugistics, Oracle, SAP, AspenTech, etc. Fig. 1 shows the structure of a generic APS.
For the sake of brevity all of these systems are not discussed in detail; however, some advanced ones, such as Aspen Plant Scheduler (http://www.aspentech.com) or SAP Advanced Planner and Optimizer (SAP-APO) (http:// http://help.sap.com) are worth to be mentioned, since they exhibit a high degree of integration of their scheduling component with the other planning modules comprising the system. Regarding solution methodologies included in the scheduling element, the first of these systems uses decision rules and heuristics, while SAP-APO employs the ILOG´s optimizer library, as

other vendors, but pre-defined model formulations are not made completely explicit. SAP APO offers several solution approaches, among them mixed-integer linear programming (MILP), constraint programming (CP) models and evolutionary algorithms.

Embedded optimization models seem to be a problem associated with APSs since explicit, high level representations are not available yet. On the contrary, SAP APO as well as other ASP providers, have coded their models in C or C++. As suggested by Kallrath and Maindl (2006), this fact has several consequences: (i) Before starting an implementation of an APS, a thorough investigation of the applicability of the vendor's models is necessary, and (ii) knowledgeable users in modeling techniques are required to embed pre-defined SAP-APO models, as well as "own" optimization models. The users (consultants, schedulers) not only would need to understand the underlying scheduling approach, but also know how to make it interoperate with the transactional and master data that the APS handles.

Despite advances in the integration of the various planning functions, APSs do not seem to have specific tools to address reactive scheduling problems. To properly tackle them, besides having specific solution approaches, the integration of the APS with the Manufacturing Execution System (MES) is required. As pointed out by Harjunkoski et al. (2008), this integration is mandatory for the correct management of manufacturing operations and to perform control activities. Such integration should rely on the pillars of standard representations (ANSI ISA-95, ISA-88), business process modeling and service oriented architectures.

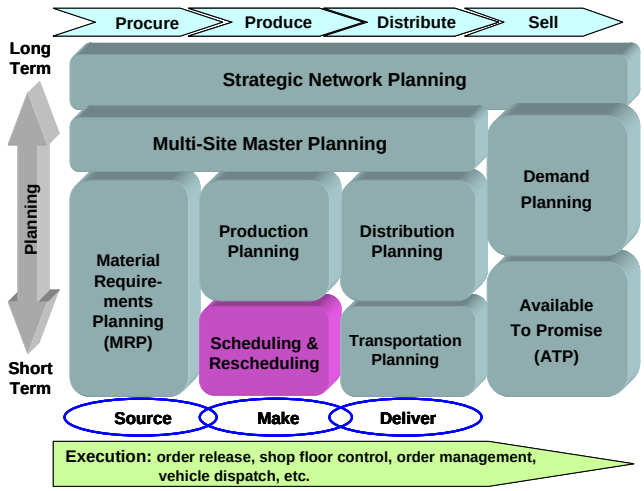


Fig.1. Components of a Generic APS

In addition to APSs, many commercial SSs have matured over the last decade. Over 100 software applications of Finite Capacity Scheduling (FCS) systems are available according to APICS (http://www.apics.org). However, the use of these systems is not widely spread yet. Among them is worth to mention some applications that are academic spin-offs. For instance, applications like VirtECS Schedule from Advanced Process Combinatorics (http://www.combination.com) and SchedulePro from Intelligen (http://www.intelligen.com) have powerful graphical capabilities and include interactive scheduling tools (ISTs), features that are always demanded in the industrial world.

However, the integration of these systems, similarly to other commercial packages, with the organization planning modules and the MES, does not seem straightforward.
Finally, it should be mentioned that most APSs and commercial SSs do not have explicit support for basic OI functionalities, like process traceability as well as the automatic and seamless generation of metrics and key performance indicators.

3. Research Efforts: Strengths and limitations of the academic proposals

Nowadays, the academic community is able to address more difficult and larger scheduling problems than those tackled twenty years ago. The most significant research efforts are highlighted in some excellent reviews, like the ones of Li and Ierapetritou (2008), Méndez et al. (2006), Floudas and Lin (2004), Kallrath (2002), Pekny and Reklaitis (1998), Pinto and Grossmann (1998) and Shah (1998). Usually, problems are solved by resorting to MILP formulations. Optimal solutions are many times reached for small and medium size problems, with reasonable computational efforts. This success basically comes from the remarkable advances in modeling techniques, algorithmic solutions and computing technologies that occurred in the last decade.
One of the key elements of the various MILP approaches involves time representation; current formulations can be classified into discrete-time and continuous-time models. Other essential issues are related to the characteristics of the production environment being considered: sequential or network process, operation modes, storage policies, type of critical resources, changeovers, etc. This wide scope of features has led to a variety of formulations that are generally oriented to properly address particular classes of problems. Despite recent advances, there is no general formulation that can tackle the various classes of problems with the same efficiency. For instance, it is not clear the extent to which methods aimed at addressing complex network processes, can also be successfully applied to other structures commonly found in industry, such as the multistage one. Efficiency is also highly dependent on the adopted objective function; however, no extensive evaluations have been done in this respect. Nowadays, nobody denies that schedule measurement is a critical aspect of the scheduling problem, which is multiobjective, multiattribute in nature. Schedules mean different things to different people and are used by various organizational groups in distinct ways. Even having a single scheduler involved in the evaluation process, the question of how to deal with the tradeoffs between the different metrics of interest is hard to address without keeping the human involved (Kempf et al., 2000). Few advances in this respect have been made by the academic world.
Formulations based on CP have also gained interest recently (Maravelias and Grossmann, 2004; Roe et al., 2005). The PSE community has started to adopt CP, mainly in combination with MILP models, because these techniques are complementary to each other when a cost or profit related objective function is chosen. Nevertheless, this hybrid approach has a poor computational performance when other objective functions related to time are tackled.
An analysis of current proposals shows that they are becoming more mature and progress is made along various research lines. Some works have improved the representations from a computational point of view (Janak and Floudas, 2008). Others have concentrated on the expressive power of the model, to be able to capture more realistic features (Giménez et al., 2009). Moreover, recent proposals have extended previous formulations to move the frontiers of the problems being tackled, thus addressing the integration of scheduling and design (Castro et al., 2005), planning and scheduling (Maravelias and Sung, 2008), of mid and short-term scheduling (Janak et al.,

2008a), etc. In fact, there is a new tendency to generalize basic formulations to address a wider scope of problems. Fortunately, the latest reactive scheduling proposals (Vin and Ierapetritou, 2000; Méndez and Cerdá, 2004; Janak et al., 2008b, etc.) follow this trend and share the basic representation with their predictive scheduling counterparts, facilitating the integration of predictive and reactive scheduling components.

4. Emerging challenges and opportunities

As seen, the industrial and academic worlds have given considerable attention to the scheduling problem and enormous advances have been achieved. Due to lack of space certain important problem features, such as uncertainty, or some solution techniques, as metaheuristics and evolutionary approaches, have not been treated in this work.

Despite this progress, most of the theoretical research done during the last 15 years is of limited use in the real world. In fact, while researchers provide answers to toy cases, practitioners have difficulties in explaining their needs and exploiting the opportunities that could come out from the utilization of advanced scheduling approaches. Naturally, this academic-industrial dichotomy has several causes. Some are pointed out below:

- Currently, a SS is one of the blocks of an APS, which in turn is part of a SCM system. This fact imposes requisites of applications' integration and business process coordination. On the other hand, academic proposals have focused on algorithms and computational performance; thus, they lack a business process view and do not have an explicit domain representation, which would facilitate integration. For instance, most MILP approaches rely on very sophisticated formulations (grounded on STN and RTN models, and artificial notions like time points, events and time grids) that do not have a direct mapping to ordinary problem elements, as they are perceived by industrial practitioners and to master data, as handled by industrial information systems. In consequence, to bring in industry representations which are oriented towards efficient problem solving, there is a need to revalorize the design of interfaces (CAPE-OPEN, 2001), as well as the development of ontologies and meta-languages.
- The role of the human scheduler needs to be stressed at the various problem solving stages: Problem definition, solution process and results analysis (Weirs, 1997). Schedulers must be able to manipulate, customize and tailor models/solution approaches as part of their day-to-day work. This manipulation should be done by means of high level constructs and visual languages. Visual modeling is nowadays used for many purposes in several industrial sectors, but not in the scheduling field yet. Besides, schedulers need to achieve independence from IT and domain specialists. Modeling should become almost as natural as sketching and scribbling. Moreover, the use of high level graphical tools for solution analysis must also be a routine practice. Finally, cooperative work among the various stakeholders must also be supported.
- Academic scheduling approaches need to be comprehensively evaluated. The current focus on just computational performance and solution quality needs to expanded to cover other features like ease of implementation and usability, robustness, extensibility, etc. (Pekny and Reklaitis, 1998). The adoption of a scheduling benchmarking service (Cavalieri et al., 2007) and the definition of new industrial-size benchmark problems (Kallrath, 2002) would certainly help in this respect.

By addressing the two first issues, an extra-added value will be obtained. The number of errors attributed to miscommunication between various organizational areas would drop due to the seamless integration of enterprise data into the SS and the proper dissemination of scheduling results. Similarly, the number of errors attributed to the scheduler would decrease substantially with his/her full articulation within the SS.

References

CAPE-OPEN, 2001, Open Interface Specification: Planning and Scheduling Interface.

P. Castro, A. P. Barbosa-Póvoa, A. Q. Novais, 2005, Simultaneous design and scheduling of multipurpose plants using resource task network based continuous-time formulations. Ind. Eng. Chem. Res., 44, 343-357.

S. Cavalieri, S. Terzi, M. Macchi, 2007, A Benchmarking Service for the evaluation and comparison of scheduling techniques, Computers in Industry, 58, 656-666.

C. Floudas, X. Lin, 2004, Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review. Computers and Chemical Engineering, 28, 2109–2129.

D. M. Giménez, G. P. Henning, C. T. Maravelias, 2009, A novel network-based continuous-time representation for process scheduling: Part I. Main concepts and mathematical formulation. Computers and Chemical Engineering. http://dx.doi.org/10.1016/j.compchemeng.2009.03.007

I. Harjunkoski, R. Nyström, A. Horch, 2008, Integration of Scheduling and Control – Theory or Practice? In Proceedings: Foundations of Computer-aided Process Operations, Boston, MA.

S. L., Janak, C. A. Floudas, J. Kallrath, N. Vormbrock, 2006a, Production Scheduling of a Large-Scale Industrial Batch Plant. I. Short-Term and Medium-Term Scheduling. Ind. Eng. Chem. Res., 25, 8234-8252.

S. L., Janak, C. A. Floudas, J. Kallrath, N. Vormbrock, 2006b, Production Scheduling of a Large-Scale Industrial Batch Plant. II. Reactive Scheduling. Ind. Eng. Chem. Res., 25, 8253-8269.

S. L. Janak, C. A. Floudas, 2008, Improving unit-specific event based continuous-time approaches for batch processes: Integrality gap and task splitting. Computers and Chemical Engineering, 32, 913-955.

J. Kallrath, 2002, Planning and scheduling in the process industry. OR Spectrum, 24, 219–250.

J. Kallrath, T. I. Maindl, 2006, Real Optimization with SAP APO, Springer.

K. Kempf, R. Uzsoy, S. Smith, K. Gary, 2000, Evaluation and comparison of production schedules, Computers in Industry, 42, 203–220.

Z. Li, M. Ierapetritou, 2008, Process Scheduling under uncertainty: Review and challenges, Computers and Chemical Engineering, 32, 715-727.

C. T. Maravelias, I. E. Grossmann, 2004, A Hybrid MILP/CP Decomposition Approach for the Continuous Time Scheduling of Multipurpose Batch Plants. Computers and Chemical Engineering, 28, 1921-1949.

C.T. Maravelias, C. Sung, 2008, Integration of production planning and scheduling: Overview, Challenges and Opportunities. In Proceedings: Foundations of Computer-aided Process Operations, Boston, MA.

C. A. Méndez, J. Cerdá, 2004, An MILP framework for batch reactive scheduling with limited discrete resources. Computers and Chemical Engineering, 28, 1059–1068.

C. A. Méndez, J. Cerda, I. E. Grossmann, I. Harjunkoski, M. Fahl, 2006, State-of-the-art Review of Optimization Methods for Short-term Scheduling of Batch Processes. Computers and Chemical Engineering, 30, 913-946.

J.F. Pekny, G. V. Reklaitis, 1998, Towards the convergence of theory and practice: A technology guide for scheduling/planning methodology. Proceedings Third Intern. Conf. on Foundations of Computer-Aided Process Operations, J. F. Pekny and G. E. Blau (Eds.), 91–111.

J. Pinto, I. E. Grossmann, 1998, Assignments and sequencing models of the scheduling of process systems. Annals of Operations Research, 81, 433–466.

B. Roe, L.G. Papageorgiou, N. Shah, 2005, A hybrid MILP/CLP algorithm for multipurpose batch process scheduling. Computers and Chemical Engineering, 29, 1277-1291.

N. Shah, 1998. Single- and multisite planning and scheduling: Current status and future challenges. Proceedings Third Intern. Conf. on Foundations of Computer-Aided Process Operations, J. F. Pekny and G. E. Blau (Eds.), 75–90.

H. Stadtler, C. Kilger, 2005, Supply chain management and advanced planning: concepts, models, software and case studies, Springer.

J. P. Vin, M. G. Ierapetritou, 2000, A new approach for efficient rescheduling of multiproduct batch plants. Ind. Eng. Chem. Res., 39, 4228-4238.

V. Wiers, 1997, Human-computer interaction in production scheduling. Analysis and design of decision support systems for production scheduling tasks. Ph.D. Thesis. T.U. Eindhoven.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Chemical Logistics – Going Beyond Intra-Plant Excellence

I. A. Karimi

Department of Chemical & Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117576

Abstract

While intra-plant logistics is critical to the effective function of a chemical plant, extra-plant logistics is pivotal to the survival and competence of a chemical supply chain. For years, the chemical companies overlooked the latter, assuming it just a fixed cost. Growing pressures of competition, globalization, integration, outsourcing, and resilience are making a sea change. Most chemical companies are now discovering that logistics is more than just a line-item charge. Producers and (service) providers alike are taking a fresh look at the art, science, and cost of chemical logistics. While the logistics for dry goods and merchandise has received the lion's share of attention in the open literature, the same for chemicals has not. In the pursuit of lean, integrated, and excellent enterprises and supply chains, some PSE researchers have rightly begun to study a variety of unaddressed issues and challenges in chemical logistics. This paper describes selected novel applications of PSE in maritime chemical logistics.

Keywords: Chemical logistics, supply chain management, enterprise-wide optimization, maritime transport.

1. Introduction

The Council of Supply Chain Management Professionals defines logistics as the part of SCM that plans, implements, and controls the efficient, effective, forward, and reverse flow, and storage of goods, services, and related information between the point of origin and the point of consumption in order to meet customers' requirements. Transportation moves goods and people, but logistics is much more than transportation. Over the last four decades, it has evolved through the integration of traditional tasks of materials management and physical distribution. It includes materials management (procurement, receipt, storage, handling, sales, delivery, disposal, cleaning, training, emergency response, labeling ...), warehousing, demand forecasting, purchasing and supplier management, industrial packaging and assembly, transportation, freight payment, transshipment, order management (documentation, customs clearance, brokerage …), production planning, inventory management, reverse logistics, etc. After years of optimizing intra-plant logistics, the chemical industry has rightly begun to focus (Karimi et al., 2002) on extra-plant (enterprise-wide or supply chain) logistics as a strategic rather than tactical issue.

Due to its global nature, the chemical industry globally sources, not just raw materials and catalysts, but also many other materials, equipment, and services, such as processing units, instruments, indirect materials (stationery and supplies), safety equipment, brokerage, order management, supplier management, customs and documentation, etc. that are crucial for its day-to-day operations. With its huge production levels and the necessity of moving hazardous materials in a safe manner, it is

not surprising that the chemicals-related industries rank the highest among all sectors in supply chain costs (Braithwaite, 2002). According to Mark Kaiser, the CEO of Cendian Corporation (Hoffman, 2002), "the US$1.5 trillion chemical industry spends US$160 billion annually on logistics, and has among the highest average supply chain costs (12% of revenues, compared to 10% for pharmaceutical companies and 9% for automotive manufacturers)." In 2004, the North American chemical producers spent US$33.2 billion transporting chemicals worth US$516.2 billion (Morris, 2006). However, when we consider the logistics costs as fractions of the net value-add, the figures for the chemicals-related industries are much higher than the other asset-intensive sectors. For instance, they are 43% for petroleum, 37% for chemicals, 36% for food and beverages, and 30% for paper, as compared to 28% for the automobile industry. While this is partly attributable to the need for maintaining high safety standards in moving hazardous chemicals, rising fuel prices and increasing regulations are placing even bigger demands on the logistics costs in the chemical industry. A study [ACHEMA, 2006] by EPCA (European Petrochemical Association) claims that an improvement of around 3-5% in the supply chain would be enough to boost noticeably the competitive standing of chemical companies. Clearly, efficient and cost-effective management of chemical logistics is crucial to the current and future financial success of global chemical supply chains. Whether it is a chemical company that owns and manages its logistics operations, or a shipping or third party logistics provider (3PL) that serves the chemical company, the ultimate cost of logistics has a significant impact on the bottom-lines of a chemical supply chain and its various entities.

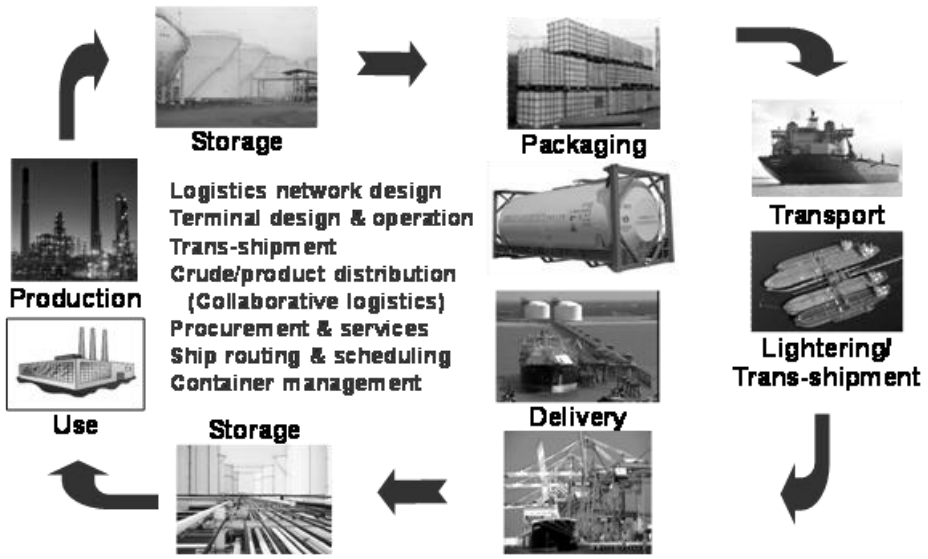


Fig. 1: Overview and scope of chemical logistics

Figure 1 gives an overview of logistics in a generic chemical supply chain. For years, most major chemical companies considered logistics as a core competency. Even now, many own or lease ships, tankers, tank trucks, rail cars, tank containers, etc. and run their own logistics functions. However, the growing trend is towards focusing on the core competency of chemical manufacturing, and whether to in-source or out-source logistics partially or fully has become a major issue in chemical supply chain

management. While transportation, warehousing, and customs brokerage top the list of most outsourced activities in general, many others such as transshipment, trans-loading, storage, distribution, packaging, freight forwarding, track-n-trace, invoice and payment services, scheduling, sourcing, inventory management, sales & distribution, customs documentation, packaging, blending, labeling, marking, drumming, break-bulk, etc. are increasingly being outsourced by chemical companies.

Chemical logistics employs a wide mix of means such as pipelines, rail, trucks, ships, barges, tankers, and planes. Pipelines are a common and important option in the chemical industry due to their low cost and higher capacity. A total of 265,440 km of oil and natural gas pipelines (as at the end of June 2008) exists in 37 Asian countries, of which roughly 45 percent carry oil/refined products/condensate (Venkatesh, 2008). Even in other regions such as South America, multiproduct pipelines transport a variety of petroleum products. This has led several PSE researchers (Rejowski & Pinto, 2008; Cafaro & Cerda, 2008; Relvas et al., 2009) to address the challenges in distributing petroleum products via multi-product pipelines. While transport via pipelines is quite a natural option for Chemical Engineer, it is not always possible or economical. For instance, rail and truck transports dominate Europe and North America. However, we are not aware of any work on the issues related to these in the PSE community. For products such as crude oil and LNG (Liquefied Natural Gas) with a few major sources, maritime logistics is preferred, as pipelines become uneconomical for transport over long distances. Maritime transport is the undisputed leader for Asian and increasingly global chemical industry. While land transport modes such as pipelines, rail, and road are important in themselves, and critical partners for maritime logistics, we focus mainly on maritime logistics due to the space constraints and the fact that maritime logistics is the workhorse of increasingly global chemical industry. Due to the lack of space, we confine our discussion to maritime transportation only, although PSE work exists on issues in other areas such as storage (Tay et al., 2005), network design (Bansal et al., 2008) etc.

2. Marine Transportation

Marine transportation in general has attracted the interest (Christiansen et al., 2004) of academic researchers since the 1950s. The following describes some recent work on marine transportation via various carriers suitable for a spectrum of chemical materials such as crude carriers for crude; LNG carriers for LNG; chemical tankers for many organic, inorganic, and petrochemical products; small parcel tankers for regional transport; and container ships for solvents, vegetable oils, fruit juices, etc. Lastly, we touch on the very timely issue of bunker fuel planning in view of the recent hikes in fuel prices.

2.1. Crude Transportation

Fully loaded crude oil tankers such as Very large crude carrier (VLCC) and Ultra Large Crude Carrier (ULCC) cannot pass through shallow channels or dock at shallow ports due to shallow drafts, narrow entrances, or small berths. Such tankers are “lightered” by small vessels at offshore deep-sea to enable their entry into shallow channels or ports. Tanker lightering involves the direct ship-to-ship transfer of crude oil from large tankers to small vessels. It is a critical step in the petroleum supply chain and a special case of the general transshipment described later.

Cheng & Duran (2004) addressed the global maritime transport of crude oil. Huang & Karimi (2006b) developed models that address a general and realistic form of the tanker-lightering problem with several realistic and practical features ignored by

previous work (Lin et al., 2003). These features include multi-compartment service vessels picking up different crude parcels during one voyage and making multiple visits to different VLCCs during one voyage, options of selecting crudes to lighter, accounting for the impact of crude densities, demurrage and time-charter costs, etc.

2.2. LNG Transportation

LNG is the most economical way of transporting natural gas over long distances. Being a cryogen, its temperature must be maintained at –163 °C during transportation by letting small amounts of LNG boil-off. To keep the ship tanks cold during the return voyage, a small amount of LNG is left to boil in the tank. This is known as heel. Minimizing boil-off losses during the entire round-trip and during loading and unloading is important. Recently, Hasan et al. (2009) performed rigorous, realistic, detailed, and extensive dynamic simulations of boil-off during various steps of LNG transportation, and studied the effects of various factors such as nitrogen content, tank pressure, ambient temperature, voyage length, etc. on boil-off losses. From their simulations, they determined the optimal heels for several scenarios of lean LNG transportation. They showed that heel can be reduced by up to 40% for a typical long voyage of 20 days, compared to the usual industrial practice of 5% of the cargo, and the reduction is significantly more for shorter voyages. Their results suggest savings of millions of dollars from heel optimization in LNG transportation.

2.3. Routing and Scheduling of Chemical Tankers

A chemical or parcel tanker employs multiple tanks to hold multiple liquid chemical cargoes unmixed. Efficient cargo assignment, routing and scheduling of parcel tankers are crucial to the financial success of a tanker company. An optimal assignment of cargos and schedules to a fleet of carriers is a complex combinatorial problem. It is accentuated by the need to comply with two safety regulations concerning cargo-tank and cargo-cargo compatibilities. Nevertheless, many small to medium-size shipping companies still route and schedule their ships manually. Hence, the potential for improving the scheduling process in maritime transportation is considerable.

Jetlund & Karimi (2004) developed a maximum-profit model for routing and scheduling parcel tankers engaged in the shipping of bulk liquid chemicals with cargo pickup time-windows (or pickup laycans). Their model can decide the most profitable spot cargos, their allocations to various ships, and voyage details of a fleet of ships. They proposed a heuristic decomposition algorithm to obtain the fleet schedule by repeatedly solving the base model for a single ship. Neo et al. (2006) presented a rudimentary approach to ensure ship stability by imposing limits on net longitudinal and lateral moments of each vessel. Their new model is an extension of the single-ship model of Jetlund & Karimi (2004) and involves deciding which ports should a ship visit and in which sequence, when and which cargoes it should pickup and unload, when and which tanks each cargo should use, etc. over the entire trip to maximize the profit. While their model does address some important safety issues, it is far from being practical, and still does not address factors such as fuel costs, delivery laycans, etc. These considerations, as discussed below, are extremely important in chemical business and we have developed a heuristic methodology (Oh, 2008) for accommodating these very practical features.

2.4. Chemical Transshipment

The major players in the ocean shipping industry normally operate fleets of large or deep-sea carriers with a focus on deep-sea trading. Over the years, there is a growing need for supplying and distributing chemicals to and from the industries within regions that large ships cannot serve. The increase in deep-sea and short-sea shipping activities globally and the myriad of mergers, acquisitions, and collaboration are increasing the

demand for trans-shipment operations that involve direct ship-to-ship transfers of chemicals from large to small ships in deep sea.
Huang & Karimi (2006a) introduced a scheduling model for a general trans-shipment scenario for regional distribution, where multiple large donor-carriers trans-ship bulk liquid cargos to multiple small recipient-carriers at a trans-shipment location. Their model determines the sequence in, the sides (larboard or starboard) from, and the times at which, each recipient ship should receive cargos to minimize the total time-charter costs of all ships. Compared to the manual procedures used in practice for such problems, their models could reduce the total operation costs. In a simple case study conducted by them, the costs reduced by 6.3%, which could be much higher depending on the scenario and its complexity.

2.5. Tank Container Management

Iso-tank containers carried by container ships are normally used to transport small quantities of liquid chemicals. These are cylindrical tanks set inside the frames of standard dry containers. Mismatch in product supply and demand results in an imbalance in the container flows across different regions. Empty containers accumulating at the demand centers must be "repositioned" to the production centers. Karimi et al. (2005) reported a comprehensive study of this important short-term tank-container management problem. They used an innovative, event-based, "pull" approach to develop a novel linear model for the minimum-cost or maximum-profit scheduling of the transport and cleaning of multiproduct tank containers (loaded and empty) given a set of projected shipment orders in the short-term. The authors illustrated the application potential of their models by using it to solve large and industrially relevant problems with key practical considerations such as alternate ship schedules, delivery time windows, and inter-modal transport routes.

2.6. Bunker Fuel Planning

Fuel expenses can be up to 90% of the daily operating cost of a tanker. Recent high fuel prices forced some companies to slow ships to save fuel. A sound management of vessel's fuel consumption is crucial to the profitability of tanker owners. However, bunker fuel prices are highly unpredictable and can exhibit significant variations across refueling ports. Furthermore, the fuel consumption rates of tankers are nonlinear functions of voyage speeds. Therefore, decisions on refueling and vessel speed require more than the experience and judgment of individuals. Oh (2008) has formulated this problem as a mixed-integer linear programming model with key operational constraints and a scenario-based approach for uncertain fuel prices. The constraints relate to pickup time windows, fuel level, and tonnage limits. The MILP model determines the optimal refueling ports, amounts, and probabilistic distribution for voyage profit. Oh (2008) has illustrated the approach on a realistic scenario with one tanker, ten port visits, three speed options, three potential refueling options at the end of each port visit, and 30 discrete scenarios of unit fuel prices for each refueling option. The stochastic MILP involved 1,597 continuous variables, 630 binary variables, 3,408 constraints, 18,514 nonzeros, and required 48 s of CPU time using GAMS/CPLEX on a Pentium 4 Xeon (2.8 GHz, 2 GB RAM) workstation. Although this model is intended for tankers for bulk maritime transport, it is useful for other vessel types including container ships, reefers, etc.

3. Conclusion

Chemical logistics beyond the plant borders is an important frontier in chemical supply chain management. The PSE community has the right tools, methods, and domain

knowledge to advance this relatively unexplored front. With increasing global chemical trade and corresponding logistics needs, the time is only ripe for us to address the practical needs of various stakeholders such as port operators, logistics service providers, freight forwarders, and chemical companies.

4. Acknowledgement

The author wishes to acknowledge partial financial support for this work from the Maritime Port Authority of Singapore (MPA) and National University of Singapore.

References

ACHEMA, 2006, Opportunities for Europe Trend Report 10: Supply Chain Management, 28th International Exhibition-Congress on Chemical Engineering, Environmental Protection and Biotechnology, Frankfurt, Germany, May 15-19.

M. Bansal,I. A. Karimi, R. Srinivasan, 2008, Outsourcing and Selection of Third-Party Logistics Contracts for Chemical Companies, Ind. Eng. Chem. Res., 47, 8301–8316.

A. Braithwaite, 2002, Achieving World Class Supply Chain and Logistics in the Chemical Industry, EPCA Logistics Meeting, Monaco, November.

D. C. Cafaro, & J. Cerdá, 2008, Dynamic scheduling of multiproduct pipelines with multiple delivery due dates. Comp. Chem. Eng., 32, 728–753.

L. F. Cheng, & M. A. Duran, 2004, Logistics for Worldwide Crude Oil Transportation using Discrete Event Simulation and Optimal Control, Comp. Chem. Eng., 28, 6-7, 897-911.

M. Christiansen, K. Fagerholt, & D. Ronen, 2004, Ship Routing & Scheduling: Status and Perspectives, Transportation Science, 38, 1, 1-18.

M. M. F. Hasan, A. M. Zheng, & I. A. Karimi, 2009, Minimizing oil-off losses in liquefied natural gas transportation, Ind. Eng. Chem. Res., in press.

W. Hoffman, 2002, Better Chemistry through Technology, Supply Chain Tech News, May.

I. A. Karimi, M. Sharafali, & H. Mahalingam, 2005, Scheduling Tank Container Movements for Chemical Logistics", AIChE J., 51, 1, 178-197.

I. A. Karimi, R. Srinivasan, & L. H. Por, 2002, Unlock Supply Chain Improvements through Effective logistics, Chem. Eng. Prog., 98, 5, 32-38.

X. Lin,E. D. Chajakis,& C. A. Floudas, Scheduling of tanker lightering via a novel continuous-time optimization framework, Ind. Eng. Chem. Res., 42, 4441.

Gregory D. L. Morris, 2006, Chemical Producers Put New Emphasis on Logistics Safety and Security, Inbound Logistics, June, 60-64.

K-H. Neo, H-C. Oh, & I. A. Karimi, 2006. Routing and Cargo Allocation Planning of Parcel Tanker. Presented at 16th European Symposium on Computer Aided Process Engineering and 9th International Symposium on Process Systems Engineering, July 9-13, Garmisch-Partenkirchen, Germany.

H-C. Oh, 2008, PhD Thesis, Dept of Chem & Biomol Eng, National University of Singapore.

H-C. Oh, & I. A. Karimi, 2007, Routing and Scheduling of Parcel Tankers: State of the Art and Opportunities. Presented at 2nd International Maritime-Port Technology and Development Conference, 26 to 28 September, Singapore.

R. Rejowski, Jr., & J. M. Pinto, 2008, A continuous time representation for scheduling of pipeline systems with pumping yield rate constraints. Comp. Chem. Eng., 32, 1042–1066.

S. Relvas, A.P.F.D. Barbosa-Povoa, & H. A. Matos,2009, Heuristic batch sequencing on a multi-product oil distribution system, Comp. Chem. Eng., 33, 712-730.

H. L. Tay, I. A. Karimi, C. K. Peck, and X. L. Peh, 2005, Contract Selection and Tank Allocation in a Terminaling and Storage Facility, Ind. Eng. Chem. Res., 44, 7435-7450.

G. Venkatesh, 2008, Oil & Gas Logistics, Logistics Insight Asia (logasiamag.com), November.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Systems Problems in Biorenewables Processing

Wolfgang Marquardt

AVT Process Systems Engineering, RWTH Aachen University, 52056 Aachen, Germany

Abstract

Different scenarios have been published in recent years which predict the depletion of fossil carbon resources for the production of fuels and chemicals in face of the demand of a growing world population. The switch from fossil to biorenewable carbon feedstocks is inevitable. This switch of feedstock will have massive impact on the chemicals and fuels value chain and the associated production technologies. This short paper will introduce some of the relevant systems engineering challenges to be tackled in the next decades. Furthermore, the objectives and the research strategy of a large collaborative research effort at RWTH Aachen University will be reviewed.

Keywords: Biomass, bioeconomy, biorefinery, biobased industrial products, biofuels.

1. Introduction

In the last few years several studies have been published which discuss different scenarios regarding the possible depletion of global fossil carbon resources comprising petroleum, natural gas and coal. An in-depth analysis of the current situation has recently been carried out by Pfennig and co-workers (Fayazz et al., 2009). These authors have formulated a coarse dynamic model with global coverage to balance the various raw material and energy flows. The model covers the known and the predicted amounts of fossil carbon resources on the one and the predicted consumption rate of the growing world population on the other hand. One objective of the model is to forecast the remaining availability of the three types of fossil carbon resources. According to this study, petroleum, natural gas and coal are expected to be depleted by roughly 2035, 2045 and 2070, respectively. One may debate the validity of this forecast given the unavoidable uncertainty in the data used in the model. However, the study clearly shows that a shift from fossil to alternative carbon resources for the production of industrial products is inevitable. Furthermore, serious research efforts have to start now given the scientific and technological challenges associated with such a major change which will affect all industrial segments.

There are only two alternatives to fossil carbon resources, carbon dioxide (CO_2) in the atmosphere and land-based or aquatic biomass, which are both available in abundance and which are continuously replaced by nature as long as solar energy is available on planet earth. Both carbon resources require novel processing technologies to convert them into useful industrial products such as fuels, chemicals and polymer materials. Despite some success with respect to promising catalytic pathways in the recent past (Sakakura et al., 2007), the direct conversion of CO_2 into chemicals and fuels is still economically not viable because of the low energy level of the molecule which results in high energy demand for its chemical conversion. Hence, the focus of most research programs is rather on biomass as a carbon feedstock.

Only biomass which does neither compete with the food chain nor result in a carbon depth (Fargione et al., 2008) should be considered. This requirement can only be

fulfilled if either waste biomass (e.g. straw, forestry residues, sewage sludge, organic waste from animal farming, organic municipal solid waste, etc.) or whole plants are used as feedstock rather than only (the often edible) parts of the plant like seeds or nuts as in the production of first-generation biofuels such as biodiesel and bioethanol. Alternatively, algae could be cultivated in open or closed aquatic systems to provide the necessary carbon source (e.g. Wackett (2008) for their use in fuels production) to avoid competition with the land-based production of agricultural food products. In this contribution, biorenewable carbon resources are consequently understood to consist of either waste biomass, lignocellulosic biomass or of aqueous algae suspensions.

2. Challenges and opportunities of biorenewables processing

There are few major established alternative routes for the processing of biorenewables (e.g. Huber et al., 2006) including (i) biomass gasification to synthesis gas and subsequent water-gas shift to hydrogen, Fischer-Tropsch synthesis to alkanes, or syntheses to methanol and other platform chemicals and (ii) liquid-phase processing starting with some pretreatment of the biomass (Sierra et al., 2008) followed by either enzymatic hydrolysis and conversion (Stephanopoulos, 2007, Wackett, 2008) or by chemical transformations (Corma et al., 2007) to fuel components and chemicals.

While gasification and subsequent synthesis gas conversion has the advantage of just replacing the very early steps from carbon raw materials to industrial products by novel processes but keeping the rest of the chemical supply chain intact, it does not fully exploit the synthesis power of nature. In particular, complex and highly structured biomolecules present in any biorenwables feedstock are decomposed to the smallest building blocks possible (i.e. H_2 and CO) during biomass gasification which are then assembled again to form functional molecules by means of a series of subsequent chemical transformations. Obviously, from a thermodynamics point of view it would be more favorable to directly isolate suitable functional building blocks from the biorenewable raw materials and to selectively refunctionalize them rather than first destroying the rich native molecular structures. Such an approach builds on low-temperature liquid-phase processing where particulate biomass is dissolved or dispersed in a solvent first and then transformed into products by bio- or chemocatalysis or a combination of both.

The deoxygenation of the oxygen-rich biomass can either lead to CO_2 or H_2O as reaction by-products. Obviously, the latter is preferred because it does not contribute to global warming and the resulting climate change. However, this approach to deoxygenation requires large amounts of hydrogen, which has to be produced in a sustainable way with minimum CO_2 footprint. Electrical water splitting using solar or wind energy as the primary energy source is presently considered to be a viable alternative. This strategy is favored over thermochemical water splitting requiring high temperature heat which can only be provided by solar or nuclear power plants.

While biomass gasification largely builds on existing process technology and targets at the established product families, liquid-phase biomass conversion requires novel processing technologies and might even aim at novel product families. In contrast to state-of-the-art carbon feedstock processing, liquid-phase reaction systems will be preferred over gas-phase reaction systems and low-temperature separations such as membrane processes, extraction and crystallization will be preferred over high temperature separations such as distillation. The viscosity of the biomass-derived liquid systems is typically very high. The raw material streams (derived from native biomass after pretreatment) show a much higher chemical diversity than petroleum or natural

gas. This is due to the variety of biomass feedstocks and the multitude of organic and anorganic chemical species the biomass consists of. Water or other solvents will be required in large amounts for liquid-phase processing. Hence, solvent recycling and management will constitute a major effort in addition to molecular transformation and product isolation and purification. Novel products with molecular structures different from the ones preferred at present and close to the native molecular structure prevalent in biomass feedstock will be derived according to users' specifications.

This short list suggests that the manufacturing technology of nowadays petroleum-based chemical supply chains will eventually be replaced by novel technologies which are tailored to the particular needs of biorenewables processing involving a variety of raw materials, alternative product families, novel reaction pathways and catalysts, different conversion processes and intensified processing equipment. Such a paradigm shift constitutes a massive scientific and technological challenge and presents a tremendous opportunity to replace current technology by sustainable production systems.

3. The systems nature of biorenewables processing

The short discussion of the various alternatives, challenges and opportunities reveal the systems nature of biorenewables processing. Any process systems engineering problem (Klatt & Marquardt, 2009) involves the representation, the analysis and the design of artifacts which comprise an often large-scale real-world entity, the process system, which is composed of interconnected subsystems showing numerous interactions. The complexity can only be dealt with, if abstraction across various hierarchical layers and suitable decomposition strategies are pursued. System design problems are often ill-defined because the objectives are numerous and often in conflict. Furthermore, the design objectives typically evolve with time during the design process. All these features are present in biorenewables processing.

We have argued that the ultimate replacement of fossil carbon raw materials by biorenewables suggests or even requires the reinvention of the chemis-tree and the chemical supply chain from raw materials to products. The system to be dealt with spans from raw materials to products and their use in diverse applications. Hence, biorenewables processing can be considered an integrated decision making problem which involves not only (i) process design – one of the classical problems in process systems engineering – but also (ii) product design (Gani, 2004), deciding on the molecular structure which is at the same time close to the structures present in biorenewable raw materials but still reflecting the product properties specified by the end user, (iii) the selection or even the design of chemical reaction pathways linking products and raw materials by a number of molecular transformation steps, and (iv) the selection of even the design of raw material streams with close links to plant sciences and biomass production. This decision making problem will be further explained and illustrated using biofuels as an example (cf. Fig. 1).

The replacement of petroleum-based fuels should target at a high quality designer biofuel with most favorable product properties. In particular, close-to-zero particulate and NO_x emissions and the efficient conversion of the chemical energy into mechanical energy are highly desirable. Obviously, this objective links the fuel properties not only with the combustion process but also with engine design. Most likely, the ideal future fuel will not consist of an ill-defined mixture of many chemical components as today but of a blend of carefully selected chemicals to meet all the requirements of the combustion process. The design of such a fuel product is therefore inevitably intertwined with the design of the combustion process, which itself is intertwined with

the design of the engine and even of the complete vehicle drive train. Hence, we have identified a first product-process design problem of a complex nature where the process refers to the "use process" rather than the "manufacturing process" of the product.

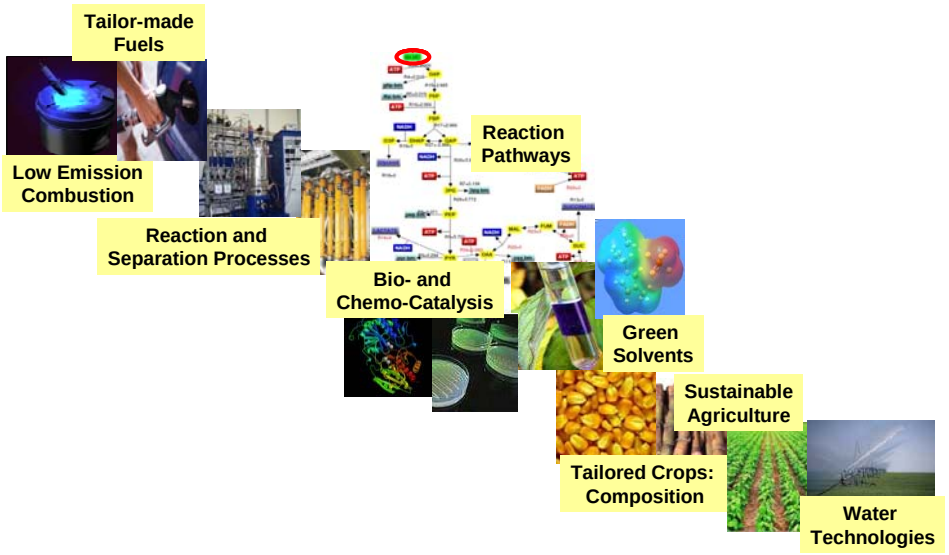


Fig. 1: Biofuels processing – a set of interacting systems problems

The components of the future biofuel have to be produced in a sustainable manner reconciling economical, ecological and societal needs. In particular, the carbon efficiency has to be maximized: the amount of CO_2 released in the whole supply chain from the production of the biomass to the wheel of the vehicle should be as close as possible to the amount inevitably produced in the combustion process.

The requirement for high carbon efficiency can not only be tackled by an integrated process design involving intensified equipment to minimize raw material consumption and energy demand (Klatt & Marquardt, 2009, van Gervent & Stankiewicz, 2009). Rather, the biofuel production process will rely on a novel chemistry to facilitate selective catalytic conversion of raw materials into platform chemicals by refunctionalization of the molecular structures present in the native biomass (Corma et al., 2008). New reaction pathways with highly selective catalytic transformation steps relying on novel catalytic systems and possibly non-classical green solvents such as supercritical fluids or ionic liquids besides water are required in this case. The selection of an appropriate reaction pathway together with suitable solvent systems minimizing solvent replacement steps is definitely another systems problem. It is determined by many factors including the available raw material components to start the molecular transformations with, the thermodynamics of the reaction steps involved and the availability and cost of highly selective bio- or chemo-catalysts which allow for highly selective conversion in every single transformation of the inevitably multi-step reaction network.

The type of raw materials and their chemical composition are by no means a given. The type of plant, the soil and climate of the geographical region, the cultivation and even harvesting conditions determine the chemical composition and the meso-scale structure of the biomass to a large extent. Therefore, the biomass feedstock can be tailored to some extent by genetic engineering and by the biomass production process.

Obviously, there is also a close link to intensified biomass production, using either land-based agricultural or aquatic cultivation systems including biodiversity requirements. Water and soil management are key issues for the sustainability of the intensified biomass production.
Last but not least, the low volumetric energy content of biomass requires sophisticated raw materials procurement and fuel production and distribution networks to minimize the amount of energy required to satisfy the needs of the logistics and shipping processes. Partly decentralized production plants are an attractive alternative for the preprocessing of the biomass feedstock into platform chemicals. These intermediates can then be shipped to large-scale fuels processing plants with a distribution network to the filling stations. Alternatively, on board processing of the platform chemicals into fuel components may evolve into an attractive alternative.
This discussion of the biofuels case clearly shows the complexity of the problems associated with a shift from fossil to biomass carbon raw materials. A number of integrated product-process design problems can easily identified: they link the various intermediate products, auxiliary materials and final products by different production and use processes. Integrated frameworks for model-based decision making are clearly required to mitigate risk and to reach truly sustainable solutions in a multi-disciplinary systems problem solving process.

4. TMFB @ RWTH Aachen University

A coordinated research effort addressing one of the prevalent systems problems in biorenewables processing has recently been initiated as a cluster of excellence (CoE) at RWTH Aachen University as part of the so-called Excellence Initiative (ExIni) of the German Federal and State Governments. The ExIni aims to promote top-level research, to improve the quality of German universities and to make Germany a more attractive and internationally competitive research location.
The objective and vision of the CoE “Tailor-made fuels from biomass” (TMFB) is to establish innovative and sustainable processes for the conversion of whole plants into fuels which are tailor-made for novel low-temperature combustion engine processes with high efficiency and low pollutant emissions, paving the way to the 3rd generation of biomass fuels. A tailor-made fuel is understood to be a well-defined blend of distinct molecular components with optimised physicochemical properties for future combustion systems, which can be produced by sustainable and economically viable production processes.
The definition of optimal fuel characteristics is understood as an inverse product design problem in the first place. Since the design of the combustion engine itself is not fixed and subject to modification, a first integrated product-process design problem results which calls for interdisciplinary interaction between combustion engineers and fuel design experts. This product-process design problem is focusing on the end use properties of the process where the properties of the fuel are considered to be the design parameters of the combustion process in a future engine. Another product-process design problem is related to fuel production process design which relies on the selection of chemical pathways and novel catalytic conversion processes. Likewise, these chemistry-centered decisions contribute to the decision variables of fuel production.
Both interrelated product-process design problems are addressed by a multi-disciplinary research team involving combustion engineering, physical chemistry and molecular thermodyamics, chemo- and bio-catalysis, process and reaction engineering as well as systems engineering and applied mathematics.

The research in the CoE has been started in late 2007 and is planned to involve more than 50 PhD students and post-doctoral researchers. It is structured into integrated research fields (IRF) supported by core interaction fields (CIF). The three IRFs focus on molecular transformations, process and reaction engineering as well as combustion engineering while the two CIFs provide methods and tools for physical property estimation and model-based fuel design. Despite its relevance, plant genetics and biomass production have been excluded form the cluster activities but are considered in collaboration with other research groups.

5. Summary

This short paper briefly introduced the systems problems related to biorenewables processing with an emphasis on biofuels production. This paper does not aim at a comprehensive literature survey or at the presentation of most recent research results. Rather, it intends to raise the awareness for the tremendous opportunities for process systems engineers beyond the classical issues related to process design given a chemical pathway, the plant capacity and product specifications. Sustainable solutions require a much wider perspective which at least addresses the various interdependent process and product design problems in an integrated manner. While process systems engineers do not have the expertise to contribute to all aspects of the overall problem equally well, they are in a very good position to address the multi-faceted and cross-displinary decision-making problems depicted in Fig. 1 in close collaboration with chemists, plant scientists, (bio-)chemical and process engineers as well as application experts.
An extended version of this paper will be prepared to give a more complete literature survey. This future paper will review selected research results to even better illustrate the opportunities and challenges for process systems engineering research.

References

A. Corma, S. Iborra, A. Velty (2007), Chemical routes for the transformation of biomass into chemicals, Chem. Rev., 107, 2411-2502.

J. Fargione, J. Hill, D. Tilman, S. Polasky, P. Hawthorne (2008), Land clearing and the biofuel carbon debt, Science, 319, 1235-1238.

R. Gani (2004), Chemical product design: challenges and opportunities, Computers & Chemical Engineering, 28, 2441-2457.

T. van Gervent, A. Stankiewicz (2009), Structure, energy, synergy, time: The fundamentals of process intensification, Ind. Eng. Chem. Res., 48, 2465-2474.

G. W. Huber, S. Iborra, A. Corma (2006), Synthesis of transportiation fuels from biomass: chemistry, catalysts and engineering, Chem. Rev., 106, 4044-4098.

K.-U. Klatt, W. Marquardt (2009), Perspectives for process systems engineering: Personal views from academia and industry. Computers & Chemical Engineering, 33, 536-550.

S. Fayazz, P. Frenzel, M. Köster, B. Kollmeier, J. McIntyre, K. Meier, M. Müller, J. Schmidt, P. Schmidt, I. Somoza, N. Weber, T. Weinert, J. Ayesterán, N. Kopriwa, A. Pfennig (2009), Wie können wir zukünftig ausreichend Energie nachhaltig bereitstellen? CLB Chemie in Labor und Biotechnik, 60, 01-02/2009, 32-39.

T. Sakakura, J.-C. Choi, H. Yasuda (2007), Transformation of carbon dioxide, Chem. Rev., 107, 2365-2387.

R. Sierra, A. Smith, C. Granda, M.T. Holzapple (2008), Producing fuels and chemicals from lignocellulosic biomass, Chem. Eng. Progr., August 2008, S10-S18.

G. Stephanopoulos (2007), Challenges in engineering microbes for biofuels production, Science, 315, 801-804.

L. P. Wackett, Microbial-based motor fuels: science and technology Microbial Biotechnology (2008) 1(3), 211–225.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimization in the Petroleum Refining industry – the Virtual Refinery

Lincoln F. Lautenschläger Moro

PETROBRAS SA, Av Chile 65, Rio de janeiro RJ 20031-915, Brasil

Abstract

As a result of the recent world economy slowdown many investment plans for refinery capacity expansions were postponed, and the optimization of existing process units, that had lost priority in relation to the design of new ones, came again to prominence. We cannot be sure if this priority will change in the near future but it seems sure that petroleum companies must be prepared to cope with a very uncertain scenario. This means that the abilities and the corresponding tools need to be well adapted to optimization and design activities, and the transition from one to the other has to be as quick and smooth as possible.

The tools used for both activities, e.g. process simulators, are usually diverse and the same applies to the teams in charge of optimization and design.

The present work tries to point out the main improvements that will have to be performed in the process engineering tools so that they may be better adapted to the this scenario.

The main improvement we point out is the so called “virtual refinery”, which is basically a rigorous dynamic model incorporating every piece of equipment in the plant. This model would be kept running continually in order to mimic the actual behavior of the plant and would be periodically upgraded to include new equipment and different operating modes. This model would then be able to represent the process in a comprehensive and detailed way and could supply the models used in simpler and narrowly oriented tools.

In order to make this vision come true several problems demand further investigation and development, and some of them are described.

Keywords: optimization, process control, petroleum, refining.

1. Introduction

For several years since the beginning of this decade the demand for petroleum products kept increasing sharply and, as a result, the crude oil prices as well as the prices of most products also increased. For example, prices for North-sea Crude Brent have increased from less than US$ 30 a barrel in 2003 to almost US$ 150 in mid 2008. Due to that several companies invested heavily in building new refining units as well as in expanding capacity in the existing ones. After an interval of several decades even grass-roots refineries were planned and built.

Then the situation reversed. The world economic growth slowed down and the demand of oil products flattened or even decreased. Many investments plans for capacity expansions were postponed. The crude oil prices fell to the levels of 2004, the same happening to the prices of most products. Despite that, the price margins, that is, the

difference between the price of a product (e.g. gasoline or diesel) and that of crude oil, after also falling sharply, quickly recovered.
As a result, the optimization of existing process units, that had lost priority in relation to the design of new ones, came again to prominence. At the present moment, we do not risk to forecast the world economy behavior in the medium term, i.e., if, after a pause, it will quickly recover the sharp growth patterns of the last years or if it will remain weak for quite some time.
Anyway, it became clear that flexibility is the keyword for everybody involved in the industrial sector, as is the case of process engineering in the refining industry. We cannot be sure if the priority in the near future will be the design of new facilities or the optimization of existing ones. This means that the abilities and the corresponding tools need to be well adapted to both and the transition from one activity to the other has to be as quick and smooth as possible.
In practice, this is not what we see. The tools used for these activities, e.g. process simulators, are usually diverse and the same applies to the optimization and design teams.
This situation holds true even when we look at mature and well-established activities like production planning and capital investment evaluation. Although the underlying technology is the same- linear programming- the commercial solutions available for both are different and are in general used by separate teams.
The present work tries to point out the main improvements that will have to be performed in the process engineering tools so that they may be better adapted to the this uncertain situation.

2. Current Process Engineering Tools

According to Magalhães (2004), the optimization of the refining business involves decisions that take place at three levels: strategic, tactical and operational.
At the strategic level, where the time horizon is normally years, these decisions deal with the long-term corporate objectives. Examples of decisions taken at this level are new products and markets to operate, new plant location, revamps for increasing capacity or flexibility, acquisitions and merges, partnerships, and so on. Typically the strategic plan covers a horizon of 10 to 15 years and is updated annually.
Tactical decisions cover the mid-term horizon (months, quarters) and emphasize the deployment of corporate strategies into a more detailed plan. In the refining industry, the corporate planning is an instance of these tactical decisions and involves several refineries, whose integration can provide significant synergies for the company as a whole. This plan covers, as a rule, a time span of a whole year, but treats with more detail the first month or quarter. Decisions such as crude purchase, target market to be supplied, products to be maximized in each facility and transfers of intermediate products among facilities are typical of this level.
Operational decisions are of the short-term type, where the time involved may vary from hourly to a monthly span. The main examples of decisions at this level are: refinery production and operation planning and scheduling, inventory management, transportation network utilization, process units' optimization and product blending.
There is no refiner nowadays that does not use advanced process engineering tools to improve business results (Moro 2001). Such tools range from advanced process control up to corporate long-term planning, passing through process optimization, scheduling, and short-term planning. Supporting these applications, one can find a myriad of tools like soft-sensors, data historians, relational databases, dynamic simulators, unit

monitoring systems, etc. The traditional structure of these applications is shown on figure 1.

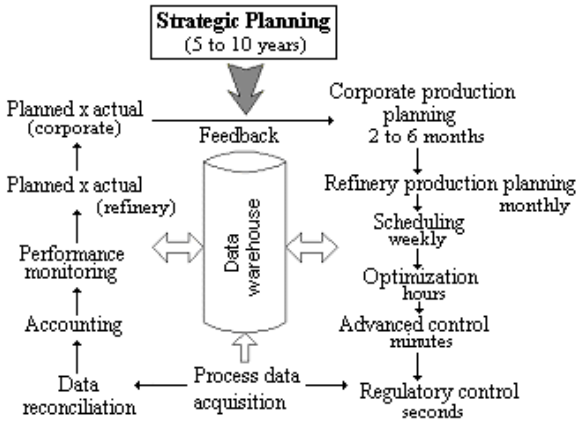


Figure 1 - Traditional structure of process engineering tools.

Very frequently these technological solutions are provided by different suppliers and are based on very different process representations. Even if one chooses to work with only one solution provider, this heterogeneity problem does not disappear because most commercial solutions were developed by different specialized teams and rely on different paradigms.

One can see in figure 1 that the piece that unifies this entire structure is the "Data warehouse". This can be roughly described as a relational and historical database responsible for feeding any application with the necessary input information and for storing its results for subsequent use by other applications. It is assumed that this architecture is able to guarantee an accurate representation of existing and planned assets in terms of expected performance.

In practice this is not good enough. Each application has a different way to model the process, e.g. linear programming, static rigorous models, dynamic rigorous and simplified models etc. This gives rise to several difficulties, since one cannot guarantee that all these models will give compatible results.

As an example of this situation, we can cite the strategic and tactical planning process in corporate and refinery levels. Companies usually use linear models for these activities but, very frequently these models were developed and are maintained and used by different teams, each one using its own tool. Due to that, although trying to represent the same production facilities, the models use different formulations and input and output variables and, understandably, give different results.

Analogous situation can be found when one looks at the rigorous process models used for the different purposes. Typical uses of such models are process design, case-studies and real-time optimization.

In general, design activities use process models covering a single process unit or even subsections of a process unit and the objective is to size new equipments or to define the modifications to be performed in existing ones. Sequential modular (S.M.) process simulators are considered to be the most adequate tool for these tasks. One of the main reasons that make them preferable over Equation-Oriented (E.O.) simulators is the

easier initialization. The simulators are usually used in simulation mode, i.e., the solver only looks for a viable solution.

We define opportunities or case studies as the activity performed by process engineers when searching for more profitable ways to operate existing facilities by changing operating conditions, streams routing and, eventually, feed quality. Best practices in this area indicate that one should use rigorous models covering an entire refinery or, at least, a few mutually interdependent process units. As a rule, engineers use sequential modular simulators capable to represent these several units in the same flowsheet, including reactors and separation equipment. Typically the simulators are used in optimization mode, with just a few variables free to be optimized by the solver.

Real time optimization (RTO), on the other hand, is better performed by Equation-Oriented, also called open-equations simulators, which solve, simultaneously, equations written in residual format. This format is very well adapted to changing modes of operation, which is a typical demand for RTO systems, since it is necessary to run them in a cycle that involves parameter estimation and optimization. The main difficulty with E.O. simulators relates to the initialization of variables because, since the equations are solved simultaneously, it is necessary to provide a reasonable starting point for every variable and there are thousands of them and many do not have a very well defined allowable range. This is not a big issue in Real-time optimization (RTO), since the previous solution provides a good starting point, but may pose significant practical difficulties for the design of new equipment.

Table 1 – Process engineering simulation and optimization tools.

Activity	Scope	Model	Time representation	Typical studies
Strategic Planning	Several plants	LP	Static	New hardware
Corporate production planning	Several plants	LP	Multiperiod	Feedstocks, products
Refinery production planning	Single plant	LP	Multiperiod	Intermediate streams routing
Refinery scheduling	Single plant	LP, MILP	Dynamic	Tank allocations
Optimization studies	Single plant	Rigorous-SM	Static	Operating conditions
Process design	Single unit	Rigorous-SM	Static	Operating conditions
Real-time Optimization	Single unit	Rigorous-OE	Static	Operating conditions
Operator training	Single unit	Rigorous	Dynamic	Controllability, Startups and Shutdowns
Multivariable predictive control	Single Unit	Transfer functions	Dynamic	Control

When it comes to dynamic simulators, the situation is similar. Typically, such simulators are used for operator training and for design of control strategies. These simulators are also rigorous but are usually not supplied by the same companies responsible for the static ones. The result is that the refiners end up with other tools,

used by other groups and there is no guarantee that the results are compatible and that the modifications in the real process will be adequately propagated through the several tools.
Table 1 describes some aspects of tools used for several process engineering activities. We can see that the scope may include several plants, which can be spread throughout the world or only a subsystem of a processing unit in a plant. Different kinds of models are also used, ranging from linear programming to rigorous dynamic models, based on differential equations. The time variable representation may be explicit, as in the dynamic models, or not even be considered, as is the case of the static models. An intermediate category is represented by the multi-period models that consist of a sequence of static models connected by inventories.

3. Virtual Refinery

In order to solve the main problems of the traditional structure described in the previous section we need a unifying tool that, working together with the data warehouse, is able to represent the process in a comprehensive and detailed way. This tool, which we chose to name "Virtual Refinery", can then supply the models used in simpler and narrower oriented tools.
In the limit the virtual refinery would basically be a rigorous dynamic model incorporating every piece of equipment in the plant, including storage tanks, piping, reactors, heat exchangers, pumps, columns etc. This model would be kept continually running in order to mimic the actual behavior of the plant and would be periodically upgraded to include new equipment and different operating modes.
This model would have to be continually updated and tuned to represent the dynamic behavior and the current situation of the production facilities. Besides ensuring that the model closely matches the plant behavior, this procedure would also supply information about process performance like, for instance, compressor efficiency, catalyst activity, and heat transfer coefficients. The model could, of course, be also used for engineer and operator training, controllability studies, design etc.
On the other hand, dynamic process simulators are not well adapted to some engineering activities. In fact, most optimization studies are performed off-line using static simulators that may cover a whole plant, but more frequently represent just a single unit. This means that it must be possible to quickly and easily transfer the dynamic model configuration to a rigorous static simulator. It can also be expected that engineers will eventually modify the static model configuration and such changes must also be available for the dynamic model. So the reverse configuration transfer, i.e., from the static to the dynamic must also be possible. The best approach would probably be to build a model database capable to store complete model configurations that could be accessed by both dynamic and static simulators, each uploading and downloading only the information necessary for its particular use. Current commercial tools do not provide such functionalities.
The rigorous static model could then be used to build simplified linear or non-linear programming tools for planning and scheduling activities. Analogously, the dynamic model could provide step-response curves to be directly used in the MPC's or, at least, to be compared with the curves obtained from step-testing procedures.
Figure 2 provides an overview of the relationship between the tools previously described.

4. Conclusions

At the first glance, the vision proposed in this paper seems easily achievable. Despite that, we do not expect to see it becoming true any time soon.

Some factors contribute to this apparent contradiction. The academic community seems to consider that all the interesting problems involved in this proposed structure have already been solved. When it comes to the solution providers, we can recognize two different situations. For those relatively small niche-oriented companies focusing on specialized solutions there is no incentive to invest in such integration issues. On the other hand, the solutions offered by large companies seem to be technologically divergent.

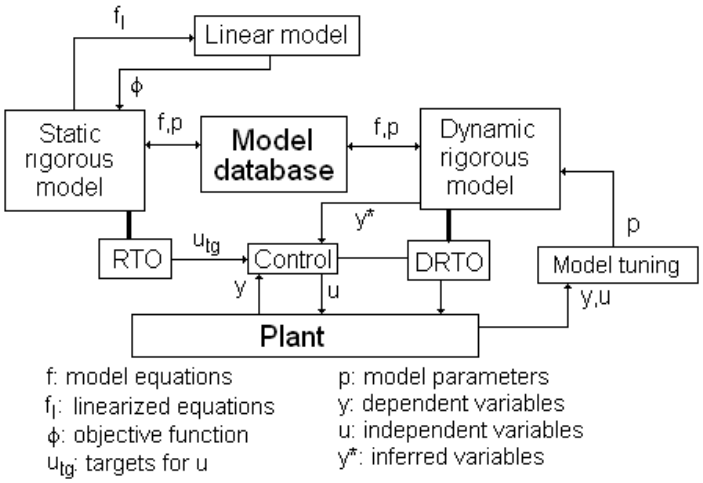


Figure 2 - Proposed structure of process engineering tools.

Among the several problems that demand further investigation and development in order that the proposed vision can someday be fulfilled, a few can be cited here:

- Procedures to tune dynamic models from live plant data;
- Dynamic modeling of reactors and other crucial processes;
- Direct use of dynamic models for on-line optimization, i.e. Dynamic RTO or DRTO;
- Development of a standard model configuration, well adapted to static and dynamic process simulators;

We hope that these and some other problems that hamper the attainment of the Virtual Refinery vision will be tackled by the academic community, and that the suppliers will identify business opportunities justifying the investment in solutions that contribute to this same goal.

References

M.V.O. Magalhães, 2004, Refinery Scheduling, PhD Thesis, Imperial College.

L.F.L Moro, 2001, Process technology in the petroleum refining industry - current situation and future trends, Computers and Chemical Engineering, 27, 1303/1305.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Nonlinear Dynamic Data Reconciliation in Real Time in Actual Processes

Diego Martinez Prata, Enrique Luis Lima, José Carlos Pinto

Programa de Engenharia Química/ COPPE / Universidade Federal do Rio de Janeiro – Cidade Universitária – CP 68502 – CEP: 21945-972 – Rio de Janeiro –RJ – Brazil

Abstract

In this work, some recent developments regarding the real time monitoring of industrial processes based on robust data reconciliation and gross error detection are reviewed. Particularly, robust data reconciliation and gross error detection are performed on line and in real time in a real polymerization process.

Keywords: Data Reconciliation, Robust Estimators, Gross Errors, Industrial Data.

1. Background

The rigorous monitoring of variables that describe the present and the long-term behavior of industrial processes is one of the first conditions required to guarantee the optimal operation of the plant. For this reason, process measurements are taken and used in process control procedures and for evaluation of process performance. With the advent of modern desktop computers, hundreds or thousands of process measurements are simultaneously measured and stored in massive storage media for the continuous monitoring of the process behavior and for performing process studies, such as the development of robust process models and the online optimization of the process operation. Therefore, modern industrial plants may provide very detailed and rich data sets for development of fundamental modeling studies. However, as information concerning such variables comes from real measurements, numerical values are normally corrupted by different kinds of errors. As a consequence, the collected data generally do not satisfy the process constraints, such as the mass and energy balances. For this reason, implementation of data rectification procedures is essential for obtainment of reliable information about the process behavior.

The data rectification procedure comprises different steps, such as variable classification, gross error detection and data reconciliation (including parameter estimation in complex models). Variable classification is required in order to determine if the available information is sufficient for solving the proposed reconciliation problem and to identify sets of observable and non-observable process variables. Different solutions have been proposed for the variable classification problem during the last four decades and detailed description of published procedures can be found in Romagnoli and Sanchez (2000).

Procedures for gross error detection are required to identify (and possibly remove from the set of measured data) observed values that do not follow the statistical distribution of the measured data, as the presence of gross errors may lead to significant biases during data reconciliation and/or parameter estimation and also to poor decisions. Gross Error Detection (GED) has been (and still is) thoroughly analyzed in the literature using different theoretical and numerical approaches, as described in detail by Narasimhan and Jordache (2000). Finally, during data reconciliation the measured data

are adjusted in order to satisfy the required process constraints (or the process model) and obtain reliable estimates for the unknown desired parameters.

Mathematical models play an important role in modern process engineering because they can be used for evaluation of process states, including prediction of non-measured variables (soft sensors). Besides, reliable process models are fundamental for implementation of on-line optimization and advanced process control techniques. As industry is rapidly shifting towards real time optimization and some crucial process parameters (such as heat transfer coefficients and catalyst activities) are likely to change at plant site along the time, estimation of model parameters in real time based on available process measurements is becoming an increasingly important issue.

The scientific literature shows that joint data reconciliation and parameter estimation with simultaneous gross error detection has emerged as a fundamental tool for on-line industrial applications. Thus, the development and implementation of these powerful procedures in real time may be of paramount importance for certain process control and real-time optimization applications. However, the implementation of monitoring procedures in-line and in real time, with the simultaneous data reconciliation, estimation of process parameters, evaluation of non-measured process variables and multiple gross error detection, has seldomly been performed in real world processes due to the existence of some inherent numerical difficulties. Some of these difficulties are related to the necessary treatment of huge amounts of data in short periods of time (for real time applications), while other difficulties are related to the robustness of available numerical techniques (convergence problems are unacceptable in most real time applications), especially when multiple gross errors are present. This work presents a brief overview on data reconciliation and multiple gross error detection applied for real time industrial process applications.

2. Background

Data reconciliation (DR) procedures can be used in a broad range of distinct applications, such as process control, real time optimization, process monitoring, safety analysis, improvement of operational efficiency, detection of gross errors and process malfunctions, among many others. As a consequence, DR represents an important step for many engineering activities in industrial process.

The use of DR procedures for analysis of process data was originally proposed in the early sixties, encouraged by the increasing capacity of digital computers. DR may be defined as the adjustment of process measurements and parameters in order to satisfy a certain set of conservation laws and process constraints (as described by a mathematical model), while minimizing some sort of objective function that weighs the deviations between corrected and observed plant values. (In general terms, the estimation of model parameters from available plant data is certainly within this context.) For this reason, it may be said that joint data reconciliation and parameter estimation (DRPE) is useful for the simultaneous analysis of process data (allowing for identification and removal of inconsistent data sets) and process model building. Different solutions have been proposed for the DR problem during the last four decades and good reviews are provided by Crowe (1996a), Romagnoli and Sanchez (2000) and Narasimhan and Jordache (2000).

GED is of fundamental importance for adequate model building and interpretation of process data, as poor input data generally lead to very poor model responses. The commonest methods used for GED are based on statistical hypothesis testing, which requires the definition of a proper statistics for the test, assuming that the error

distribution is known *a priori* (Narashimhan and Jordache, 2000). A gross error is detected if the computed statistical test exceeds a critical value, which is calculated with the assumed distribution. It is frequently assumed that measurement errors follow the well-known Normal (Gaussian) distribution, with zero mean. In this particular case, application of the maximum likelihood principle leads to the *Weighted Least Squares* (WLS) DR problem. Although the assumption of normality of the error distribution is usual, it is not necessary for DR and GED (Crowe, 1996b).

When multiple gross errors are present, a strategy (not only a method) for gross error detection may be necessary. Three strategies are normally used to deal with multiple gross error detection (MGED): Serial Elimination, Serial Compensation and Simultaneous (or Collective) Compensation (Narashimhan and Jordache, 2000). Principal component tests have also been used for MGED, but significant improvement of the MGED has not been observed (Jiang *et al*., 1999). The main disadvantage of these MGED strategies is that they are based on residuals obtained from regression. As residuals may be heavily biased by the presence of gross errors, the implementation of iterative or combinatorial procedures may be necessary, increasing the computational effort. Despite that, these strategies have been widely used for analysis of steady-state or quasi-steady-state process conditions. As real time industrial applications require frequent MGED analysis of nonlinear dynamic models, the usage of iterative and/or combinatorial procedures may be inappropriate. (In practical implementations, it is common to use a time varying moving window in order to capture the recent behavior of the process and reduce the size of the optimization problem to manageable dimensions during the solution of dynamic DR problems.)

In order to deal with the joint nonlinear dynamic data reconciliation (NDDR) and MGED problems, alternative approaches based on cluster analysis (Chen and Romagnoli, 1998; Abu-El-Zeet *et al*., 2002), artificial neural networks (Vachhani *et al*., 2001) and robust estimators (Özyurt and Pike, 2004; Prata *et al*., 2008a) have been proposed and successfully applied in the literature. Among these alternative approaches, robust data reconciliation (RDR) has been used more frequently.

There are several classes of robust estimators. The most important for DR are the M-estimators, which are generalizations of maximum-likelihood estimators. Robust estimators tend to look at the bulk of the data and ignore atypical values (gross errors), due their mathematical structure. Table 1 summarizes the works that applied robust M-estimators for industrial process monitoring, based on RDR procedures. This strategy has been recently applied for in-line and real time monitoring of real industrial nonlinear dynamic systems by Prata *et al*. (2008b,c). Comparative analysis of the performances of distinct robust M-estimators were performed by Özyurt and Pike (2004) and Prata *et al*. (2008a) for steady state and dynamic systems, respectively. Prata *et al*. (2008a) concluded that the Welsch M-estimator (which presents continuous first and second derivatives) shows better global performance, reducing the negative effects of gross errors upon the obtained estimates.

It is important to notice that the implementation of RDR approaches can be difficult sometimes, as some robust estimators are non-convex, requiring very good variable initialization in order to avoid local optima. Besides, some robust estimators present discontinuous first and/or second derivatives. As the computation of first and second derivatives may be necessary during the optimization (depending on the numerical technique implemented for minimization of the objective function), this may lead to lack of numerical stability and require the usage of smoothing functions (making numerical implementation and calculation more complex).

In order to overcome the numerical difficulties, Wongrat *et al*. (2005) proposed the use of Genetic Algorithms (GA) to perform the optimization of RDR. Other non-deterministic optimization methods, such as Tabu Search (TS), Simulated Annealing (SA) and the Particle Swarm (PS), can also be used. These methods present some additional advantages, such as the global character of the search (which avoid local minima), the unnecessary computation of derivatives and the simplicity of the implementation, although they are usually characterized by the high number of objective function evaluations, which may require more CPU time than conventional methods. However, the computed values of the objective function can be used for rigorous statistical analyses of the confidence regions of parameter estimates, which can also constitute an important benefit of these algorithms (Schwaab *et al.*, 2008).

Reimers *et al*. (2008) applied non-deterministic optimization methods (GA, TS and SA) to solve steady-state DR problems in a real industrial case. Prata *et al*. (2008b) also applied a non-deterministic optimization method (PS) to solve the robust NDDR problem in a real industrial plant, in-line and in real time. As a consequence, one may conclude that non-deterministic procedures can be used successfully for solution of DR problems in real time.

Table 1: Industrial applications of robust M-estimators for DR and MGED.

Author (Year)	Industrial Scenario	M-Estimator
Zhang *et al*. (1995)	Sulfuric acid plant.	Contaminated Normal
Chen *et al*. (1998)	Sulfuric acid plant.	Fair and Lorenzian
Bourouis *et al*. (1998)	Multistage flash desalination plant.	Contaminated Normal
Özyurt and Pike (2004)	Sulfuric acid plant.	Contaminated Normal, Cauchy, Fair, Hampel, Logistic and Lorenzian.
Faber *et al*. (2006)	Industrial coke-oven-gas purification process	Li *et al*. (2000)
Faber *et al*. (2007)	Industrial coke-oven-gas purification process	Li *et al*. (2000)
Schladt and Hu (2007)	Industrial distillation column.	Contaminated Normal
Lid and Skogestad (2008)	Industrial naphtha reformer	Contaminated Normal
Prata *et al*. (2008b)	Industrial Polymerization Reactor.	Welsch

It is important to emphasize that both traditional and robust formulations of DR problems depend on the relative magnitudes of the elements of the variance/covariance matrix of measured data. Robust estimators can also be used for the simultaneous characterization of the variance/covariance matrix of measurements and elimination of gross errors (Chen *et al*., 1997; Morad *et al*., 1999), resulting in improved characterization of the statistics of the process data.

3. Analysis of Real Industrial Data

Some of the obtained results are presented below for the bulk propylene polymerization performed in a stirred tank reactor (model and process description can be found in Prata *et al*., 2008d). Raw process data were obtained over a time interval of

7h, with a sampling time of 5 min (300s), using a time moving window of 1h (samples collected every 5 min for 22 measured variables, resulting in 286 reconciled variables in each window). The DRPE were conduced in two different computers using the *WLS* and *Welsch* estimators. The symbols (○), (●), (●) e (●) represent the plant measurements, laboratory analyses (obtained with delays of 8h, but synchronized after reconciliation), *WLS* reconciled values and reconciled values by the *Welsch* estimator, respectively. The letter (**E**) represents the gross errors detected by the *Welsch* estimator. The magnitudes of the measurement errors (associated with the instrument) and of the process noise (associated with the operation) were evaluated for all process signals based on the available data, but are not presented here for confidential reasons.

Figures 1 and 2 show measured and reconciled data for two variables: the propane concentration in the recycle stream and the reactor temperature, respectively. It can be observed that the results provided by the *Welsch* estimator are better than the results provided by *WLS* estimator, as the first one is much less sensitive to the presence of gross errors than the second one. It is important to note that deviations from the temperature setpoint (340.15 K) can not be greater than 1K (dot line); otherwise, control actions are taken in order to prevent potentially dangerous operation conditions. Figure 2 shows that some *WLS* reconciled values would violate this general rule, generating unnecessary control actions.

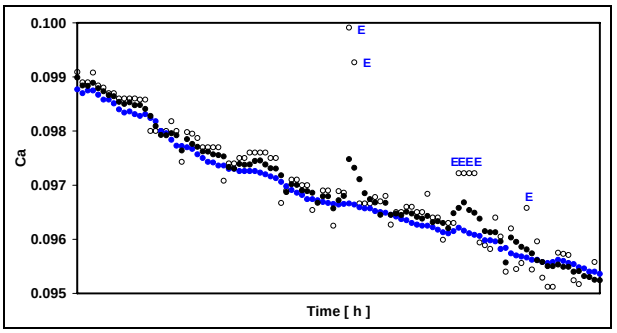


Figure 1. Ca – Propane concentration in the recycle stream.

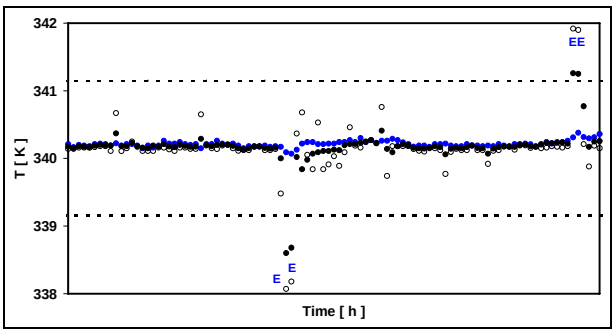


Figure 2. T – Reactor temperature.

Figure 3 shows measured, reconciled and laboratory data for the Melt Index of the polymer product. Again, it can be observed that the results obtained with the *Welsch* estimator are better than the results obtained with the *WLS* estimator. The in-line and real time MI inferences provided by the *Welsch* estimator are underpinned by the laboratory analyses (after synchronization), showing that in-line information provided

by a process rheometer is biased. It seems clear that the proposed RDR scheme based on *Welsch* estimators can provide better results than the in-line process rheometer, which requires frequently re-calibrations during process operation. Therefore, the RDR scheme can improve the safety and performance of the process operation, aggregating value for decision making.

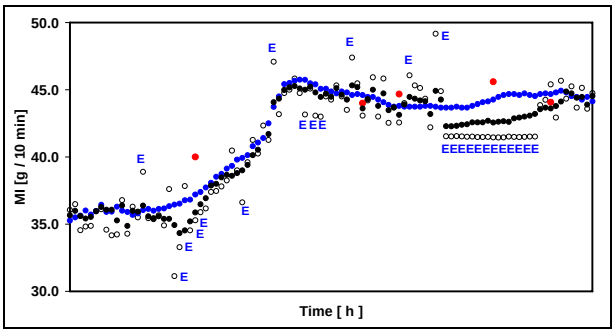


Figure 3. MI – Melt Index (after industrial extruder).

It is important to notice that the computational time required to accomplish all the calculations for each moving window (size=1h) using the *Welsch* robust estimator was roughly equal to 260s (250s for solution of the nonlinear DRPE problem with PS and 10s for data handling) in a standard Pentium 4 (single core 3.0 GHz) with 1024 MB memory, allowing for real time implementation of the code.

4. Conclusions

Some recent developments regarding the real time monitoring of industrial processes based on robust data reconciliation and gross error detection were reviewed. Particularly, a robust method for the joint nonlinear data reconciliation and parameter estimation with simultaneous gross error detection, based on the particle swarm optimization algorithm, was implemented. The proposed method was used to describe the operation of an industrial bulk propylene polymerization process and was validated with actual industrial data in-line and in real time. The negative influence of gross error was minimized with the help of a *Welsch* estimator. Off-line laboratory analyses confirmed the effectiveness of the proposed procedure, which indicated frequent failures of the process rheometer and allowed for improvement of the process performance. The numerical procedures required 260s of CPU in standard desktop computers, allowing for implementation of advanced process monitoring and control in real time.

References

Z. H., Abu–El–Zeet, V. M., Becerra, P.D., Roberts, 2002, "Combined Bias and Outlier Identification in Dynamic Data Reconciliation", *Computers and Chemical Engineering*, v.26, pp. 921-935.

M., Bourouis, L., Pibouleau, P., Floquet, *et al*., 1998, " Simulation and data validation in multistage flash desalination plants", *Desalination.*,v.115, pp. 1-14.

J., Chen, A., Bandoni, J.A., Romagnoli, 1997, "Robust estimation of measurement error variance/covariance from process sampling data". *Computers and Chemical Engineering*, v. 21, pp. 593-600.

J., Chen, J. A., Romagnoli, 1998, "A Strategy for Simultaneous Dynamic Data Reconciliation and Outlier Detection", *Computers and Chemical Engineering*, v. 22, pp. 559-562.

X., Chen, R. W., Pike, T. A., Hertwig, J. R., Hopper, 1998, "Optimal Implementation of On-Line Optimization", *Computers and Chemical Engineering*, v.22, pp. S435-S442.

C. M., Crowe, 1996a, "Data Reconciliation - Progress and Challenges", *Journal Process Control*, v. 6, pp. 89-98.

C. M., Crowe, 1996b, "Formulation of Linear Data Reconciliation Using Information Theory", *Chemical Engineering Science*, v. 51, pp. 3359-3366.

R., Faber, H., Arellano-Garcia, P., Li, G., Wozny, 2007, "An optimization framework for parameter estimation of large-scale systems", *Chemical Engineering and Processing*, v. 46, pp. 1085-1095.

R., Faber, B., Li, P., Li, G., Wozny, 2006, "Data reconciliation for real-time optimization of an industrial coke-oven-gas purification process", *Simulation Modelling Practice and Theory*, v. 14, pp. 1121–1134.

Q., Jiang, M., Sánchez, M., Bagajewicz, 1999, "On the Performance of Principal Component Analysis in Multiple Gross Error Identification", *Industrial Engineering and Chemical Research*, v.38, pp.2005-2012.

B., Li, B., Chen, D., Duan, M., Kong, 2000, "Correction coefficient method for gross error detection based on temporal redundancy", *Qinghua Daxue Xuebao, Ziran Kexueban*, v.40, pp. 65-69.

T., Lid, S., Skogestad, 2008, "Scaled steady State Models for Effective On-Line Applications", *Computers and Chemical Engineering*, v.32, pp. 990-999.

K., Morad, W.Y., Svrcek, I., McKAY, 1999, "A robust direct approach for calculating measurement error covariance matrix". *Computers and Chemical Engineering*, v. 23, pp. 889-897.

S., Narashimhan, C., Jordache, 2000, *Data Reconciliation and Gross Error Detection: An Intelligent Use of Process Data*, Gulf Professional Publishing. Houston, TX,

D. B., Özyurt, R. W., Pike, 2004, "Theory and practice of simultaneous data reconciliation and gross error detection for chemical process", *Computers and Chemical Engineering*, v. 28, pp. 381-402.

D. M., Prata, E. L., Lima, J. C., Pinto, 2008d, "In-Line Monitoring of Bulk Polypropylene Reactors Based on Data Reconciliation Procedures", *Macromolecular Symposia*, v. 271, p. 26-37.

D. M., Prata, J. C., Pinto, E. L., Lima, 2008a, "Comparative Analysis of Robust Estimators on Nonlinear Dynamic data Reconciliation", *Computer-Aided Chemical Engineering*, v. 25, p. 501-506.

D. M., Prata, J. C., Pinto, E. L., Lima, 2008c, "Robust Data Reconciliation and Gross Error Detection through Particle Swarm Optimization for an Industrial Polymerization Reactor" (*submited*).

D. M., Prata, M., Schwaab, J. C., Pinto, E. L., Lima, 2008b, "Abordagem Híbrida para Detecção de Erros Grosseiros e Reconciliação Robusta de Dados: Aplicação Industrial", In*: Proceedings of the COBEQ 2008*, Recife, Setembro (*in portuguese*).

C., Reimers, J., Werther, G., Gruhn, 2008, "Flowsheet simulation of solids processes Data reconciliation and adjustment of model parameters", *Chemical Engineering and Processing*, v. 47, pp. 138–158.

J. A., Romagnoli, M. C., Sanchez, 2000, *Data Processing and Reconciliation for Chemical Process Operations*, Academic Press. San Diego.

M., Schladt, B., Hu, 2007, "Soft Sensor Based on Nonlinear Steady-State Data Reconciliation in the Process Industry", *Chemical Engineering and Processing*, v.46, pp. 1107-1115.

M., Schwaab, E. C., Biscaia Jr., J. L., Monteiro, J. C., Pinto, 2008, "Nonlinear Parameter Estimation through Particle Swarm Optimization", *Chemical Engineering Science*, v. 63, pp. 1542-1552.

P., Vachhani, R., Rengaswamy, V., Venkatasubramanian, 2001, "A Framework for Integrating Diagnostic Knowledge with Nonlinear Optimization for Data Reconciliation and Parameter Estimation in Dynamic Systems", *Chemical Engineering Science*, v. 56, pp. 2133-2148.

M., Wongrat, T., Srinophakun, P., Srinophakun, 2005, "Modified Genetic Algorithm for Nonlinear Data Reconciliation" *Computers and Chemical Engineering*, v. 29, pp. 1059-1067.

Z., Zhang, R.W., Pike, T., Herting, 1995, “Source Reduction from Chemical Plants Using On-line Optimization”, *Waste Management*, v. 15, pp.183-191.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Energy Systems Engineering

Paul I. Barton

Process Systems Engineering Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.

Abstract

Meeting current and future demands for energy is often quoted as the most pressing problem facing humanity in the 21st century. This problem has to be addressed in a global environment where resources are constrained and diminishing, while at the same time raising standards of living for a growing population and preventing catastrophic climate change. Recent experience has shown that energy production is increasingly competing with other demands on the same resources, such as commodity chemicals production and food production, and all compete for water as a primary resource. In moving to a sustainable future, radical changes will have to take place in how we produce, transport and utilize energy and commodity chemicals. The solutions will be many and diverse, and highly dependent on geographic location and patterns of consumption.

In many respects, chemical engineers, and process systems engineers in particular, are uniquely qualified to address this grand challenge. The design, optimization and operation of a diverse range of novel, highly integrated processes and supply chains is our core competence. Furthermore, process systems engineering tools, techniques and formulations are highly developed and ready to meet this challenge, so the creative application of the best of these will have a tremendous impact. On the other hand, new challenges and applications also spawn new formulations and tools. The purpose of this talk is to use several examples from ongoing projects in the Process Systems Engineering Laboratory at MIT to illustrate how the energy systems engineering challenge is fostering both the creative application of the state-of-the-art and new formulations and tools.

Gasification of low-value carbonaceous feedstocks (e.g., coal, lignocellulosic biomass, oil shale, etc.) is a promising route to the *polygeneration* of electrical power, transportation fuels and chemicals. In order to reduce or even reverse carbon dioxide emissions to the atmosphere, novel processes have to be invented with integrated carbon capture and sequestration. We have recently used PSE tools to invent a novel high-efficiency process for power production from coal that almost eliminates the efficiency penalty of carbon capture and sequestration, and dramatically reduces the water consumption of the process. This process should be particularly attractive in water-stressed regions with large coal reserves, such as Northern China and the American West. The notion of polygeneration enables a plant to exploit fluctuations in prices and demands for several products (e.g., electrical power, transportation fuels) on several time scales. However, in this setting it is no longer optimal to design for a single steady state. Therefore, we are developing novel optimization formulations and algorithms for integrated design and operation that yield optimal profitability over a range of price and demand scenarios.

The productivity of transportation fuels and chemicals from the aforementioned processes can be significantly enhanced by the introduction of a carbon-free hydrogen source, specifically water splitting using heat and/or electrical power from nuclear and solar sources. Moreover, large amounts of carbon dioxide are produced from current and future hydrogen demands. One of the most promising processes for the production of hydrogen using heat from a nuclear reactor is the sulfur-iodine (SI) thermo-chemical cycle. At the heart of this process is the Bunsen reactor, in which water, sulfur dioxide and iodine are combined to form sulfuric acid and hydrogen iodide. Due to a fortuitous liquid-liquid phase split, the sulfuric acid and hydrogen iodide can be separated and thermally decomposed to generate oxygen and hydrogen, respectively. The key to designing and optimizing this process is accurate thermophysical property models that can predict phase and chemical equilibrium of multi-electrolyte, multi-solvent solutions. This challenge has prompted us to develop a new version of the electrolyte-NRTL model suitable for such solutions, and large-scale optimization techniques to develop self-consistent thermodynamic databases for parameter estimation. In addition, the challenge of fitting binary parameters to such complex phase diagrams required us to propose a new formulation for the parameter estimation problem as a bilevel program; the inner programs determine the stable phase splits by Gibbs free energy minimization, and the outer program minimizes the least squares error. Solution of this formulation requires our recent developments in algorithms for the solution of bilevel programs with nonconvex inner programs. Moreover, this advance has very broad implications for how binary parameters should be fitted to any phase equilibrium data.

The emergence of a global liquefied natural gas (LNG) market is facilitating the development of novel supply chains that enable the exploitation of remote natural gas combined with carbon capture, sequestration and enhanced oil recovery. The implementation of these supply chains require novel process designs at several stages in the overall supply chain, and a common feature of these processes is the need for heat and power integration between streams at varying temperatures and pressures. This problem is considerably more complex than conventional heat integration, and leads to novel nonconvex optimization formulations that vary temperature and pressure levels simultaneously to find the optimal process design. These formulations have prompted us to develop the novel notion of *global optimization of algorithms*, in which the process model is evaluated as an algorithm, and the algorithm is relaxed directly in the global optimization procedure. Such advances are making these complex problems tractable within a deterministic global optimization framework.

The quest for new sources of oil to meet ever-increasing demand is motivating the development of technologies for oil exploration and production in ultra-deep water (depths of one mile or more). At such depths of water, production facilities have to be located on the seabed and operated remotely. Due to the extreme costs associated with maintenance and repair of subsea production processes, an autonomous process must be designed that is guaranteed to be robust to input disturbances. These guarantees must be provided a priori at the design stage, which is motivating the development new formulations and optimization algorithms for robust process design.

Finally, the design and operation of processes for the production of transportation fuels from biomass via fermentation is motivating new dynamic optimization formulations and algorithms. The bioreactor can be modeled using conventional differential equation formulations, but this must be coupled with predictive models of the response of the

microorganism(s) to their environment. A common approach to this is a flux balance model that represents the behavior of a microorganism as a linear program (LP); increasingly, large-scale flux balance models are being constructed from genomic data. The overall model is a system of differential equations with LPs embedded, which can be reformulated as an equivalent complementarity system, a type of hybrid (discrete/continuous) system. This motivates the development of novel simulation and optimization algorithms for such complementarity systems, in particular drawing on ideas and techniques from nonsmooth analysis.

Keywords: coal, carbon capture and sequestration, nuclear hydrogen production, liquefied natural gas, biofuels.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Thermally Coupled Distillation

José A. Caballero

Department of Chemical Engineering, University of Alicante, Ap. Correos 99, 03080 Alicante, Spain

Abstract

This paper presents a brief overview in the design of mixed thermally coupled-heat integrated distillation sequences. The approach considers from conventional columns (each distillation column with a condenser and a reboiler) to fully thermally coupled systems (only one reboiler and one condenser in the entire system). The advantages and disadvantages of thermally coupled distillation in comparison with conventional columns are discussed; the consequence is that the optimal configuration is usually an intermediate point between fully thermally coupled systems and conventional systems. The paper also includes 1. Structural relations between separation tasks and number of column sections, logical relationships that separation tasks must obey in order to select only feasible solutions avoiding configurations that are always sub-optimal. 2. The convenience of using a task based approach instead a column based approach to avoid the redundancy due to thermally couple configurations. 3. How to take advantage of thermodynamically equivalent configurations to synthesize more operable configurations. 4. Heat integrated – thermally coupled distillation systems.

Keywords: Process synthesis, Distillation, Thermally coupled distillation.

1. Introduction

Distillation is likely the most studied unit operation in the history of Chemical Engineering. This research effort is amply justified because distillation has been and continues to be the most widely used separation technique. The main disadvantage of distillation is the high energy requirement. Integration of distillation with the overall process can result in important energy savings[1,2] However, the heat integration between distillation and the rest of the process is often limited.

From the point of view of energy requirements, separation sequences using conventional columns (a single feed with two product streams, condenser and reboiler) suffer from an inherent inefficiency produced by the thermodynamic irreversibility during the mixing of streams at the feed, top and bottom of the column[3]. This remixing is inherent to any separation that involves an intermediate boiling component and can be generalized to an N component mixture. Theoretical studies developed by Petluyk and coworkers[3], showed that this inefficiency can be improved by removing some heat exchangers and introducing thermal coupling between columns. If a heat exchanger is removed the liquid reflux (or vapor load) is provided by a new stream that is withdrawn from another column, in this way it is possible to reduce the energy consumption and under some circumstances also the capital costs. A fully thermally coupled configuration (FTC) is reached when the entire vapor load is provided by a single

reboiler and all the reflux by a single condenser. Different researchers [4], [5], [6] have shown that thermally coupled configurations are capable of typically achieving 30% in energy reduction when compared to conventional systems. Halvorsen and Skogestad [7, 8, 9] proved that the minimum energy consumption for an N component mixture is always obtained for the FTC configuration. However the number of alternatives and the complexity for generating feasible –possible optimal- configurations greatly increases.
In spite of the reduction in energy consumption, there is a price to be paid when using thermally coupled systems: 1. The energy must be supplied in the worst conditions, at the highest temperature in the reboiler and removed at the lowest temperature in the condenser, preventing in most cases the uses of cheaper utilities. 2. When using conventional columns it is common to constrain the alternatives to sharp separations, for example in a three component mixture (ABC) –sorted by volatilities-, we postulate initially separations A from BC (A/BC) and AB from C (AB/C). However, in fully thermally coupled distillation we could introduce sloppy separations –an intermediate product is allowed to distribute along the column- and therefore, we are increasing the number of column sections. This increase does not imply an increase in the number of columns, but usually an increase in the total number of trays. A detailed discussion on number of columns section needed for a given separation can be found in Agrawal [10] and Caballero and Grossmann [11]. 3. In FTC systems, the minimum vapor flow is that of the most difficult separation [8, 9] and therefore some column sections will have large diameters. 4. Operation is more difficult due to the large number of interconnections between columns. Therefore, it cannot be concluded that complex configurations are always superior compared with sequences of simple columns. Instead, the optimum configuration will depend on the specific mixture and feed conditions and it is usually an intermediate situation between FTC and conventional columns.

2. 2. Structural Considerations

In the synthesis of distillation sequences using complex columns it is convenient to separate the concept of separation task from the equipment in which a separation task is performed. In fact, in complex columns, a single column usually performs more than a separation task. This aspect can be illustrated with a simple three component mixture (ABC). Using conventional columns we have two options. The direct sequence: separate first A (A/BC) and then B from C (B/C) or the indirect sequence separate first AB from C (AB/C) and then A from B (A/B). Note that in this case there is a one to one relationship between tasks and columns. However, in a fully thermally coupled configuration (sequence of tasks AB/BC, indicating that the component B distributes along the column, A/B and B/C), the distribution of tasks in columns depends on the alternative chosen among all the thermodynamically equivalent.
Of course, a sloppy separation is not constrained to columns with thermal links. The classical pre-fractionation arrangement shows this last case. However, having established that we have a source for the vapor load (i.e. a reboiler or a vapor side stream from another column) and a source for the reflux, the model of the first column in the pre-fractionation configuration is identical to the model of the first column in the Petlyuk arrangement. Two important consequences follow for the modeling of the problem:
1. *We can separate the heat exchange from the separation task.* In fact the heat exchange can be defined as a new task. Note that in systems with no heat integration, there is a one to one relationship between the heat exchange task and the physical

device that performs that task (the heat exchanger). However, in heat integrated systems, two heat exchange tasks are able to be joined in a single device.

2. *We can study the separation structure (sequence of tasks) without taking into account the presence or not of heat exchangers* (except those related to the streams that leave the systems, as we will comment latter).

A column section can be defined as the portion of a distillation column which is not interrupted by entering or exiting streams or heat flows. A separation task is formed by two column sections, a rectifying and a stripping section, and therefore, it is convenient to consider a separation task as a pseudo-column by similarity with conventional distillation columns, although in the final assignment of column sections to actual columns the two sections of a given separation task do not necessarily belong to the same column. From a structural point of view, thermally coupled configurations are much more complex than sequences of conventional columns. Agrawal [10]established that, in contrast to systems with conventional columns, the number of column sections is not fixed. Consider, for example, the sequence in Figure 1a. (This sequence is performing the separation tasks AB/C – A/B) It is possible to remove the heat exchanger related to stream AB and introduce a thermal couple between both columns. The resulting configuration is shown in Figure 1b. In this configuration the sequence of separation tasks is the same than in the original. It is possible to remove the heat exchanger B to create a FTC configuration. However, in this case it is necessary to introduce two more column sections (the resulting configuration is shown in Figure 1c), this modification also changes the sequence of separation tasks performed by the columns (AB/CD; A/C; C/D). Therefore, there is a trade off between the energy consumption and the extra number of column sections.

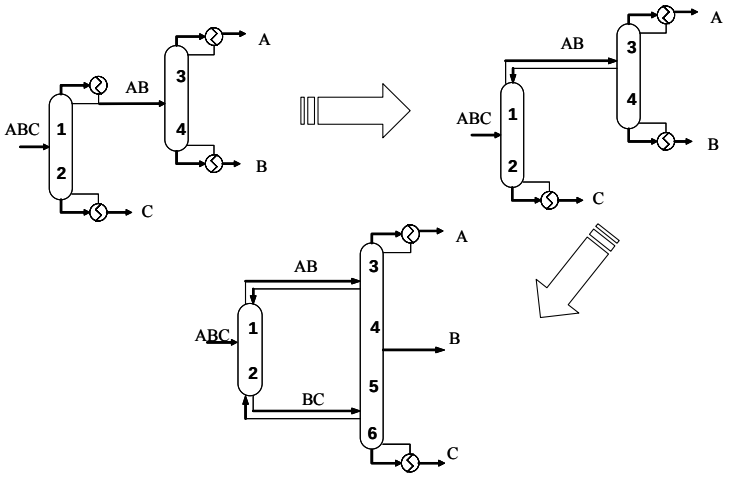


Figure 1. Graphical representation of the steps for going from a conventional system to a FTC system.

The minimum number of column sections (or tasks, each task produces two column sections) is related to the number of heat exchangers associated to final products (or more in general to streams that leave the system)[10, 12]. In an N component mixture in which all the final products have a heat exchanger there are 2(N-1) column sections – (N-1) separation tasks- the previous relationship is always true independently of the existence or not of heat exchangers associated to intermediate streams The heat exchangers associated to final products (pure components or streams that leave the

system) of intermediate volatilities can be removed, but in this case we have to add two more sections for each heat exchanger removed. In fully thermally coupled systems the minimum number of separation tasks is (4N-6)/2 [10]. However, it is possible to generate structures with up to N(N-1)/2 separation tasks. At difference of distillation sequences formed by conventional columns in which to generate valid separation sequences and eventually generate a superstructure is straightforward, this is not the case in thermally couple distillation, at least if we are not interested in including suboptimal separations like is the case of structures with duplication of key components or structures with duplication of separation tasks that has been shown that are always suboptimal solutions. The structural relations between the tasks require a detailed presentation that is out of the scope of this brief review, but it can be found in references [13-16] It is interesting remark that those relations between tasks can be easily transformed in logical rules that can be translated to algebraic equations in terms of binary variables and included in a mathematical programming code to extract the best sequence of columns. Besides, if we are interested in simply generate feasible alternatives it is possible algorithmically adapt those rules to generate alternatives [16].

3. Thermodynamically equivalent configurations

In the design of thermally coupled distillation systems an extra difficulty appears. Two configurations apparently different can be structurally equivalent (Thermodynamically equivalent). Consider Figure 2(b to e). This Figure shows different arrangements of two columns, but the sequence of separation tasks is the same. For example, it is possible to go from the configuration in Figure 2b to 1e by changing section 6 from column 2 to column 1, but flows, temperatures, pressures of the sections remain unchanged. This concept can be generalized to N component mixtures[15]. Note that in these cases a single column can perform more than one separation task. Therefore, in a preliminary design it is important deal only with thermodynamically different alternatives, because differences between thermodynamically equivalent alternatives are usually negligible and only important at the operational level.

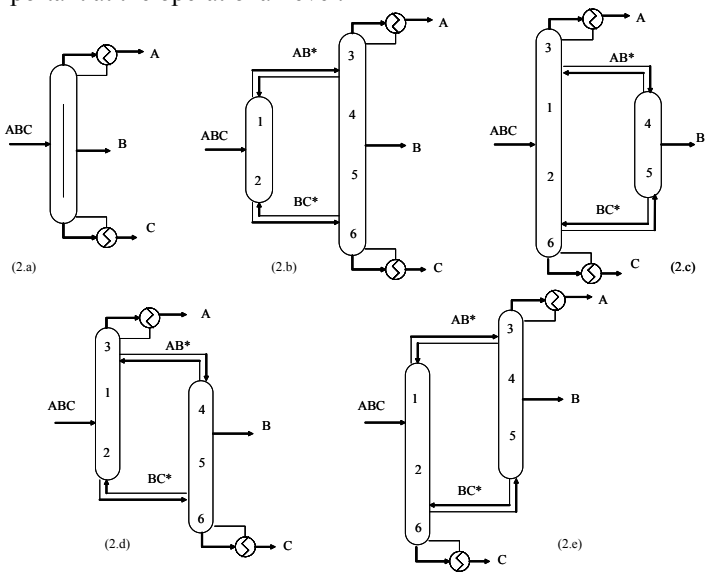


Figure 2. Thermodynamically equivalent alternatives for a three component FTC configuration.

For a given sequence of separation tasks it is of interest being able of generating all the different alternatives in which those tasks can be arranged in actual columns. Caballero and Grossmann showed [15] that the number of thermodynamically equivalent alternatives is equal to $2^{N^{\circ}\ of\ Thermal\ couples}$.
The differences in total cost between two thermodynamically equivalent configurations are not very important, at least when we compare with other structurally different sequences, but they have significant operational differences. Let us consider, for example, the Peltlyuk configuration of Figure 2b. In this configuration the vapor has to be transferred back, from the main or product column to the prefractionator, and forth, from the prefractionator to the main column. In these conditions the vapor flow rate has a difficult control, because neither of the columns has a higher nor lower pressure than the other column. Therefore, the bottom part of the prefractionator has to be at lower pressure than the lower section of the main column and the pressure in the upper section of the main column has to be lower than the pressure in the upper section of the prefractionator. The thermodynamically equivalent configurations (Figures 2d,e) have not this problem. Caballero and Grossmann[15] postulated that there is always a thermodynamically equivalent configuration in which it is possible to force the vapor to go from columns at higher to columns at lower pressure, and proposed a set of logical rules to extract these 'easier to control' configurations among all the thermodynamically equivalent alternatives.

4. Heat Integrated – Thermally coupled distillation sequences

Thermally coupled designs have important savings in energy and capital costs when compared with conventional column designs. However Rev et al[17] showed that for ternary mixtures the Petlyuk configuration is usually not superior to energy integrated systems even in energy savings. This is corroborated by Agrawal[18], in the case of three component mixtures. He found that the total heat duty of the double effect fully coupled configuration is always substantially lower than that of double-effect configuration, and that this results could easily be extrapolated to multicomponent mixtures. Although, the situation is much more complex in multi-component mixtures, again depending on the feed composition, available utilities, difference in volatilities etc; the optimal solution is expected to be an intermediate situation including thermal links as well as heat integrated columns, however, it has been also observed that in heat integrated designs the number of thermal couples decreases when compare with no-heat integrated sequences[16]. This fact complicate the design, because in general it is not truth the heuristics that says that the best no-heat integrated sequence is also the best when heat integration is considered. Therefore, the determination of the best columns sequence and heat integration must be considered simultaneously and not sequenctially.

References

1. Smith R, Linnhoff B. The Design Of Separators In The Context Of Overall Processes. *Chemical Engineering Research & Design.* May 1988;66(3):195-228.
2. Mizsey P, Hau NT, Benko N, Kalmar I, Fonyo Z. Process control for energy integrated distillation schemes. *Computers & Chemical Engineering.* 1998;22(Supplement 1):S427.
3. Petlyuk FB, Platonov VM, Slavinsk.Dm. Thermodynamically Optimal Method For Separating Multicomponent Mixtures. *International Chemical Engineering.* 1965;5(3):555-&.
4. Rudd H. Thermal Coupling For Energy Efficiency. *Chemical Engineer-London.* Aug 27 1992(525):S14-S15.

5. Triantafyllou C, Smith R. The Design And Optimization Of Fully Thermally Coupled Distillation-Columns. *Chemical Engineering Research & Design.* Mar 1992;70(2):118-132.
6. Fidkowski ZT, Agrawal R. Multicomponent thermally coupled systems of distillation columns at minimum reflux. *Aiche Journal.* Dec 2001;47(12):2713-2724.
7. Halvorsen IJ, Skogestad S. Minimum energy consumption in multicomponent distillation. 1. V-min diagram for a two-product column. *Industrial & Engineering Chemistry Research.* Feb 5 2003;42(3):596-604.
8. Halvorsen IJ, Skogestad S. Minimum energy consumption in multicomponent distillation. 2. Three-product Petlyuk arrangements. *Industrial & Engineering Chemistry Research.* Feb 5 2003;42(3):605-615.
9. Halvorsen IJ, Skogestad S. Minimum energy consumption in multicomponent distillation. 3. More than three products and generalized Petlyuk arrangements. *Industrial & Engineering Chemistry Research.* Feb 5 2003;42(3):616-629.
10. Agrawal R. Synthesis of distillation column configurations for a multicomponent separation. *Industrial & Engineering Chemistry Research.* Apr 1996;35(4):1059-1071.
11. Caballero JA, Grossmann IE. Generalized disjunctive programming model for the optimal synthesis of thermally linked distillation columns. *Industrial & Engineering Chemistry Research.* May 16 2001;40(10):2260-2274.
12. Caballero JA, Grossmann IE. Design of distillation sequences: from conventional to fully thermally coupled distillation systems. *Computers & Chemical Engineering.* Oct 15 2004;28(11):2307-2329.
13. Agrawal R, Xu JG. Separation Devices For Gas Mixing. *Aiche Journal.* Dec 1995;41(12):2585-2602.
14. Caballero JA, Grossmann IE. Generalized Disjunctive Programming Model For The Synthesis of Thermally Linked Distillation Systems. Paper presented at: European Symposium on Computer Aided Process Engineering -11, 2001; Kolding, Denmark.
15. Caballero JA, Grossmann IE. Thermodynamically equivalent configurations for thermally coupled distillation. *Aiche Journal.* Nov 2003;49(11):2864-2884.
16. Caballero JA, Grossmann IE. Structural considerations and modeling in the synthesis of heat-integrated-thermally coupled distillation sequences. *Industrial & Engineering Chemistry Research.* Dec 6 2006;45(25):8454-8474.
17. Rev E, Emtir M, Szitkai Z, Mizsey P, Fonyo Z. Energy savings of integrated and coupled distillation systems. *Computers & Chemical Engineering.* 2001;25(1):119.
18. Agrawal R. Multieffect distillation for thermally coupled configurations. *Aiche Journal.* Nov 2000;46(11):2211-2224.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Challenges with Advanced Control Technologies

Mario C.M. Campos,[a] Marcos V. C. Gomes,[b] and Herbert Teixeira[a]

[a]*PETROBRAS / CENPES / Engenharia Básica Abastecimento e Gás&Energia/ Automação, Equipamentos Dinâmicos e Confiabilidade (e-mail: mariocampos@petrobras.com.br).*
[b]*PETROBRAS / CENPES / P&D de Gás, Energia e Desenvolvimento Sustentável / Gás Natural / Célula de Otimização, Av. Horácio Macedo, 950 - Cidade Universitária, Ilha do Fundão, Rio de Janeiro, 21949-915, Brazil.*

Abstract

Model predictive control (MPC) has become a standard solution in the process industries, but there are still many open issues related to the acceleration of new implementations and the maintenance of good performance along the years. This paper will discuss some of these issues.

Keywords: performance assessment, regulatory control, advanced control system.

1. Introduction

The advanced control systems in oil & gas and petrochemical plants are an industrial reality (Qin and Badgwell, 2003). These advanced systems provide many advantages for the process units, as improved stability and safety, respect to constraints and higher profitability. PETROBRAS has been investing in the development of these systems for several years. Advanced control system is already a consolidated technology in its refineries with many model predictive controllers implemented (Zanin and Moro, 2004). To install and maintain these advanced systems with good performance is a great challenge. Its performance is influenced by instrumentation problems, bad tuning of the regulatory and advanced control, unreliable process dynamic models (Ender, 1993)(Kern, 2007), unmeasured disturbances, etc. This article will discuss the problems and challenges of advanced control in petroleum industries nowadays. It discusses some tools for diagnosis and tuning of the regulatory and advanced control. In spite of the several tools in the market that deal with industrial control, PETROBRAS has decided to invest on the development of its own tools and solutions in many situations, usually in association with some Brazilian universities. The goal of this paper is to show some challenges faced, solutions and results obtained in PETROBRAS facilities.

2. Regulatory Control Level

Process control aims to maintain certain variables within their desirable operational limits and could be visualized as a pyramid. In the base of this pyramid, the first level is the regulatory control, that uses PID controllers (Campos and Teixeira, 2006)(Ogata, 1982) and is configured in the digital systems (DCS - Distributed control system or PLC - Programmable logical controllers). In a second level, we have the advanced control systems that use for instance Model Predictive Control (MPC). This algorithm considers the interaction between control loops, and includes an optimization layer of the industrial plant. These algorithms are usually implemented in a process computer that

communicates with DCS or PLC systems by the use of OPC protocol (OPC, 2008). The outputs of this advanced control are usually the set points of the PID controllers. The architecture is conceived in such a way that if there is a failure in the advanced control level, the plant operation continues with the last PID set points in the DCS.

An advanced control system won't reach the expected benefits if is turned off constantly for the operators. Therefore, the instruments, valves and the regulatory control loops (PIDs) should operate appropriately. Hence, the performance of the regulatory control is fundamental for the success of the advanced control system. An industrial plant usually has hundreds of control loops, and less and less engineers to maintain the system. Therefore, the industries need tools to perform automatic analysis and diagnoses of the problems associated with the regulatory control. For example, these tools should be able to detect failures with the instrumentation (miscalibration, badly sizing, sensor noisy, out of scale, measurement resolution, etc.), non linear behavior in the process due to changes in the operational point, bad PID tuning (oscillation, stability, etc.) and control strategy problems (coupling between control loops, degrees of freedom, etc.).

There are several tools in the market that help engineers to maintain the regulatory control, but most of them require a well-trained engineers to interpret, analyze and define the correct actions, for instance: to change a control valve, tune PID controllers or to implement a new control strategy (decoupling, feedforward), etc. These engineers should also know very well the process in order to evaluate the better actions to be taken.

The great challenge for these tools will be to incorporate more "intelligence" to help engineers in the definition of the better actions. For instance, in certain case, only PID tuning could reach 80% of improvement in process variability reduction, and in some case, the process performance would improve only 10%. A lot of times in industries the engineer spends time and money with an action that won't bring great results. So, it is clear the importance of a tool that could perform the automatic diagnosis and assessment of the regulatory control (Farenzena et al., 2006). The most important features of this tool should be to have automatic ways to prioritize the actions for each process that might result in a better performance, and also to provide a standardized metric to compare different actions in different processes, even in different scales such as economical, environmental or safety (Harris, 1989)(Kempf, 2003)(Farenzena and Trierweiler, 2008). These features are a great development challenge for these tools.

Despite the several tools in the market, PETROBRAS and Federal University of Rio Grande do Sul (UFRGS) have developed their own tool, the software called “BR-PerfX”. BR-PerfX main purpose is to compute some universal key performance indicators that reduce the subjectivity in the analysis and help engineers in their assessments and decisions about problems affecting the regulatory control.

In order to face the PID tuning problem, PETROBRAS and the Federal University of Campina Grande (UFCG) developed the software "BR-Tuning" (Schmidt et al., 2008)(Arruda and Barros, 2003) “BR Tuning” is comprised by a group of techniques regarding open and close loop identification and the proposition of new tuning parameters. It communicates directly with the process automation system (DCS or PLC) using the OPC protocol.

As it was said previously, the challenge is to develop an "intelligent" layer that helps to make a diagnosis based on several indexes or indicators. The integration between different tools is also an important concern. The use of the OPC standard for the exchange of information could be an option. So, each tool could make available their indicators to others tools through OPC. This way, the engineers' work would be facilitated, avoiding losses of time and money.

The challenges in relation to controllers' tuning are associated mainly with the identification of the models, the determination of the non-linearities of the process, interaction between control loops, as well as defining the desired performance for each control loop.
There are some processes where the disturbances' pattern can change with the time, as in some off-shore petroleum platform. The slug flow can change its intensity for example due to changes in the gas-lift. So, we don't have a PID tuning tools and method that are good for all these different situations. In this case, it was developed an "intelligent" system that supervises the process plant and changes the PID tuning automatically when necessary. This control strategy is equivalent a "gain-scheduling" where the control performance (deviation between the process variable and the setpoint) is evaluated during a time, and the system decides what is the best tuning for that moment. All the possible values for the PID tuning are chosen off-line. This system was installed in several PETROBRAS's platforms. The figure 1 shows the system changing the PID tuning and the level PID performance. This project used a tool called MPA, which was developed by Catholic University of Rio de Janeiro (PUC-RJ) to PETROBRAS (Campos et al., 2008).
Another challenge is the development of non-linear controllers for some special cases, for example to pH control in certain plants, although PID will continue to be the algorithm more used in this regulatory layer control for several years.
Researches and developments for the regulatory control level are still necessary, and they can bring great economical earnings. For example, an application of these tools (evaluation, tuning and changes in control strategy) allows an increased of about 9% in the production of LGN (Liquefied Natural Gas) in a natural gas plant (Campos et al., 2007).

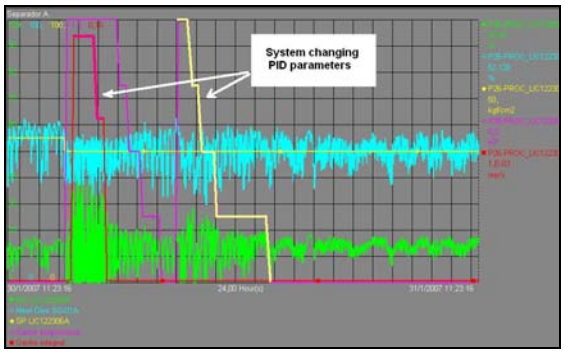


Figure 1– Performance of this control in production platform (1 day).

3. Advanced Control Level

The multivariable predictive controllers (MPCs) are powerful tools for the process optimization and are available in many industrial plants. This system can increase feed and preferred product rates, reduce energy consumption and waste material. These benefits are more visible in complex processes where challenging dynamic responses (significant time delays, non-minimum phase responses, control loop interaction, etc.) due to disturbances (feed flow and composition, energy integration, usefulness, etc.) that must be dealt with while taking into account process constraints and trying to pursue the best economic performance. As an example of the benefits achieved, figure 2 shows an increase of about 16% in the LPG yield due to the implementation of an APC system in a natural gas plant.

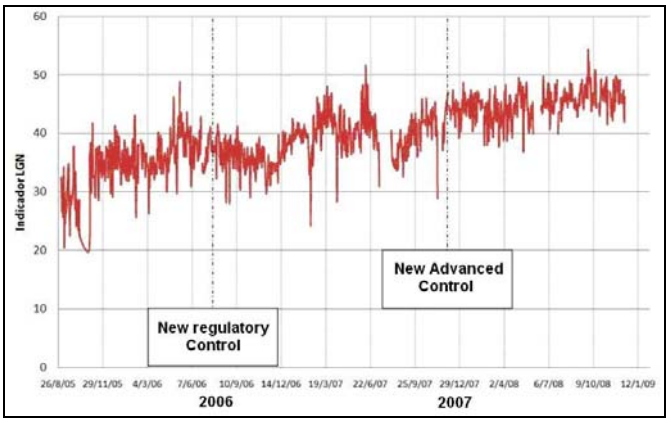


Figure 2– LPG yield increase in a natural gas plant due to MPC.

However, even if MPC systems are nowadays seen as a commodity, there is still much to be done, due to the significant gap between the recent MPC technologies development in the academy and those effectively used on industrial plants. Most industrial MPC applications are based on the most traditional approaches: linear algorithms based on step-response models obtained through traditional step tests.

MPC maintenance

MPC performance decay throughout time is a well-known and widely reported fact (figure 3). If no maintenance work is done, the operators end up turning them off. There are many causes for this behaviour:

- Changes in the units operational objectives;
- Equipments efficiency losses (fouling);
- Changes in the feed quality;
- Problems in instruments and in the inferences;
- Lacks of qualified personnel for the controller's maintenance.

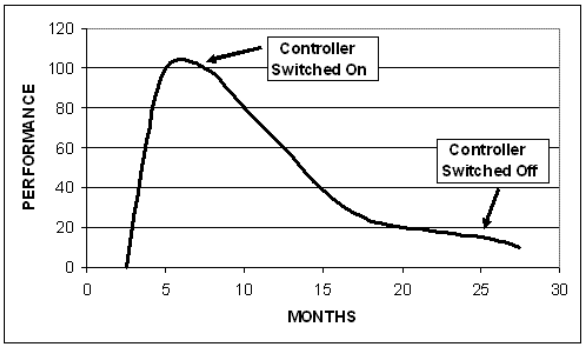


Figure 3 – Advanced Control Performance during the time.

Therefore, the first great challenge associated with MPC control is to have reliable tools to keep performance and diagnose problems. So, industry needs better tools to help maintenance personnel to answer the following questions:

Is advanced control system accomplishing their objectives? What is its performance? Is the process optimized? What are the benefits? How is the level of disturbances? What is operational factor of the controller? How are the operators adjusting the limits of the manipulated variables? Are manipulated variables very limited? What is the variability of the main controlled variable? Is the process operating close to the constraints?

It is necessary a tool not only to answer these questions, but the system point out the causes of the bad performance: bad models, bad controller tuning, inference problems, non-linearities, frequent changes in the operation point, new constraints not considered in the design?

Nonlinear models, Identification and Model mismatch

Many different and even sophisticated approaches have been proposed in order to allow MPC algorithms to cope with process nonlinearity, see Bequette (2007) for a recent review on the subject. However, despite all this effort, industrial NMPC applications are relatively few, and most of these are based on the simplest approaches.

One possible reason for that might be simply that the nonlinear behaviour is not known, and any lack of performance is seen as a typical model mismatch.

Another possibility might be that the nonlinear behaviour is known, but can not be easily determined with traditional plant tests. One way to overcome these problems might be the use of rigorous dynamic simulators, to improve the understanding of the process behaviour. Information obtained with dynamic simulation could be combined to the existing linear model in order to provide a reliable nonlinear one. Dynamic simulation might be useful also to find out the best way to characterize the observed nonlinearity. Once more, although there is availability of dynamic simulators, there is not much use of them in industrial applications.

Process identification of complex processes is still a hard task, where a significant part of the effort on MPC implementation is spent.

In order to address this problem, some commercial tools have been conceived in this decade for closed-loop identification. These tools are based on efficient ways to perform step tests allied to modelling strategies for minimization of the model order. While this approach has proved to be useful and promising, it is still a hard task to apply these techniques to complex processes, especially when dealing with noisy data. It seems to be a lot of space for development in this area.

Another interesting way to reduce implementation time can be the use of algorithms for automation of the plant test.

Tuning

MPC tuning is another interesting issue, where new technologies might help to reduce implementation time and also on the maintenance task.

Some interesting ideas have been proposed by Trierweiller and Farina (2003) that try to combine desired and achievable performances. However, the controller tuning still consume time and is critical points for controller performance. Normally, all MPC tuning methods consider a square controlled variables x manipulated variables matrix, but, in fact all controller has a rectangular matrix that means different tuning scenarios depending on which constraints is active.

Another big challenge is to reduce the application time and maintenance time. For this, it is believed that the main critical points are:

- Tools for the development of inferences:
  - Use of rigorous dynamic simulators, or statistical methods for better inferences using less laboratory analysis data.
- Dynamic models identification:
  - Automation of the identification tests, minimizing problems and loss of data;
  - Efficient tools for closed loop identification;
  - Characterization and identification of the non linearities of the process.
- Better tools for tuning the predictive controller:

- How to define the priorities in the several operating points of the controller and change automatically the tuning parameters. This activity is still done by trial and error in many industrial cases.

New advanced controllers that contemplate these aspects will help the users to implement and maintain these industrial systems.

4. Conclusions

This article has discussed some ideas about the improvement of MPC implementation, monitoring and maintenance. The first important issue to be taken into account is the requirement of expert tools regarding the regulatory control level. Aspects related to the MPC model identification and updating, and the conception of new tools for MPC tuning generation and assessment are seen as key points to help the users on implementation and maintenance tasks.

References

Arruda, G. and Barros, P., 2003, "Relay based gain and phase margins PI controller design", IEEE Transactions on Inst. And Meas. Tech., 52(5), 1548-1553.

Bequette, B. W, 2007, "Non-linear Model Predictive Control: A Personal Retrospective", The Canadian Journal of Chemical Engineer, 85, 408-415.

Campos, M. and Teixeira, H., 2006, "Controles típicos de equipamentos e processos industriais", Ed. Edgard Blücher, São Paulo.

Campos, M., Vasconcellos, L., Neto, J. and Souza, A., 2007, "Ganhos econômicos devidos à melhoria no controle De uma planta de processamento de gás natural", IV Congresso Rio Automação 2007, IBP, Rio de Janeiro.

Campos, M., Costa, L., Torres, A. and Schmidt, D., 2008, "Controle Avançado dos Níveis dos Separadores de Plataformas de Produção", CICAP – Congresso de Instrumentação, Controle e Automação da PETROBRAS, May, Rio de Janeiro.

Ender, D., 1993, "Process Control Performance: Not as good as you think", Control Eng., pp.180.

Farenzena, M. and Trierweiler, J., 2008, "Fronteiras e desafios em gerenciamento de malhas de controle", In: COBEQ 2008 - Congresso Brasileiro de Engenharia Química, Recife.

Farenzena, M., Teixeira, H. and Trierweiler, J., 2006, "Using the Inference Model Approach to Quantify the Loop Performance and Robustness". SICOP 2006 - International Workshop on Solving Industrial Control and Optimization Problems, Gramado.

Harris, 1989, "Assessment of Control Loop Performance", The Can. J. of Chemical Engineering, Vol. 67, pp856-86.

Kempf, A., 2003, "Avaliação de Desempenho de Malhas de Controle", Dissertação de Mestrado, Departamento Eng. Química, Universidade Federal do Rio Grande do Sul, UFRGS.

Kern, G., 2007, "Summiting with multivariable predictive control", Hydrocarbon Processing.

Ogata, K., 1982, "Engenharia de Controle Moderno", Ed. Prentice/Hall do Brasil.

OPC Foundation, 2008. Site: <http://www.opcfoundation.org/>, Accessed in: 03/17/08.

Qin, S. and Badgwell, T., 2003, "A survey of industrial model predictive control technology", Control Engineering Practice 11, 733-764.

Schmidt, D., Carvalho, R., Campos, M. and Santos, N., 2008, "BR-Tuning Ferramenta para sintonia de controladores PID", Primeiro CICAP – Congresso de Instrumentação, Controle e Automação da PETROBRAS, May, Rio de Janeiro.

Trierweiler, J. and Farina, L., 2003, "RPN tuning strategy for model predictive control", Journal of Process Control, Oxford- Inglaterra - Elsevier, v. 13, p. 591-598.

Zanin, A. and Moro, L., 2004, "Gestão da Automação Industrial no Refino", Rio Oil&Gás 2004, IBP, Rio de Janeiro.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

An Industrial Perspective on Pharmaceutical PSE

Peter Crafts

Principal Process Engineering, AstraZeneca Pharmaceuticals, Global Process R&D, Macclesfield, UK.

Abstract

If we, or one of our loved ones, become ill then access to safe and effective drugs to treat disease, alleviate pain and enable surgery becomes a high priority. AstraZeneca's business model is to discover, produce and market new pharmaceutical compounds that best meet these needs.

The industry average success rate for taking a new pharmaceutical compound through the clinical trials process and into the market is only ~8% however. This translates to a process development environment where many new processes compete for limited resource in stop-start projects, driven by the clinical trials programme. Within this environment PSE is most beneficial in focusing laboratory effort to areas most likely to succeed. Extrapolating small data sets to design sustainable and optimised processes is highly prized.

This paper explores the current role of PSE in AstraZeneca. Contemporary case studies are presented to illustrate both successes and areas where current knowledge and technology are insufficient. The examples cover:

- Solvent selection for reactions;
- Batch distillation for the exchange of process solvents;
- Crystallisation process design, solubility prediction and phase diagrams;
- The prediction of oiling-out during crystallization.

A common theme running through all these cases is the requirement for accurate phase equilibrium models, and this area will be explored in more detail.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Incorporating Molecular Signature Descriptors in Reverse Problem Formulations

Nishanth G. Chemmangattuvalappil, Charles C. Solvason,
Susilpa Bommareddy, Mario R. Eden
Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA

Abstract

The recently introduced concept of reverse problem formulation has been helpful for the solution of integrated process and product design problems from a properties perspective. A new algorithm has been developed that utilizes molecular property operators based on molecular signatures to obtain the molecular structures corresponding to the property targets estimated during the process design step. The algorithm allows property models based on different Topological indices (TI) and group contribution models to be represented on the same property platform.

Keywords: Molecular design, Signature description, Reverse problem.

1. Theoretical background

1.1. Reverse problem formulation

Reverse problem formulation (RPF) is a technique used to reduce the complexity of integrated process and product design problems. The application of RPF for identifying the property targets corresponding to optimum process performance can be found in the work published by Eden et al. (2004). In RPF, the integrated problem is broken down into two reverse problems. The first reverse problem provides the property targets to achieve optimum process performance and the second reverse problem is used to generate the molecular structures that meet the property targets. However, the current algorithms used for solving the second reverse problem are limited to very few properties which can be predicted using group contribution methods (GCM). Therefore, there is a need for an algorithm which is more general and can be extended to property models beyond group contribution methods.

1.2. Molecular signature descriptors

Molecular signature is a molecular descriptor used for representing the atoms in a molecule using the extended valancies to a pre-defined height (Visco *et al.,* 2002, Faulon *et al.,* 2003a). If G is a molecular graph and x is an atom of G, the atomic signature of height h of x is a canonical representation of the sub graph of G containing all atoms that are at a distance h from x. The signature of a molecule is defined as the linear combination of its atomic signatures. The detailed procedure for generating signature can be found in the work published by Visco *et al.* (2002) and Faulon *et al.* (2003a). Faulon *et al.,* (2003a) identified the relationship between TIs and signatures. If k is a constant, ${}^{h}\alpha_{G}$ is the vector of occurrences of atomic signature of height h and TI (root (${}^{h}\Sigma$)) is the vector of TI values calculated for each root of atomic signature:

$$TI(G) = k \cdot {}^{h}\alpha_{G} \cdot TI(root({}^{h}\Sigma)) \qquad (1)$$

2. Reverse problem formulation using molecular signatures

Since many existing QSPR expressions can be reformulated in terms of molecular signatures, algorithms based on signatures will be able to track a wide variety of property targets. In this approach, the property targets identified in the first RPF will form the input to the molecular design problem.

2.1. General problem statement

Identify the molecules with the best dominant property, which also satisfy a set of property constraints. The first step is to identify the QSAR/QSPR expressions corresponding to all the given properties/activities. In the next step, identify the heights of molecular signature corresponding to the TIs used in QSPR. If θ is the property function of property *P,* the property operator corresponding to *P* is estimated as follows. The property operator of a property is a function of the actual property which obeys linear mixing rules (Shelley and El-Halwagi, 2000):

$$\theta = f(TI) \tag{2}$$

$$TI = \sum_{i=1}^{N} {}^{h}\alpha_i \cdot TI(root({}^{h}\Sigma)) = \sum_{i=1}^{N} {}^{h}\alpha_i L_i \tag{3}$$

$$\Omega(P) = \sum_{i=1}^{N} x_i L_i \tag{4}$$

The dominant property, which is expressed in terms of the occurrences of atomic signatures, can be optimized subject to the property constraints. If Ω_j is the property operator corresponding to the dominant property and Ω_{ij} is the normalized property operator of molecule *i*, an optimization problem can be formulated as follows:

$$Max / Min \ \ \Omega_j \tag{5}$$

$$\Omega_j^{\min} \le \Omega_j \le \Omega_j^{\max} \tag{6}$$

2.2. Structural constraints

The solution of Eqs. 2-6 will provide a collection of signatures that give the best dominant property. However, to ensure that the collection of signatures will form a molecule without any free bonds, the *handshaking lemma* from graph theory is used (Trinajstic, 1992). According to this, in a graph, if *D(i)* is the number of degrees of each vertex *i* and *M* is the total number of edges:

$$\sum_{i=1}^{N} D(i) = 2M \tag{7}$$

In the reverse problem formulation, since the final molecular structure is unknown, it is not possible to know the number of bonds. Therefore, there should be a way to relate the number of signatures to the edges in the structure. Consider a graph without any circuits. If N_i is the total number of signatures, N_{Di}, N_{Mi} and N_{Ti} are the signatures with one double bond, two double bonds and one triple bond in the parent level, the following relationship has been developed between the number of signatures and edges:

$$M = \left[\sum_{i=1}^{N} x_i + \frac{1}{2} \sum_{i=0}^{N_{Di}} x_i + \sum_{i=0}^{N_{Mi}} x_i + \sum_{i=1}^{N_{Ti}} x_i \right] - 1 \qquad (8)$$

If the number of rings in the final structure is N_r, the following relation can be used:

$$M = \left[\sum_{i=1}^{N} x_i + \frac{1}{2} \sum_{i=0}^{N_{Di}} x_i + \sum_{i=0}^{N_{Mi}} x_i + \sum_{i=1}^{N_{Ti}} x_i \right] - 1 + N_r \qquad (9)$$

While designing molecules with both single and multiple bonds, the signatures with single and multiple bonds can form independent structures even with the theoretical potential for connectivity. This situation can be avoided by inserting a constraint which makes the formation of molecules only based on single bonds infeasible. If M_f is the number of edges in the signature combination with only single bonds:

$$M_f \geq 1 \qquad (10)$$

Since Eqs.8 or 9 ensures the entire connectivity of the combination of signatures, the aforementioned constraint will produce a molecular structure with both single and multiple bonds.

In order to differentiate between different types of atoms, vertex coloring has been used. The coloring function has to be selected based on the types of atoms and the nature of the final molecule. For signatures with double and triple bonds, the type of bond also should be the part of the coloring function. The coloring should start from the root atom to all atoms up to level *h-1*. For instance, in the design of an amine molecule, the coloring function is the degree of C or N atoms, the type of bond and the atom type. An example is shown in Figure 1 where the signature of C atom with NH_2 bond is colored:

CH_3-CH_2-C (NH_2) =CH_2 → C (=CC(C) N) → C4 (=C2N1C2)

Figure 1. Coloring of signatures

In a complete molecule, since each edge is shared by two vertices, the colors of the edge that joins the two vertices must be the same for both vertices. However, the order of colors will be different for both the vertices since the reading of color has to start from the root. In addition, the bond type should be consistent in the joined signatures. For instance, consider the signature in Figure 1. Here, the reading of colors are 4→ (=2), 4→N1 and 4→2. It can be seen that these orders are appearing in the reverse order while writing the signatures of attached atoms. The presence of both edges ensures that there is a link between the vertices. If $(l_i \rightarrow l_j)_h$ is one coloring sequence $l_i \rightarrow l_j$ at a level h:

$$\sum (l_i \rightarrow l_j)_h = \sum (l_j \rightarrow l_i)_h \qquad (11)$$

If the color sequence in one signature is $l_i \rightarrow l_j$ with $i=j$, then, there must be another signature present in the set of signatures with the same color sequence to complement the previous one. If η is the number of color sequences $l_i \rightarrow l_j$ on one signature with $i=j$ and x is the number of such color sequences and K is an integer, then:

$$\sum_{i=j} \eta_i x_i = 2K \qquad (12)$$

In order to form a connected tree, it must be ensured that the total number of signatures, where the degree of the vertex at a higher level is more than the degree at a lower level

should be less than the total number of vertices with the higher degree. If n_i is the number of child vertices with a higher degree than the parent vertex, i and j represents the child and parent colors:

$$\sum x_i n_i \leq \sum x_j \tag{13}$$

The dominant property function can be optimized subjected to the constraints. Integer cuts can be used to form other feasible solutions. An algorithm has been developed by Faulon *et al.* (2003b) to enumerate the molecules from the identified signatures.

2.3. Expression of group contribution (GC) models with signatures

In GCM, the property function of a compound is estimated as the sum of property contributions of all the molecular groups present in the molecule. The molecular signatures of sufficient height can be used to re-write GC expressions. This transformation will allow us to solve the property models based on TIs and GC models simultaneously. The height and number of the signatures used to write the molecular groups in the GC model depends on the number of atoms and the nature of the final molecule. Consider the design of a molecule with amine and alkyl groups. The maximum number of amine groups on each alkyl group and attached carbon atom on each amine group is limited to one. The potential groups available in GCM are (Marrero and Gani, 2001):

CH_2 NH_2 CH_2NH_2 $CHNH_2$ CNH_2 CH_3 CH_2 CH C

The root atom on every signature can be colored with two numbers - the first color is the number of neighboring C atoms and the second color is the number of neighboring N atoms. Signature of height 2 is required for complete coloring. All signatures with root N can be used to re-write different amine groups. If the neighboring C atom of an N root has three neighbors, then, the root N atom will be equivalent to a CNH_2 group. Similarly, all N root signatures can be assigned the property contributions of CH_2NH_2 and $CHNH_2$ groups based on the colors of their nearest neighbors. The signatures with root C atoms will form the alkyl groups. Here, the signatures with root C atom having N in any of its nearest neighbor should not be considered as a group since that signature has been taken care of in the amine groups. For the rest of the signatures with root C, property contributions can be assigned based on the numbers of neighboring C atoms.

2.4. Property models with different signature heights

In a molecular design problem with multiple property constraints, different TIs may be describing different properties of interest. It is also possible to have one OSAR/QSPR containing different TIs. If the signature heights of the TIs are different, the signatures corresponding to the largest height have to be enumerated first and the number of signatures of smaller height can be obtained from the higher signatures. Consider the situation where QSPR is based on signatures of heights one and two respectively. The property operator for property Y can be written in terms of signatures as follows:

$$f(Y) = \sum_i L_i{}^1 \alpha_i + \sum_j L_j{}^2 \alpha_j \tag{14}$$

If the height of the largest signature of interest is h and height of the lower signature is h-m, then the total number of each h-m level signatures can be obtained by adding the h level signatures under same color at level h-m. For example, if the signatures of interest are N1(C), N2 (CC) and N3 (CCC) which are signatures with height one. Now, signatures of height two with root vertex N can be divided into three sets S_1, S_2 and S_3.

The elements in these sets are signatures of height two with N vertex and vertex color 1, 2 and 3 respectively. The occurrences of signature N1(C) can be obtained as follows:

$$N(C):\ {}^{1}\alpha_{N(C)} = \sum_{S_1} {}^{2}\alpha_i \qquad (15)$$

The number of occurrences of other signatures of height one is similarly estimated.

3. Case Study: Design of substituent for a fungicide

The optimal substituent selection for dialkyldithiolanylidenemalonate (DD), which was originally solved by Raman and Maranas (1998) has been reworked here. The target properties are affinity (*log* (V_E)), mobility (*log* (μ)) and retention (*log* [*R/(1-R)*]). These properties are correlated to hydrophobic factor (Uchida, 1980), which is correlated to a first order molecular connectivity index (Murray *et al.,* 1975). In this work, an additional property constraint of toxicity (LC_{50}) is included. The GCM is used to predict LC_{50} (Martin and Young, 2001). The property models are shown in Table 1. The objective of this case study is to identify the alkyl substituents of DD that give maximum affinity subject to the constraints based on mobility, retention and toxicity. Here, substituents with multiple bonds are also considered. The first order connectivity index can be re-written in terms of molecular signatures of height two (Visco *et al.,* 2002) according to Eq. 16:

$${}^{1}\chi = \sum_{i=1}^{K_G}\left(\frac{1}{2}\sum_{u\in {}^{l}V_2({}^{h}\chi_i)}\left[\deg(u)\deg({}^{-1}\nu_\sigma(u))\right]^{-\frac{1}{2}}\right)x_i = \sum_{i=1}^{K_G} L_i x_i \qquad (16)$$

It is possible to form 120 signatures of height 2. The property targets and the normalized property operators are shown in Table 1. The molecular design problem can be written in terms of signatures using Eqs. 17-20.

Table 1. Property operators and targets

Property	*Property model*	*Property operator*	*Ref.*	*L.B.*	*U.B.*
Mobility	$-0.6984({}^{1}\chi)+2.0143$	$1.432(2.0143-\log(\mu))$	1	-0.3	0.3
Retention	$0.787({}^{1}\chi)-2$	$1.271\left(\log\left(\frac{R}{1-R}\right)+2\right)$	1	-0.3	1.0
LC50	$\exp\left(-\sum_{i=1}^{N} n_i\alpha_i\right)$	$-\log(LC_{50})$	1	-	0.04
Affinity	$0.5751({}^{1}\chi)-0.2942$	$1.739(\log(V_E)+0.2942)$	1	-	-

$$Max\ \Omega_{V_E} \qquad (17)$$

$$2.455 \le \sum_{i=1}^{120} h_i x_i \le 3.314 \quad \text{(Includes the constraints on mobility and retention)} \qquad (18)$$

$$3.129 \le \sum_{i=1}^{3} c_i x_i \quad \text{(Constraint on LC50)} \qquad (19)$$

$$\sum_{i=1}^{120} D_i x_i = 2\left[\left(\sum_{i=1}^{120} x_i + \frac{1}{2}\sum_{i=66}^{107} x_i + \sum_{i=108}^{120} x_i\right) - 1\right] \text{ (Handshaking lemma)} \qquad (20)$$

The first 65 signatures are the signatures with only single bonds in their parent atom, next 43 signatures have one double bond in their parent atom and the rest of the signatures have either two double bonds or one triple bond in the parent atom.
To differentiate among different types of carbon atoms, the vertices are colored with the degree of each carbon atom. Equations 11-13 are applied to ensure that the signatures will complement each other. For illustration, one of the equations from Eq. 11 is shown. Here, x_{68}, x_{98}, x_{99} *and* x_{100} are the occurrences of the signatures C3(≡C4), C4(≡C3C2), C4(≡C3C3) and C4(≡C3C4). It can be seen that the number of appearances of C3≡C4 is equated to the number of appearances of C4≡C3:

$$x_{68} = x_{98} + x_{99} + x_{100} \qquad (21)$$

Equations 17-21 are solved to obtain the number of appearances of each signature and the molecular structures are enumerated from the signatures using the algorithm provided by Faulon *et al.* (2003b). The best three solutions and the estimated properties are shown in Table 2.

Table 2. Final solution

Affinity	*Mobility*	*Retention*	*Toxicity*	R_1	R_2
				methyl	3-methyl-butyl
1.608	-0.296	0.603	0.035	methyl	2-pentyl
				ethyl	*sec*-butyl
1.598	-0.2833	0.5895	0.018	*Iso*-propyl	3-methyl-1-propenyl
1.5979	-0.2832	0.5893	0.015	Vinyl	4-methyl-2,3-butenyl

4. Conclusions

The newly introduced concept of molecular signature descriptor has been used as tool for the molecular design part of the general reverse problem formulation framework. The ability of signature to represent a wide variety of topological indices allowed the application of RPF into a variety of property targets. Since all property models are expressed as a function of occurrences of signatures, the problem can be solved for a single descriptor even if the TIs are described with signatures of different heights.

References

M.R. Eden, S.B. Jorgensen, R. Gani and M.M. El-Halwagi, (2004), *Chemical Engineering & Processing* 43, 595-608

J.L. Faulon, D. P. Visco Jr., R. S. Pophale, *J. Chem. Inf. Compt. Sci.* 2003a, 43, 707-720

J.L. Faulon, C.J.Churchwell, D. P. Visco Jr., *J.Chem. Inf. Compt. Sci*. 2003b, 43, 721-734.

J. Marrero and R. Gani, (2001), *Fluid Phase Equilibria*, 183/184, 183-208.

T.M. Martin and D.M.Young,(2001), *Chem.Res.Toxicol.* 2001,14,1378-1385.

W.J. Murray, L.H. Hall and L.B. Kier (1975), *J. Pharmaceutical Sci.* 64, 1978-81.

V. S. Raman, C. D. Maranas, *Computers & Chemical Engineering* 1998, *22* (6), 747-763.

M.D. Shelley and M.M. El-Halwagi, (2000), *Computers & Chemical Engineering*, 24.

N. Trinajstic, Chemical Graph Theory, 2nd Edition, CRC press, Boca Raton, FL, 1992

M. Uchida, *Pesticide Biochemistry and Physiology* 1980, 14, 249.

D. P. Visco Jr., R. S. Pophale, M. D. Rintoul, J.L. Faulon, *Journal of Molecular Graphics and Modeling* 20 (2002) 429-438

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Online optimizing control: The link between plant economics and process control

Sebastian Engell

Process Dynamics and Operations Group, Department of Biochemical and Chemical Engineering, TU Dortmund
Emil Figge-Str. 70, 44221 Dortmund, Germany, email s.engell@bci.tu-dortmund.de

Abstract

The idea of direct online optimizing control of chemical production processes is motivated and the handling of plant-model mismatch is discussed in some detail.

Keywords: Online optimizing control, RTO, plant-model mismatch, measurement-based optimization.

1. Introduction

In the design of chemical processing plants, it has become standard practice to employ rigorous stationary models to determine the sizes of the equipment as well as the nominal operating points. But as the models used for plant design do not represent the real process exactly and the processes are subject to a large number of perturbations, e.g. varying raw material compositions and ageing equipment, even if the operating regime was optimized for a rigorous plant model during plant design this will not lead to an optimal operation of the real plant. Moreover, adaptation to varying throughputs or new product specifications has to be performed increasingly often (Backx at al., 2000). This calls for continuous dynamic reactive and proactive adaptation of the operational degrees of freedom of the plant to changing conditions. For process stability and process safety, it is necessary to keep a large number of variables close to their set-points, a task which is usually realized by decentralized low-level PI/PID-controllers. From a process engineering point of view, the main purpose of automatic feedback control is to operate the plant such that the net return is maximized in the presence of disturbances and uncertainties, exploiting the available measurements. This has been pointed out in a number of papers (see e.g. Morari, et al., 1980, Narraway and Perkins, 1991, Skogestad, 2000a). A positive effect of tight feedback control is that the back-off from the operating constraints and product specifications can be reduced, leading to a more profitable operation, but not all process variables need to or should be controlled tightly.

Linear model-predictive control has be become the standard technique in the petrochemical industry to achieve high-performance multivariable control. In recent years, the focus in research has shifted to nonlinear model-predictive control (NMPC) often based upon rigorous dynamic process models. Advances in computer hardware and in numerical algorithms (e.g. Diehl et al., 2002, 2005, Jockenhövel et al., 2003) have rendered the application of NMPC employing high-order first-principles models in real time at real plants feasible. If NMPC controllers are implemented, the cost function can be formulated to represent plant performance rather than tracking performance, see e.g. (Zanin et al., 2000, Toumi and Engell, 2004, Bartusiak, 2005, Abel et al., 2005, Rolandi et al., 2005, Busch et al., 2005). In this paper, we discuss the idea of online optimizing

control with a focus on the issue of plant-model mismatch. We conclude with a discussion of some obstacles for the widespread implementation of this approach.

2. Classical feedback control vs. RTO vs. online optimizing control

The purpose of feedback control is to adapt the operational degrees of freedom of the plant such that the operational constraints and the product specifications are met and the economic performance is optimized. This can, to a certain extent, be achieved by regulating suitably chosen process variables to suitably chosen set-points (Skogestad, 2000a, 2000b). The effect of regulatory feedback control on a profit function J in the stationary situation can measured by the difference of the profit under the influence of disturbances $\underline{d}$ and its nominal value $J(u_{nom}, 0)$(Engell et al., 2005, Engell, 2007):

$$\Delta J = \left(J(\underline{u}_{nom}, 0) - J(\underline{u}_{nom}, \underline{d}_i) \right) + \left(J(\underline{u}_{nom}, \underline{d}_i) - J(\underline{u}_{opt}, \underline{d}_i) \right) + \left(J(\underline{u}_{opt}, \underline{d}_i) - J(\underline{u}_{con}, \underline{d}_i) \right).$$

The first term is the loss (or gain) if the manipulated variables are fixed at their nominal values when disturbances occur, the second term represents the effect of an optimal adaptation of the manipulated variables in the presence of disturbances, and the third term is the difference between the optimal adaptation and the one which is realized by the chosen feedback control structure. If the first term is large compared with the other ones then changing the manipulated variables will have little benefits, neither online optimization nor feedback control is useful here. On the other side, if the third term is large then online optimizing control should be used in addition to regulatory feedback control. Note that the terms appearing in the above formula can be determined, for a given control structure and known disturbances, by an optimization of the degrees of freedom for a stationary process model which is nowadays often available during plant design. This computation reveals the potential of online optimization for a given plant.

Often, the economically optimal operation of a plant is to maximize its throughput (see e.g. Price et al., 2003). The maximum throughput is restricted by the operational constraints of the units and by their ability to meet the constraints on product purities etc. Aske at al. (2008) have recently proposed a control scheme where an optimization of plant throughput is performed based upon the information that is available from the linear models used in local linear MPC controllers. This requires models that are accurate at the steady state while the exact steady-state gain is otherwise not needed to obtain a good control performance, so this may necessitate some additional modeling effort. The scheme was already partly implemented at a large gas treatment plant (Aske, 2009).

If the economic optimum is cannot simply be established by maximizing the throughput, the common approach is to introduce a higher level optimization layer where the operating points are optimized based upon rigorous stationary models, termed RTO – real-time optimization, see (Marlin and Hrymak, 1997) and the references therein. An RTO system is a model-based, upper-level control system that is operated in closed loop and provides set-points to the lower-level control systems in order to maintain the process operation as close as possible to the economic optimum. The general structure of an RTO system is shown in Fig.1. Its hierarchical structure follows the ideas put forward already in the 1970's, see e.g. (Findeisen, et al., 1980). The introduction of an RTO system provides a clear separation of concerns and of time-scales between the RTO system and the process control system. The RTO system optimizes the plant economics on a medium time-scale (shifts to days) while the control system provides tracking and disturbance rejection on shorter time-scales from seconds to hours.

A main disadvantage of this layered approach is the delay of the optimization which is inevitable because of the use of steady-state models. The time between successive RTO steps must be large enough for the plant to reach a new steady state after the last commanded move.

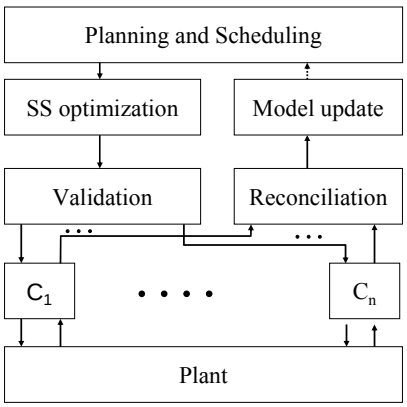


Figure 1: RTO control structure

If a step disturbance occurs in some unmeasured external input to the plant, then first the control system will regulate the plant (to the extent possible) to the set-points that were computed before the disturbance occurred. After all control loops have settled the RTO optimizer can be started, and after the results have been computed and validated, the control layer can start to regulate the plant to the new set-points. Thus it will take several times the settling time of the control layer to drive the plant to the new optimized mode of operation. In the first phase, the control system will try hard to maintain the previously optimal operating conditions even if without fixing the controlled variables to their set-points the operation of the plant would have been more profitable. If the disturbance persists for one sampling period of the RTO system plus one settling time of the regulatory layer, the use of the RTO system on the average recovers about half of the difference between the profit obtained by the regulatory system alone (with fixed set-points) and an online-optimizing controller that implements the optimal set-points within the settling time of the regulatory control layer. The combined RTO-regulatory control structure will work satisfactorily for infrequent step changes of feeds, product specifications or product quantities but it will provide no benefit for changes that occur at time scales below the RTO sampling period.

This situation calls for dynamic model-based online optimization as performed in nonlinear model-predictive control. In NMPC, the cost function is not restricted to a quadratic measure of set-point tracking, but economic goals can be included. This was probably first put forward and implemented at a real plant by Zanin et al. (2000). When dynamic online optimization is introduced, the functions of the RTO layer and of the regulatory control layer can be re-distributed. On the upper layer, key process variables are optimized dynamically over a prediction horizon in order to optimize the economic performance of the process while on the lower layer, these control moves are implemented (e.g. by flow controllers) and inventory-related and other secondary control loops are realized. Implementations at real plants have also been reported by (Toumi and Engell, 2004, Abel et al., 2005) and by Bartusiak (2005) for an industrial application. Recently, Ochoa et al. (2009) compared a direct optimizing control structure for a bioethanol plant to a conventional control structure in simulations and found significant benefits of online optimizing control.

3. Dealing with plant-model mismatch

A key aspect in the application of optimizing control is the effect of plant-model mismatch. The model that is used to dynamically optimize the performance over the prediction horizon never predicts the plant behavior exactly. The ability to cope with plant-model mismatch is one of the most important properties of well-chosen feedback con-

trol structures. In industrial reality, many controllers are implemented based only on very rough plant models but nonetheless function quite well. Feedback is usually introduced into online optimization over moving horizons by the following mechanisms:

- Restart of the optimization with a shifted horizon based upon an update of the estimated states
- Bias correction (extrapolation of the mismatch between the predictions and the measurements).

These two elements of feedback can be considered as relatively weak. The computation of the control moves relies completely on the assumed model of the behavior of the plant that is only shifted to compensate steady-state errors in order to ensure steady-state offset-free regulation. Figure 2 shows an example where an online optimizing control scheme fails temporarily due to plant-model mismatch. The process considered is a combined reaction-separation process, the Hashimoto-SMB process (Borren et al., 2006). The goal of the optimizing controller is to minimize the solvent consumption which dominates the variable production cost under hard constraints of the product purity and of the product recovery. The control scheme minimizes the solvent consumption plus a regularization term over a finite prediction horizon and employs the classical zero order prediction error extrapolation scheme, i.e. the constraints are modified by the difference of the predicted and the measured product purity and recovery at the most recent point in time.

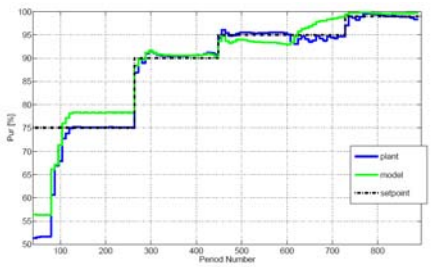


Figure 2: Predicted and measured product purities for the Hashimoto-SMB process with plant-model error mismatch (simulation) (Küpper and Engell, 2008)

It can be seen in Fig. 2 that the assumption of a constant error that results from plant-model mismatch is justified in the first two intervals while in the third interval, the measured purity hardly changes while the predicted purity reacts to the actions of the controller. The prediction error even changes its sign. The controller does not manage to keep the product purity above the constraint in periods 600 - 720.

For the Hashimoto SMB example, a careful analysis of the operating point chosen by the online optimizing controller revealed that the economically optimal operating mode is operation with a breakthrough over the recycle line, in contrast to the standard operation of SMB processes (Mazzotti et al., 1997). This operation mode is more efficient, but less robust than the conventional one. Enforcing the standard operation mode leads to robust control at the expense of an increased solvent consumption (Küpper and Engell, 2008). The example also shows that online optimization may point out opportunities for different operating modes that have not been considered in plant design.

Iterative solution and bias correction does not lead to optimality, because the model used in the optimization is not corrected and the optimality conditions are only satisfied for the model, not for the real plant. Moreover, it cannot be ruled out that due to plant-model mismatch, the plant is driven into a situation where feasibility can no longer be achieved, or the closed-loop system becomes unstable. The next paragraph discusses an approach that avoids these problems without being overly conservative.

3.1. Two-stage optimization for a set of plant models

An extension of moving horizon optimization that takes model errors into account explicitly is to formulate a min-max problem where the controller outputs are computed such that the cost function is minimized and the constraints are met for the worst possible model within a given set. This however may lead to unnecessarily conservative solutions, because in the formulation the presence of feedback that leads to a modification of the control inputs in the future after model errors or disturbances have influenced the system is ignored.

A better way to achieve non-conservative robust nonlinear model predictive control is to employ two-stage stochastic optimization, similar to the approach taken for scheduling problems e.g. in (Sand and Engell, 2004). The model uncertainty is represented by a set of possible plant models. The controlled variables are divided into two sets, *here and now* variables which are applied to all models (the first inputs), and *recourse* variables that can assume different values for each model in the set (the future inputs that will be chosen after additional information has been obtained). In the formulation proposed in (Dadhe, 2006, Dadhe and Engell, 2008), the cost function of the optimization problem is computed for the nominal model only, but the constraints must be met for all models in the set, by a suitable choice of the specific inputs for each model:

$$\min_{\substack{u_k \cdots u_{k+N_p-1} \\ u_k^{*b} \cdots u_{k+N_p-1}^{*b}}} J_k = \min \left(\phi \left(y_{k+N_p|k} \right) + \sum_{j=0}^{N_p-1} \ell \left(y_{k+j|k}, u_{k+j|k} \right) \right)$$

where the first term represents the terminal cost and the second term the incurred cost along the trajectory, under the constraints,

$$\begin{aligned} y_{k+j} &\in \mathcal{Y} \\ u_{k+j-1} &\in \mathcal{U} \\ \Delta u_{k+j-1} &\in \Delta\mathcal{U} \\ 0 &= y_{k+j} - f\left(\theta, y_{k+j-1}, \ldots, u_{k+j-1}, \ldots\right) \\ 0 &= u_{k+i} - u_{k+i}^{*b} \\ y_{k+N_p} &\in W \ominus \mathcal{W}(\alpha) \end{aligned} \qquad \begin{aligned} y_{k+j}^{*b} &\in \mathcal{Y} \\ u_{k+j-1}^{*b} &\in \mathcal{U} \\ \Delta u_{k+j-1}^{*b} &\in \Delta\mathcal{U} \\ 0 &= y_{k+j}^{*b} - f\left(\theta^{*b}, y_{k+j-1}^{*b}, \ldots, u_{k+j-1}^{*b}, \ldots\right) \end{aligned}$$

The upper index "*b" refers to the additional models that reflect the uncertainty about the plant dynamics. The index i assumes values from 0 … N_1, N_1+ 1 being the length of the first stage, and at the end of the prediction horizon, the outputs of all models must be driven to a target region *W* which is shrunk by an estimate of the model uncertainty. The two-stage formulation reflects the presence of feedback in a somewhat optimistic fashion because it is implicitly assumed that the true model will be revealed at a fixed point in time in the future. This robust non-conservative NMPC formulation was applied to sets of bootstrap models for neural net models of nonlinear systems in (Dadhe, 2006) but can be applied to any set of models that characterizes the model uncertainty. For a case study of a bioreactor model which exhibits strong nonlinearities and where the performance of a NMPC controller based upon a nominal model alone is insufficient in certain parts of the operating window, the two-stage multi-model formulation led to satisfactory control (Dadhe, 2006). Instead of considering the model uncertainty only in the constraints, the cost function can also be extended to the expected (average) value of the cost over all plant models or be formulated in a risk-conscious manner.

3.2. Achieving optimality in the presence of model uncertainty

How to use the available measurement information to achieve true optimality in the presence of plant-model mismatch is a challenging problem that has increasingly be

discussed recently (Gao and Engell, 2005, Kadam et al., 2007, Chachuat et al., 2008). The classical solution is to include a (usually small) set of model parameters into the estimation of the plant states and to adapt the model to the observations. This is efficient if the uncertainty can be tied to a limited number of crucial model parameters. If the plant-model mismatch can be eliminated by parameter estimation, optimality can be achieved for the real plant. However, if there is structural plant-model mismatch, then parameter adaption in a structurally incorrect model may make things worse, as demonstrated in (Chachuat et al., 2008).

If plant-model mismatch cannot be eliminated completely, there is a very promising scheme that addresses the issues of optimality and of feasibility: gradient adaptation (Tatjewski, 2002, Gao and Engell, 2005). The idea of Tatjewski that simplifies the Integrated System Optimization and Parameter Estimation Strategy first proposed by Roberts (1979) is to correct not only the bias but also the gradient of the cost function. The strategy requires the knowledge of the gradient of the real performance function (which is assumed to depend only on variables which are measured) with respect to the optimization parameters. Then convergence to the optimum can be assured without adaptation of model parameters. Gao and Engell (2005) extended this idea to the case where uncertainties are also present in the functions that relate the constraints to the degrees of freedom, e.g. product purity or conversion constraints. The drawback of the scheme is that the gradients of the performance indicator and of the constraints with respect to the controls have to be estimated online which may require additional control moves. Various schemes and for this are also discussed in (Gao and Engell, 2005).

4. Further critical issues

Lack of space prevents a detailed discussion of the many open issues in online optimizing control with economic cost functions here. Clearly stability is an issue, as discussed e.g. by Rawlings and Amrit (2008). From a practical point of view, modeling is the major bottleneck. To formulate, parameterize and validate dynamic models requires a major effort that has to be justified by the expected economic advantages. The recent trend towards building training simulators for all major new plants yields dynamic models for control almost for free, but then the simplification of these overly detailed models is a challenging problem.

A second practically very important limiting issue is that of reliability and transparency. It is hard, if not impossible to rule out that a nonlinear optimizer does not provide a solution which at least satisfies the constraints and gives a reasonable performance. While for RTO an inspection of the commanded set-points by the operators usually will be feasible, this is less likely in a dynamic situation. Hence automatic result filters are necessary as well as a backup scheme that stabilizes the process in the case where the result of the optimization is not considered safe. But the operators will still have to supervise the operation of the plant, so a control scheme with optimizing control should be structured into modules which are not too complex. Distributed online optimization has become an increasingly active research area recently (e.g. Venkat et al., 2008, Rawlings and Stuart, 2008, Gunnerud et al., 2009). Good interfaces to the operators that display the predicted moves and the predicted reaction of the plant and enable comparisons with their intuitive strategies are certainly essential for practical success. Other aspects are discussed in (Engell, 2007).

References

Abel, S., G. Erdem, M. Amanullah, M. Morari, M. Mazzotti and M. Morbidelli (2005). Optimizing control of simulated moving beds-experimental implementation. J. Chromatography A 1092 (1), 2-16.

Aske, E.M.B., 2009, Design of plantwide control systems with focus on maximizing throughput. Doctoral Thesis, Department of Chemical Engineering, NTNU Trondheim.

Aske, E.M.B., S. Strand and S. Skogestad, 2008, Coordinator MPC for maximizing plant throughput, Comput. Chem. Eng. 32, 195–204.

Backx, T., O. Bosgra and W. Marquardt, 2000, Integration of model predictive control and optimization of processes, Proc. IFAC ADCHEM 2000, 249–260.

Bartusiak, R.D. (2005). NMPC: A platform for optimal control of feed- or product-flexible manufacturing. Preprints Int. Workshop on Assessment and Future Directions of NMPC, Freudenstadt, 3-14.

Borren, T., J. Fricke, and H. Schmidt-Traub, 2006, Reactive liquid chromatography, in: Process Intensification by Integrated Reaction and Separation Operations, 191–240, Springer, Berlin.

Busch, J., M. Santos, J. Oldenburg, A. Cruse and W. Marquardt (2005). A framework for the mixed integer dynamic optimization of waste water treatment plants using scenario-dependent optimal control. Proc. ESCAPE 15, Elsevier, 955-960

Chachuat, B., B Srinivasan and D. Bonvin, 2008, Model parameterization tailored to real-time optimization. Proc. ESCAPE 18, Elsevier, 1-13.

Dadhe, K., and S. Engell, 2008, Robust nonlinear model predictive control: A multi-model non-conservative approach. Book of Abstracts, Int. Workshop on NMPC, Pavia, 24.

Dadhe, K., 2006, Dr.-Ing. Dissertation, Department of Biochemical and Chemical Engineering, Universität Dortmund, Shaker-Verlag, Aachen, 2007, (in German).

Diehl, M., H.G. Bock and J.P. Schlöder (2005). A real-time iteration scheme for nonlinear optimization in optimal feedback control. SIAM J. Control Optim. 43, 1714-1736.

Diehl, M., H.G. Bock, J.P. Schlöder, R. Findeisen, Z. Nagy and F. Allgöwer, 2002, Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations. J. Process Control 12, 577–585.

Engell, S., 2007, Feedback control for optimal process operation. J. Process Control, 17, 203–219.

Engell, S., T. Scharf and M. Völker (2005). A methodology for control structure selection based on rigorous process models. 16th IFAC World Congress, Paper Tu-E14-T0/6.

Findeisen, W., F.N. Bailey,M. Brdys, K.Malinowski, P. Tatjewski and A.Wozniak, 1980, Control and Coordination in Hierarchical Systems, JohnWiley & Sons.

Gao, W., and S. Engell (2005). Iterative set-point optimisation of batch chromatography. Comp. Chem. Eng. 29, 1401-1410

Gunnerud, V., B. Foss, B. Nygreen, R. Vestbø and N. C. Walberg, 2009, Dantzig-Wolfe decomposition for real-time optimization - applied to the Troll west oil rim. Proc. IFAC ADCHEM.

Jockenhövel, T., L.T. Biegler and A. Wächter (2003). Dynamic optimization of the Tennessee Eastman process using the OptControlCentre. Comp. Chem. Eng. 27, 15l3-l53l.

Kadam, J.V., M. Schlegel, B. Srinivasan, D. Bonvin and W. Marquardt, 2007, Dynamic optimization in the presence of uncertainty: From off-line nominal solution to measurement-based implementation, J. Proc. Control 17, 389–398.

Küpper, A. and S. Engell, 2008, Engineering of Online Optimizing Control - A Case Study: Reactive SMB Chromatography. Proc. 17th IFAC World Congress, Seoul, 2008, 964-969.

Marlin, T.E., and A. N. Hrymak, 1997, Real-time operations optimization of continuous processes. Proc. CPC-5 (J.C. Kantor, C.E. Garcia and B. Carnahan, Eds.), 156–164.

Mazzotti, M., G. Storti, and M. Morbidelli, 1997, Optimal operation of simulated moving bed units for nonlinear chromatographic separations. J. of Chromatography A 769, 3–24.

Morari, M., Y. Arkun and G. Stephanopoulos, 1980, Studies in the synthesis of control structures for chemical processes, Part I, AIChE Journal, 220–232.

Narraway, L.T., J.D. Perkins and G.W. Barton, 1991, Interaction between process design and process control: economic analysis of process dynamics. J. Process Control 1, 243-250.

Price, R.W., P. R. Lyman and C. Georgakis, 1994, Throughput manipulation in plantwide control structures, Ind. Eng. Chem. Res. 33, 1197–1207.

Rawlings, J.B. and R. Amrit, 2008, Optimizing process economic performance using model predictive control. Book of abstracts, Int. Workshop on NMPC, Pavia, 24.

Rawlings, J.B., and B.T. Stewart, 2007, Coordinating multiple optimization-based controllers: New opportunities and challenges, J. Process Control 18, 839-845.

Roberts, P. D. , 1979, An algorithm for steady-state system optimization and parameter estimation, Int. J. Syst. Sci. 10, 719-734.

Rolandi. P.A. and J.A. Romagnoli, 2005, A framework for online full optimising control of chemical processes, Proc. ESCAPE 2005, Elsevier, 1315–1320.

S. Ochoa, J.-U. Repke and G. Wozny, 2009, Plantwide optimizing control for the bio-ethanol process. Proc. IFAC ADCHEM.

Sand, G. and S. Engell: Modelling and solving real-time scheduling problems by stochastic integer programming. Computers and Chemical Engineering 28, 2004, 1087-1103.

Skogestad, S., 2000a, Plantwide control: the search for the self-optimizing control structure. J. Process Control 10, 487–507.

Skogestad, S., 2000b, Self-optimizing control: the missing link between steadystate optimization and control. Comp. Chem. Eng. 24, 569–575.

Tatjewski, P., 2002, Iterative optimizing set-point control - the basic principle redesigned. Proc. 15th IFAC World Congress, Paper T-Th-E16-3.

Toumi, A., and S. Engell, 2004, Optimization-based control of a reactive simulated moving bed process for glucose isomerization. Chem. Eng. Sci., 59, 3777– 3792.

Venkat, A.N., I.A. Hiskens, J.B. Rawlings, and S.J. Wright, 2008, Distributed MPC Strategies with Application to Power System Automatic Generation Control, IEEE Tr.on CST 16, 1192-1206.

Zanin, A.C., M. Tvrzska de Gouvea, and D. Odloak, (2000). Industrial implementation of a real-time optimization strategy for maximizing production of LPG in a FCC unit. Comp. Chem. Eng. 24, 525-531.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Integration of Planning and Scheduling and Consideration of Uncertainty in Process Operations

Zukui Li, Marianthi Ierapetritou

Department of Chemical and Biochemical Engineering, Rutgers University
98 Brett Road, Piscataway,NJ 08854, USA

Abstract

In the process industry, integration of planning and scheduling and addressing uncertainty have been identified as two of the most important factors that affect the efficient implementation of planning and scheduling operations for the production systems. In this paper, we review the existing work towards the integration of planning and scheduling and different methodologies of addressing uncertainty for planning and process scheduling decision making. The major challenges in planning and scheduling toward effective process operations are also presented throughout the paper.

Keywords: planning, process scheduling, integration, uncertainty.

1. Introduction

Production planning and scheduling are important subjects for modern industry, especially in the situation of facing major new challenges through increased global competition, greater regulatory pressures and uncertain prices of energy, raw materials and products. Although it has received wide research in planning and scheduling, there still exist very important and urgent needs in addressing the following two important issues for realistic application of production planning and scheduling systems. First, the traditionally strategy of generating planning and scheduling tasks separately often causes infeasible and suboptimal decisions, so it is of great necessity to integrate the planning and scheduling so as to generate really good operation decision for the production systems. Second, uncertainty appears in all the different levels of the industry from the detailed process description to multi-site manufacturing, it is necessary to systematically consider uncertainty so as to validate the use of mathematical models and preserve plant feasibility and viability during operations.
In this paper, we give a review of the existing work in the area of integration of planning and scheduling and in considering uncertainty in process operations. Different solution strategies for the integration of planning and scheduling is presented in section 2, followed by a detail description of uncertainty considerations for process scheduling and planning in section 3. Throughout the paper we present the major challenges towards effective implementation of planning and scheduling.

2. Integration of Planning and Scheduling

Production planning and scheduling belong to different decision making levels in process operations, they are also closely related since the result of planning problem is the production target of scheduling problem. The planning problem deals with longer term issues compared to scheduling with the emphasis being on the optimization of production capacity minimizing cost. In the process industry, the most commonly used

planning and scheduling decision making strategy follows a hierarchical approach, in which the planning problem is solved first to define the production targets and the scheduling problem is solved next to meet these targets. However, the main issue with this traditional strategy is the lack of communication between the two decision levels, i.e., the planning decisions generated might cause infeasible schedule subproblems. Thus at the planning level, the effects of changeovers and daily inventories are neglected, which tends to produce optimistic estimates that can not be realized at the scheduling level, i.e., a solution determined at the planning level does not necessarily lead to feasible schedules. Moreover, the optimality of the planning solution cannot be ensured because the planning level problem might not provide an accurate estimation of the production cost, which should be calculated based on the details of the scheduling problem.

Therefore, it is important and necessary to develop methodologies that can effectively integrate production planning and scheduling. However, since production planning and scheduling are dealing with different time scales, the major challenge towards the integration is dealing with the big computational complexity associated with the resulted optimization problem. To overcome the above difficulty, most of the work appeared in the literature aim at decreasing the problem scale through different types of problem reduction methodologies and developing efficient solution strategies as summarized by Grossmann et al. (2002) and Maravelias and Sung (2008). In this paper, we are aiming of reviewing the existing work in the area of planning and scheduling integration as classified in the following subsections.

2.1. Solving the Integration Problem through Hierarchical Decomposition

The first type of methods is based on decomposition in a hierarchical way through iterative solution procedure. In hierarchical decomposition of the integration problem, information is passed from the aggregate planning problem to a set of detailed scheduling problems which are solved independently using temporal decomposition. Thus, the problems that need to be solved include a relative simple planning problem and a series of scheduling subproblems. To ensure the feasibility and optimality of the solution, it is further necessary to develop effective algorithms to improve the solution using additional cuts in the planning level within an iterative solution framework (Papageorgiou and Pantelides, 1996; Bassett et al. 1996; Munawar and Gudi, 2005; Dogan and Grossmann, 2006).

2.2. Rolling Horizon Based Framework

The second type of methods, which is also called rolling horizon approach, consider a relative rough model for the far future planning periods in the integrated planning and scheduling model, whereas detailed scheduling models are only used for a few early periods and aggregate models are used for later periods. The production targets for the early periods are directly implemented, while the production targets for the later periods are updated along with the rolling horizon. Applications of this kind of strategy can be found in (Dimitriadis et al., 1997; Sand et al., 2000; Wu and Ierapetritou, 2007; Verderame and Floudas, 2008).

2.3. Campaign Planning Integrated with Periodic Scheduling

For the cases where there is no plant and market variability, campaign mode can be applied to generate an easy to implement and profitable process operations plan. In a periodic scheduling framework, the planning and scheduling integration problem is replaced by establishing an operation schedule and executing it repeatedly (Zhu and Majozi, 2001; Castro et al., 2003; Wu and Ierapetritou, 2004).

2.4. Integration with Surrogate Scheduling Model through Offline Analysis

Other than using the detailed scheduling model in the integrated planning, surrogate methods aim at deriving the scheduling feasibility and production cost function first and then incorporating them into the integrated problem. This avoids the disadvantage of large scale and complex model which directly incorporate the detailed scheduling model into aggregating planning model as shown in (Sung and Maravelias, 2007).

2.5. Mathematical Programming Based Decomposition of the Simultaneous Model

Except the different methods for the integrated planning and scheduling as summarized above, another approach is based on the study of the special structure of the mathematical programming model for the integration problem and aims at developing efficient decomposition techniques to solve the optimization problem directly. It is intuitive that the direct way for addressing the integrated planning and scheduling problems is to formulate a single simultaneous planning and scheduling model that spans the entire planning horizon of interest. However, when typical planning horizons are considered, the size of this detailed model becomes intractable, because of the potential exponential increase in the computation. Lagrangian relaxation is an approach that is often applied to models with a block diagonal structure as shown in Figure 1. In such models, distinct blocks of variables and constraints can be identified that are linked with a few "linking" constraints and variables. The major drawback of Lagrangian relaxation method is that there is duality gap between the solution of the Lagrangian dual problem and the solution of original problem, and the feasibility of the solution needs to be recovered through heuristic steps. Thus it is often only used as the bounding step in the branch and bound framework. However, the disadvantage of Lagrangian relaxation can be avoided by augmented Lagrangian relaxation (ALR) method, which has retrieved several applications in areas such as multidisciplinary optimization (Tosserams et al., 2008), power generation scheduling (Carpentier et al., 1996).

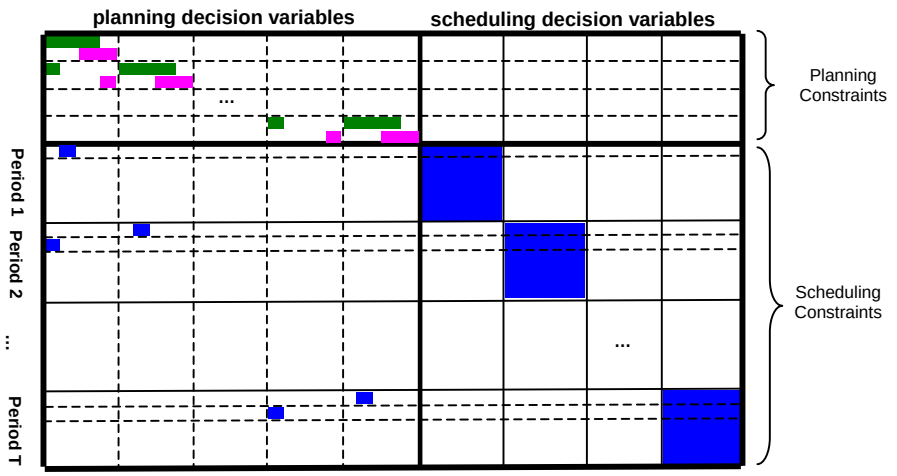


Figure 1. Structure of the constraint matrix of the simultaneous model

3. Planning and Scheduling Under Uncertainty

Traditionally, most of the work in the area of process operations deals with the deterministic optimization model where all the parameters are considered known. Along with the studies in deterministic planning and scheduling, consideration of uncertainties in these problems has got more attention in recent yeas because the inability of many

existing planning and scheduling systems to address general issue of uncertainty has been cited as a major reason for the lack of influence of planning especially scheduling research in industrial practice. It has also been recognized that having ways to systematically consider uncertainty is as important as having the deterministic model itself. Thus to achieve effective process operations, it is necessary to consider uncertainty that appears in all the different levels of the industry from the detailed process description to multi-site manufacturing.

In real plants, uncertainty is coupled with the planning and especially scheduling process since many of the parameters that are associated with these problems are not known exactly. Major uncertainties may exist in the planning and scheduling data including (Schnelle and Bassett, 2006): (i) sales and marketing (changes in product orders or order priority, changes of demand and pricing from forecasts); (ii) process (batch cycle times variability, yields, changeover times, recipe variations); (iii) production (e.g., batch or equipment failures, unavailable raw material, resource changes); (iv) new processes and products (ramp-up rates for cycle times and unit ratios). Uncertainty in those parameters are described using bounded interval form, probability distribution functions (continuous or discrete), and fuzzy description. Following the alternative description methods for uncertainty in process operations, different models and optimization approaches have been developed.

3.1. Process Scheduling Under Uncertainty

Methodologies for process scheduling under uncertainty aim at producing feasible, robust and optimal schedules. According to the different treatment of uncertainty, process scheduling methods can be classified into two groups: preventive scheduling and reactive scheduling.

3.1.1. Preventive Scheduling

Preventive scheduling generates scheduling policies before uncertainty occurs. Detail classification of preventive scheduling includes: stochastic scheduling, robust optimization method, fuzzy programming method and sensitivity analysis method. Stochastic scheduling is the most commonly used approach in the literature for preventive scheduling, in which the original deterministic scheduling model is transformed into stochastic model treating the uncertainties as stochastic parameters. Within the stochastic programming models we can distinguish the following categories: two-stage/multi-stage stochastic programming and chance constraint programming based approach (Orçun et al., 1996; Petkov and Maranas, 1997; Balasubramanian and Grossmann, 2002; Bonfill et al., 2004). Robust scheduling aims at building the preventive schedule to minimize the effects of disruptions on a performance measure. It tries to ensure that the predictive and realized schedules do not differ drastically while maintaining a high level of schedule performance (Vin and Ierapetritou, 2001; Lin et al., 2004; Janak et al., 2007; Jia and Ierapetritou, 2007). Fuzzy programming based method can be used in the case where probabilistic information is not available. Fuzzy set theory and interval arithmetic are used to describe the imprecision and uncertainties in process parameters (Balasubramanian and Grossmann, 2003; Wang, 2004; Petrovic and Duenas, 2006). An alternative way in preventive scheduling is using MILP sensitivity analysis, which is used to determine how a given model output depends upon the input parameters (Jia and Ierapetritou, 2004).

3.1.2. Reactive Scheduling

Reactive scheduling is a process of modifying the existing schedule during the process operation to adapt to changes (uncertainty) in production environment, such as disruptive events, rush order arrivals, order cancellations or machine breakdowns. For

this type of uncertainty there is not enough information prior to realization of the uncertain parameters that will allow a preventive action. There are a number of papers in the literature that focus on this kind of methodology (Huercio et al., 1995; Sanmarti et al., 1996; Honkomp et al., 1999; Vin and Ierapetritou, 2000; Ruiz et al., 2001; Méndez and Cerdá, 2003; Janak and Floudas, 2006). However, a major limitation of the existing reactive scheduling techniques is the response time, which might cause significant delay while the generation of a new schedule takes place. In Li and Ierapetritou (2008), a novel approach based on parametric programming is proposed to improve the efficiency of reactive scheduling and avoid the resolution of a complex optimization problem when uncertain event occurs during the scheduling period. Parametric programming serves as an analytic tool by mapping the uncertainties in the optimization problem to optimal solution alternatives. From this point of view, parametric programming provides the exact mathematical solution of the optimization problem under uncertainty. In the proposed method, reactive schedule is obtained from the solution of multiparametric programming problem which is solved ahead of time and covers all possible outcomes of future uncertainty as shown in Figure 2. The multiparametric programming problem is derived from a reactive scheduling formulation which integrates disruptive events (rush order and machine breakdown) as uncertain parameters in the process modeling. A restriction of the parametric method is that the computational complexity increases tremendously with the number of uncertain parameters, and thus it is appropriate for the problem with relative small number of uncertain parameters.

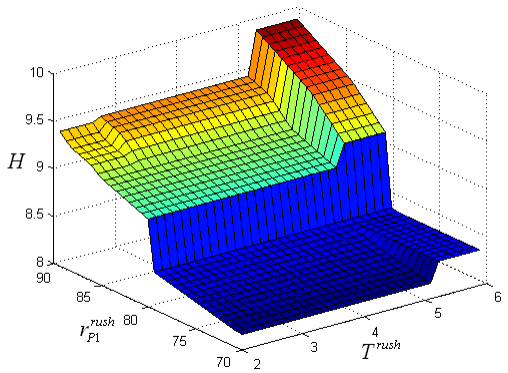


Figure 2. Parametric solution of the reactive scheduling problem, which gives the exact relationship between the new makespan and rush order information (new order amount, order arrival time)

3.2. Planning Under Uncertainty

For production and supply chain planning level problems, uncertainty is usually associated with product demand, which may lead to either unsatisfied customer demands or loss of market share or excessive inventory costs. Gupta and Maranas (2000) utilize a two-stage stochastic programming approach, where production is chosen here-and-now while distribution decisions are optimized in a wait-and-see fashion. This makes sense, since production tends to be the main contributor to lead times. Gupta et al. (2000) further investigate the trade-offs between customer demand satisfaction and production costs using a chance-constrained approach. Ryu and Pistikopoulos (2003) aim to deal with two problematic features in supply chain planning: (i) hierarchical decision structures with interdependence of the decisions of different agents and (ii) uncertainty in data. They develop a bi-level approach, which

captures the interdependence of the solutions and solve the problem using a parametric programming approach. Wenkai et al. (2003) describe a large refinery scheduling and inventory management model and introduce the concept of marginal value analysis, which identifies critical streams and operations. Neiro and Pinto (2003) extend this work to a set of refinery complexes, and also add scenarios to account for uncertainty in product prices. They demonstrate that nonlinear models reflecting process unit conditions and mixture property prediction can be used in multi-site planning models and that there is significant cost benefit in solving for the complex together rather than for the individual refineries separately. Wu and Ierapetritou (2007) address the simultaneous solution of production planning and scheduling problem through a hierarchical framework. The planning problem aggregates orders in the planning period and considers uncertainty utilizing a multi-stage stochastic programming formulation where three stages are considered with increasing level of uncertainty. The planning model includes material balances and time horizon constraints which involve a sequence factor to reflect the recipe complexity. Using a rolling horizon strategy, the production for the current stage is provided to the scheduling problem, which is solved using a continuous-time formulation. Finally, although very few work have addressed the problem of the uncertainty in integrated planning and scheduling problem (Dogan and Grossmann, 2006), it is still a great challenge since the deterministic integration of these two decision making stages is still a primarily unsolved problem for realistic cases.

4. Conclusions

In the existing planning and scheduling systems, there are few generic packages which can support the integration of planning and scheduling, and furthermore, the problem of the uncertainty is so far not well addressed in commercial systems. To achieve realistic and effective application of planning and scheduling decision making in process industry, the systematic consideration of integration and uncertainty issue is a necessary. Based on our experience, challenges towards effective implementation of planning and scheduling in the process industry include: a) the development of effective solution methods for integrated planning and scheduling problem; a) the development of new reactive scheduling methods; b) theoretical and algorithmic studies of methods for scheduling under uncertainty which can address a large number of uncertain parameters; c) the development of novel uncertainty analysis and optimization algorithms which can be applied in realistic large scale problems; d) the development of unified frameworks for integrated planning and scheduling under uncertainty. In summary, the integration of different decision making stages and the systematic consideration of uncertainty will have a tremendous impact on industrial sector increasing their production flexibility in the current dynamic and uncertain economic environment.

Acknowledgements

The authors gratefully acknowledge financial support from the National Science Foundation under Grants CTS 0625515, and 0224745.

References

J. Balasubramanian, I.E. Grossmann, 2002. A novel branch and bound algorithm for scheduling flowshop plants with uncertain processing times. Computers and Chemical Engineering 26, 41-57.

J. Balasubramanian, I.E. Grossmann, 2003. Scheduling optimization under uncertainty - an alternative approach. Computers and Chemical Engineering 27, 469-490.

M.H. Bassett, J.F. Pekny, G.V. Reklaitis, 1996. Decomposition techniques for the solution of large-scale scheduling problems. AIChE Journal 42, 3373-3387.

A. Bonfill, M. Bagajewicz, A. Espuna, L. Puigjaner, 2004. Risk Management in the Scheduling of Batch Plants under Uncertain Market Demand. Industrial and Engineering Chemistry Research 43, 741-750.

P. Carpentier, G. Cohen, J.C. Culioli, A. Renaud, 1996. Stochastic optimization of unit commitment: A new decomposition framework. IEEE Transactions on Power Systems 11, 1067-1073.

P. Castro, A. Barbosa-Povoa, H. Matos, 2003. Optimal Periodic Scheduling of Batch Plants Using RTN-based Discrete and Continuous-time Formulations: A Case Study Approach. Ind. Eng. Chem. Res. 42, 3346-3360.

A.D. Dimitriadis, N. Shah, C.C. Pantelides, 1997. RTN-based rolling horizon algorithms for medium term scheduling of multipurpose plants. Computers & Chemical Engineering 21, S1061-S1066.

M.E. Dogan, I.E. Grossmann, 2006. A Decomposition Method for the Simultaneous Planning and Scheduling of Single-Stage Continuous Multiproduct Plants. Ind. Eng. Chem. Res. 45, 299-315.

I.E. Grossmann, S.A. Van Den Heever, I. Harjunkoski, 2002. Discrete optimization methods and their role in the integration of planning and scheduling. AIChE Symp. Ser. 98, 150.

A. Gupta, C.D. Maranas, 2000a. Two-stage modeling and solution framework for multisite midterm planning under demand uncertainty Industrial and Engineering Chemistry Research 39, 3799-3813.

A. Gupta, C.D. Maranas, C.M. Mcdonald, 2000b. Mid-term supply chain planning under demand uncertainty: Customer demand satisfaction and inventory management. Computers & Chemical Engineering 24, 2613-2621.

S.J. Honkomp, L. Mockus, G.V. Reklaitis, 1999. A framework for schedule evaluation with processing uncertainty. Computers and Chemical Engineering 23, 595-609.

A. Huercio, A. Espuña, L. Puigjaner, 1995. Incorporating on-line scheduling strategies in integrated batch production control. Computers and Chemical Engineering 19, S609-S615.

S.L. Janak, C.A. Floudas, 2006. Production Scheduling of a Large-Scale Industrial Batch Plant. II. Reactive Scheduling. Industrial and Engineering Chemistry Research 45, 8253-8269.

S.L. Janak, X. Lin, C.A. Floudas, 2007. A new robust optimization approach for scheduling under uncertainty: II. Uncertainty with known probability distribution. Computers and Chemical Engineering 31, 171-195.

Z. Jia, M.G. Ierapetritou, 2004. Short-Term Scheduling under Uncertainty Using MILP Sensitivity Analysis. Industrial and Engineering Chemistry Research 43, 3782.

Z. Jia, M.G. Ierapetritou, 2007. Generate Pareto optimal solutions of scheduling problems using Normal Boundary Intersection Technique. Computers & Chemical Engineering 31, 268-280.

Z. Li, M.G. Ierapetritou, 2008. Reactive scheduling using parametric programming. AIChE Journal 54, 2610-2623.

X. Lin, S.L. Janak, C.A. Floudas, 2004. A new robust optimization approach for scheduling under uncertainty: I. bounded uncertainty. Computers and Chemical Engineering 28, 1069-1085.

C.T. Maravelias, C. Sung, 2008. Integration of production planning and scheduling: overview, challenges and opportunities. Proceedings Foundations of Computer-Aided Process Operations (FOCAPO 2008) 13-22.

C.A. Méndez, J. Cerdá, 2003. Dynamic scheduling in multiproduct batch plants. Computers and Chemical Engineering 27, 1247-1259.

S.A. Munawar, R.D. Gudi, 2005. A Multilevel, Control-Theoretic Framework for Integration of Planning, Scheduling, and Rescheduling. Ind. Eng. Chem. Res 44, 4001-4021.

S.M.S. Neiro, J.M. Pinto, 2003. Supply chain optimisation of petroleum refinery complexes. Proceedings of the fourth international conference on foundations of computer-aided process operations.

S. Orçun, K. Altinel, Ö. Hortaçsu, 1996. Scheduling of batch processes with operational uncertainties. Computers and Chemical Engineering 20, S1191-S1196.

L.G. Papageorgiou, C.C. Pantelides, 1996. Optimal campaign planning scheduling of multipurpose batch semicontinuous plants. 2. A mathematical decomposition approach. Industrial & Engineering Chemistry Research 35, 510-529.

S.B. Petkov, C.D. Maranas, 1997. Multiperiod Planning and Scheduling of Multiproduct Batch Plants under Demand Uncertainty. Industrial and Engineering Chemistry Research 36, 4864-4881.

D. Petrovic, A. Duenas, 2006. A fuzzy logic based production scheduling/rescheduling in the presence of uncertain disruptions Fuzzy Sets and Systems

D. Ruiz, J. Cantón, J.M. Nougués, A. Espuña, L. Puigjaner, 2001. On-line fault diagnosis system support for reactive scheduling in multipurpose batch chemical plants. Computers and Chemical Engineering 25, 829-837.

J.H. Ryu, E.N. Pistikopoulos, 2003. A bilevel programming framework for enterprise-wide supply chain planning problems under uncertainty. Proceedings of fourth international conference on foundations of computer-aided process operations.

G. Sand, S. Engell, A. Märkert, R. Schultz, C. Schultz, 2000. Approximation of An Ideal Online Scheduler for A Multiproduct Batch Plant. Comput. Chem. Eng. 24, 361-367.

E. Sanmarti, A. Huercio, A. Espuna, L. Puigjaner, 1996. A combined scheduling/reactive scheduling strategy to minimize the effect of process operations uncertainty in batch plants. Computers and Chemical Engineering 20, 1263-1268.

K.D. Schnelle, M.H. Bassett(2006). Batch Process Management: Planning and Scheduling. Batch Processes. E. Korovessi and A.A. Linninger, CRC Press: 389.

C. Sung, C.T. Maravelias, 2007. An attainable region approach for effective production planning of multi-product processes. AICHE Journal 53, 1298-1315.

S. Tosserams, L.F.P. Etman, J.E. Rooda, 2008. Augmented Lagrangian coordination for distributed optimal design in MDO. Int. J. Numer. Meth. Engng. 73, 1885-1910.

P.M. Verderame, C.A. Floudas, 2008. Integrated operational planning and medium-term scheduling of a large-scale industrial batch plants. Ind. Eng. Chem. Res 47, 4845-4860.

J.P. Vin, M.G. Ierapetritou, 2000. A new approach for efficient rescheduling of multiproduct batch plants. Industrial and Engineering Chemistry Research 39, 4228-4238.

J.P. Vin, M.G. Ierapetritou, 2001. Robust short-term scheduling of multiproduct batch plants under demand uncertainty. Industrial and Engineering Chemistry Research 40, 4543-4554.

J. Wang, 2004. A fuzzy robust scheduling approach for product development projects. European Journal of Operational Research 152, 180-194.

L. Wenkai, C.W. Hui, B. Hua, Z. Tong, 2003. Plant-wide scheduling and marginal value analysis for a refinery. Proceedings of fourth international conference on foundations of computer-aided process operations.

D. Wu, M. Ierapetritou, 2004. Cyclic short-term scheduling of multiproduct batch plants using continuous-time representation. Computers & Chemical Engineering 28, 2271-2286.

D. Wu, M.G. Ierapetritou, 2007. Hierarchical approach for production planning and scheduling under uncertainty. Chemical Engineering and Processing 46, 1129-1140.

X.X. Zhu, T. Majozi, 2001. Novel continuous-time MILP formulation for multipurpose batch plants. 2. Integrated planning and scheduling. Ind. Eng. Chem. Res. 40, 5621-5634.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling and design of distributed systems; methods and algorithms

Brian J. Sweetman,[a] Sukhraaj Basati,[a] Madhu Iyer,[a] Andreas A. Linninger[a,b]

[a]*University of Illinois at Chicago, Dept. of Bioengineering, 851 S. Morgan St, Chicago 60607, USA*
[b]*University of Illinois at Chicago, Dept. of Chemical engineering, 851 S. Morgan St, Chicago 60607, USA*

Abstract

Mathematical programming techniques to predict cerebrospinal fluid (CSF) flow fields and drug transport in the human brain are presented. In addition, advantageous use of distributed mathematical models to accelerate design and development of a novel volume sensor for hydrocephalus treatment is demonstrated. CSF flow in the brain can be measured analytically in three dimensions, but quantitative interpretation requires a distributed inversion problem. In the example of a volume sensor, medical imaging and rigorous mathematical analysis lead to optimal sensor design and optimal placement for highest sensitivity. Finally, prediction of drug transport in the brain leads to improved treatment options for patients suffering from neurodegenerative disorders such as Parkinson's and Alzheimer's. Infusion parameters such as flow rate and drug catheter position need to be optimized in three dimensions to achieve maximal therapeutic thresholds in the desired target area of the soft brain tissue. This example constitutes a three dimensional design problem with partial differential equation constraints. Our systematic modeling approach may improve the simulation and design of disease treatment options.

Keywords: brain physics, medical imaging, convection enhanced drug delivery.

1. Introduction

Recently in the process systems engineering community there is a trend towards analyzing biological systems or harnessing their capability for technical processes. Examples include energy systems based on algae or biomass, conversion of sunlight with biological organisms, or the integration of novel materials and processes into life sciences. In many of those processes the design challenges involve complex biological domains that are different from traditional process system engineering approaches. A specific challenge for mathematical models is the distributed nature of biological systems as observed in discrete parts of cell organelles, or the aggregation of cells in soft tissue, as well as the distributed nature of bioreactors. Design methods for distributed systems are less well developed than for lumped systems. However, we propose linking experimental data to accurate models of distributed biological systems to address the distributed nature of the process.

A multi-scale modeling approach is proposed to highlight new directions and challenges for distributed systems design with specific attention to biological systems. This method bridges the gap between analysis at the molecular level to the macroscopic level of a continuum. We propose large scale optimization methods to solve distributed

inversion problems. This leads to the discovery of unknown kinetic and molecular transport properties of biological systems.

2. Intracranial Dynamics

The first example of mathematical tools to quantify biological phenomena is the prediction of fluid transport and tissue stresses in the human brain. Intracranial dynamics describes the interactions of blood flow, CSF flow, and the deformation of porous brain tissue. While specific data about viscosity, density, and porosity about the brain and its fluids are known, there currently exists no comprehensive mathematical description of blood flow or its interaction with CSF and the flow field that results from the pulsating vasculature. While detailed aspects of normal pulsatile deformations of brain tissue can be quantified using existing medical imaging techniques, the complete picture of cause and effect is missing. Quantification of force interactions between blood, CSF, and the brain is necessary to understand normal dynamics as well as pathological conditions affecting the human brain.

With the advancement in imaging and computational analysis, understanding the complex interactions between vasculature and the soft brain tissue is feasible. Recently, our group has reconstructed the arteries and veins of a specific patient. This was achieved by manual segmentation of 1200 computer tomography (CT) images using Mimics 12.11. A precise three-dimensional surface of the main arterial system was obtained and is shown in Figure 1. A detailed description of our patient-specific image reconstruction methodology is discussed elsewhere (Linninger et al., 2007). In future work, the model of the cerebral vasculature will be used to predict blood pressures and vessel wall expansion.

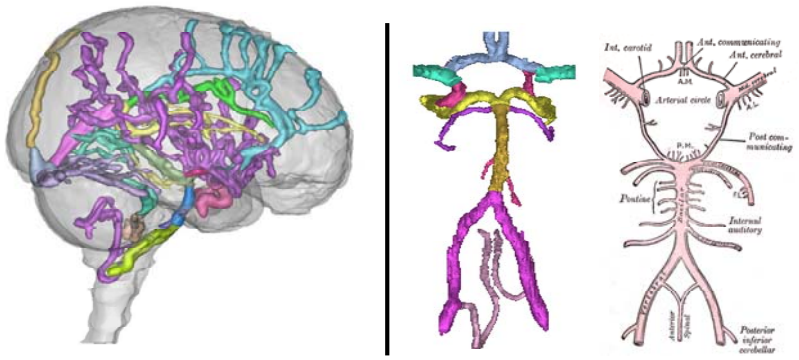


Figure 1. (Left) CT angiography images and Mimics were used to reconstruct a patient's cranial arteries/veins. (Right) Detailed view of reconstructed Circle of Willis and vertebral arteries juxtaposed with artist sketch (Gray, 1918).

Mimics was also used to develop a three-dimensional distributed model of brain tissue and CSF spaces. Reconstructed surfaces and volumes were meshed in ADINA 8.5, yielding a computational domain composed of tetrahedral balance envelopes. Boundary conditions such as constant CSF production and vasculature expansion in the brain parenchyma were included in the model. Linearized Navier-Stokes equations were solved numerically in ADINA-FSI 8.5. The three-dimensional model of CSF and brain tissue is used to predict CSF flow due to changes in cranial volume throughout each cardiac cycle. The model explains subtle changes in the flow field and quantifies other symptoms associated with pathologies of intracranial dynamics. Specific indicators we

discovered include the decrease in the prepontine stroke volume and the increase in the aqueductal flow as shown in Figure 2. We believe comprehensive models such as these will lend greater insight into normal intracranial dynamics and help us understand causes of certain brain disorders including idiopathic hydrocephalus. Future work will refine the model in the spirit of hierarchal decomposition to accurately capture the cerebral capillary network, which in the current models is grossly simplified. In a step by step process, it is believed that mathematical modeling will be able to uncover the functionality of the physical brain which to date is still elusive. More detail regarding the modeling approach and transport mechanisms are explained in Somayaji et al. (2008).

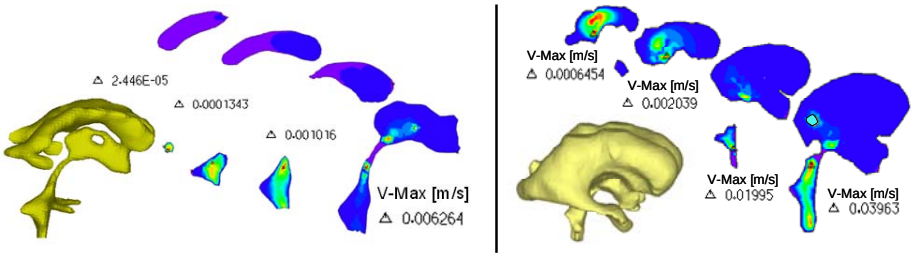


Figure 2. CSF velocity field in the cerebral ventricular system of normal subject (left) and patient with communicating hydrocephalus (right). The velocity field is displayed in sagittal cross-sections of 5 mm increments from the mid-sagittal plane. In hydrocephalus, aqueductal flow increases.

3. Computer Aided Design of a Volume Sensor

In classical biological research, there is a tendency to perform design optimization directly on animal models in an iterative trial and error process. Parameters are improved in several stages, scaling up from small animals to larger ones until a device or treatment for humans has been generated. Modeling of distributed systems accelerates the scaling procedure and design decisions typically achieved with animal experimentation. In this example, the implementation of these design decisions to fabricate a ventricular volume sensor is discussed.

Hydrocephalus, a disease associated with enlarged cerebral ventricles, may cause nausea, severe headaches, and possible death. To improve current treatment of this disease a sensor to measure ventricular volume is proposed. The proposed sensor is based on an impedance measurement system. This system renders a correlation between ventricular volume and the potential drop across sensor electrodes. The measured potential drop is due to a small electromagnetic field induced by two excitatory electrodes (Linninger, 2008c). We have shown that the potential drop is a function of the ventricular size according to the distribution of current within the cerebrospinal fluid and the adjacent volume (Linninger et al., 2008a). Due to the contrast in electrical conductance of brain tissue and CSF, the expansion of the cavity directly affects the voltage drop detected on the measurement electrodes. Sensor parameters such as electromagnetic field strength, electrode quantity, electrode spacing, catheter positioning, instrumentation frequency and gain, and biocompatibility are part of a large design problem. Instead of simultaneously optimizing all these parameters for each animal simultaneously in realistic conditions, the development process can be accelerated using image reconstruction and computational methods.

The proposed catheter is integrated into a computer model to predict ideal sensor measurements according to Maxwell's equations. The *in silico* model is used as a

virtual laboratory to test different catheter parameters such as spacing and position. The simulations help maximize device sensitivity and catheter position with respect to the enlarging ventricles. Once ideal sensor parameters are determined using computer simulations, the sensor is fabricated with microfabrication techniques. Figure 3 shows the design of an optimal sensor, a simulation to test sensor parameters, and the microfabricated sensor.

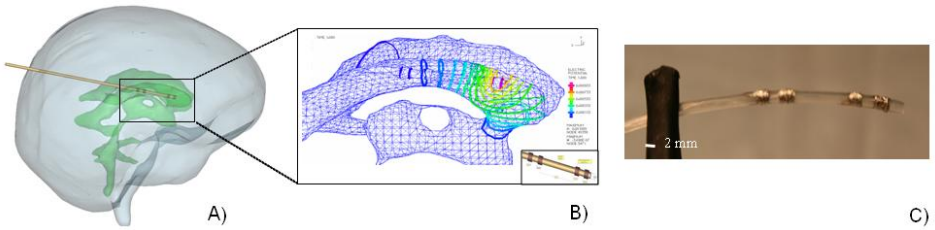


Figure 3. Computer Aided Sensor Design utilizing MRI reconstruction is shown in frame A. Simulations are done with Finite Element Analysis in frame B, and the microfabrication of the sensor is displayed in frame C.

4. Applications in Drug Delivery

4.1. Catheter Placement

The next case demonstrates the large potential of mathematical programming for rational therapy design in biomedical research. The objective is to maximize chemo-toxic agents in tumors. High concentrations of the chemotherapeutic agents in the brain can destroy the tumor and prevent reoccurrence. At the same time, the drug concentration should be low in other areas to spare critical regions in the cortex that would otherwise lead to detrimental side effects. A conceptual therapy plan given by the physician is depicted in Figure 4. The problem we wish to address with rigorous optimization is to precisely locate the optimal catheter position and design for administering the cytotoxin. Three catheter designs were available to choose from: single port, two-port, and a three-port catheter (Linninger et al., 2008b).

The distributed optimal drug delivery problem was solved with mathematical programming methods. Figure 4 compares achievable concentration profiles and distribution volumes in different brain regions with various port configurations. The best single-port catheter solution suggests an infusion site positioned near the tumor as expected. However, the single outlet configuration cannot deliver the desired therapeutic dosage to the white matter. The treatment only reached 46% of targeted volumes, I1, and I2, in sufficiently high concentration so that the risk for recurrence is unacceptable. A two-hole catheter distributes the drug over a wider range with two peak concentrations, the first near the tumor and the second inside the white matter. The program also determined the optimal distance between the outlet ports as 39 mm. Unfortunately, therapeutic thresholds were reached in 72% of the targeted tissue only. The optimal three-port catheter doses the tumor at therapeutic concentrations, I1, and simultaneously delivered effective drug concentrations into the white matter, I2. The chemotherapy delivered by a three-hole catheter dosed the tumor region as well as the adjacent white matter tracts (100% distribution volume). At the same time, toxic levels remained safe in the cortex (V1, V2). This illustration shows the viability of our optimization approach to provide pre-operative decision support to the physician to determine optimal infusion parameters such as catheter design and placement.

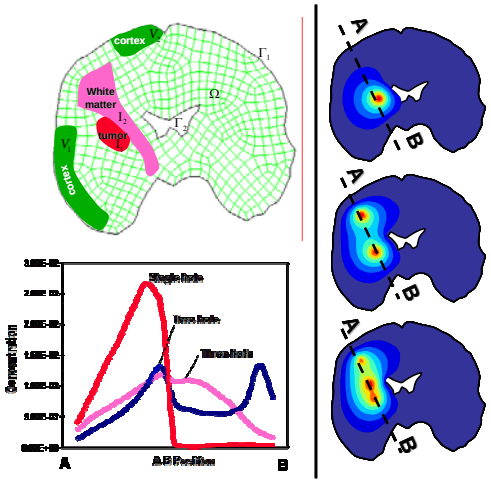


Figure 4. (Left) The conceptual treatment plan for glioblastoma targets the tumor (I1) and adjacent white matter (I2), without exceeding toxicity limits in the cortex (V1, V2). (Right) The optimal drug delivery plans with single, two- and three-port catheters.

4.2. Anisotropy of Brain

Molecular transport mechanism varies greatly in different regions of the brain. The effective diffusivity in the gray matter is essentially isotropic, whereas in white matter diffusion is anisotropic and heterogeneous (Linninger et al., 2008d). The difference in the transport properties are attributed to different properties of gray and white matter and can be quantified using diffusion tensor imaging (DTI) data of the brain tissue. DTI measures the apparent water diffusion coefficients in the brain tissue. Quantitative maps developed from DTI data can be used to characterize the degree of heterogeneity and anisotropy of the brain tissue.

Previously, data from DTI was primarily used to trace white matter tracts in the brain (Lebihan et al., 2001; Cao et al., 2003). In our work, we incorporate the three dimensional tensor quantities from the DTI data with first principles transport equations to predict drug distribution. The apparent water diffusion tensors are calculated by fitting the crude DTI data into a three-dimensional symmetric, positive-definite, second order tensor. For each voxel in the imaging plane, eigenvalues and eigenvectors of the corresponding diffusion tensor are calculated and visualized as hyper-ellipsoids with principal axes; the principle axes correspond to the lengths and orientations of the three eigenvectors as shown in Figure 5a. At each grid point in the brain mesh, six independent diffusion tensor elements are obtained by interpolation with the nearest neighbor. Simulations of drug distribution with anisotropic and heterogeneous brain models were performed using Fluent software. The drug concentration field and bulk flow field in an isotropic brain model were compared with an anisotropic brain model, visualized in Figure 5. Predicted drug distribution patterns in the anisotropic model were asymmetric and followed directional influences of the brain tissue.

5. Conclusions

This overview article highlights several applications of mathematical modeling for advancing biomedical research. We have shown three case studies in which rigorous mathematical modeling has led to insight into pathological conditions, design of

medical devices, and design of rational therapy parameters. These examples and the underlying systems methodologies provide a scientific approach for discovery about complex brain functions and is a contribution to a true interdisciplinary research effort. We hope that this overview provides a motivation for process systems engineers to venture into the field of biomedical research with methods that have been proven successful in the chemical industry and hold enormous potential to advance knowledge in biomedical research.

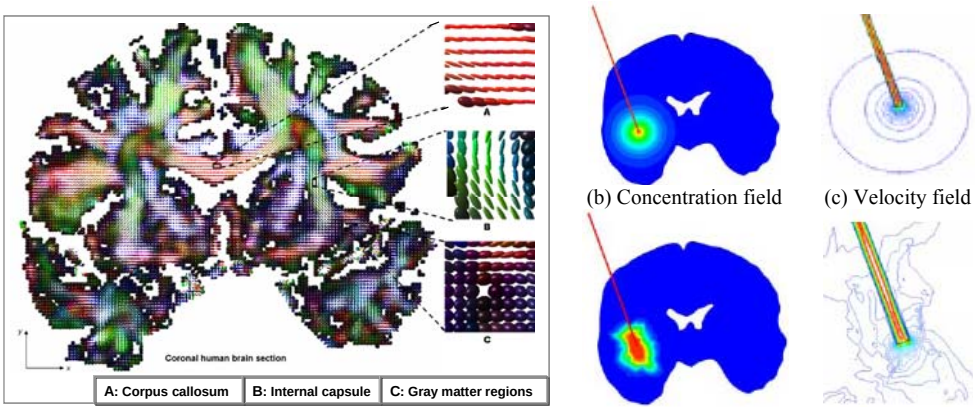


Figure 5. (a) Fractional anisotropy map of coronal human brain slices calculated from measured apparent water diffusion tensor field. Drug concentration field and bulk flow field using single hole catheter in homogeneous isotropic (b,c) tissue and anisotropic (d,e) tissue models.

Acknowledgments

The authors would like to gratefully acknowledge partial financial support of this project from NIH Grant 5R21EB004956, a grant from the Stars Kid Foundation, as well as NSF Grant CBET-0756154. Materialise is gratefully acknowledged.

References

Y. Cao, S. Whalen, J. Huang, K.L. Berger, M.C. DeLano, 2003, Asymmetry of subinsular anisotropy by in vivo diffusion tensor imaging. Human Brain Mapping, 20, 82-90.

H. Gray, 1918, Anatomy of the Human Body, 20th Ed., Lea & Febiger, Philadelphia.

D.L. Lebihan, J.F. Mangin, C. Poupon, C.A. Clark, S. Pappata, N. Molko, C. Hughes, 2001, Diffusion Tensor Imaging: Concepts and Applications. J. Magn. Reson. Imaging, 13, 534-546.

A.A. Linninger, M. Xenos, D.C. Zhu, M.R. Somayaji, R. Penn, 2007, Cerebrospinal fluid flow in the normal and hydrocephalic brain, IEEE Trans. Biomed. Eng., 54, 291-302.

A.A. Linninger, S. Basati, R. Dawe, R. Penn, 2008a, An Impedance Sensor to Monitor and Control Cerebral Ventricular Volume, Med. Eng. Phys., submitted September 2008.

A.A. Linninger, M.R. Somayaji, M. Mekarski, L. Zhang, 2008b, Prediction of convection-enhanced drug delivery to the human brain, J. Theor. Biol., 250, 125-138.

A.A. Linninger, 2008c, Monitoring and controlling hydrocephalus. US Patent 005440. 10 Jan. 2008.

A.A. Linninger, M.R. Somayaji, T. Erickson, X. Guo, R. Penn, 2008d, Computational methods for predicting drug transport in anisotropic and heterogeneous brain tissue, J. Biomech., 41, 2176–2187.

M.R. Somayaji, M. Xenos, L. Zhang, M. Mekarski, A.A. Linninger, 2008, Systematic design of drug delivery therapies. Comput. Chem. Eng., 32, 89.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Hybrid System Descriptions for Chemical Engineering Processes

Luís C. Oliveira-Lopes

FEQUI/UFU – School of Chemical Engineering, Federal University of Uberlandia, Av. João Naves de Ávila, 2121, Santa Mônica, CEP 38408-902, Uberlândia – MG , Brazil

Abstract

The hybrid description for process modeling has the advantage of allowing a more flexible description for dynamic models in the presence of state jumps and logical decisions. This work presents hybrid descriptions for processes of interest in Chemical Engineering, therefore, phenomena such as change in the number of stages/phases, processes failures, tolerant control, procedures or arrangements between stages, and jump on variables may be assessed by the verification of computational hybrid models. Cases of interest of process modeling, simulation, control and analysis are discussed.

Keywords: Hybrid systems, Modeling, Simulation, Control.

There are several entities, called model, which are able to mimic relevant features of a process or a situation one is interested in studying. A mathematical model is a simplified and abstract representation of part of the reality one desires to describe with a specific purpose. There is always a trade-off between the complexity of the description and the degree of accuracy it is designed to handle. Each different model assumes different hypothesis and in general, " it is not possible to maximize simultaneously generality, realism and precision."(Levis, 1968). A prediction that depends on very specific assumptions for validity of a model is called *fragile*, a prediction of a model is *robust* if it can be derived from a different model or from a general broader description. Besides that, in addition to make prediction, a model may be required to explain a certain phenomenon or even to incorporate knowledge in a no classical way. Many times it is necessary to combine approaches for a process representation in order to do that. For instance, the combination of a neural network model within a fuzzy framework (the neuro-fuzzy model) is often seen in the literature (Jin, 2000).

A model that presents a combination of approaches is often called as hybrid model. Nevertheless, the term hybrid system was originally proposed by Witsenhausen (1966) to describe a combination of continuous dynamical systems and discrete event systems. Hybrid Systems description comes from coupling between digital controllers, computers and subsystems built as finite-state machines, and plants modeled by differential or difference equations (Lazar, 2006). Therefore, it consists of a unified framework for describing processes evolving according to continuous, discrete dynamics, and logic rules. Hybrid systems are capable of selecting the activation of a specific mode when a discrete event occurs, handling discontinuities.

The term *hybrid* refers basically to a combination of two or more methods or objects. The expression *hybrid system* was proposed by Witsenhausen (1966) aiming to describe a combination between continuous dynamical systems and discrete event systems. Hybrid Systems come from coupling between digital controllers, computers

and subsystems built as finite-state machines, and plants modeled by partial or ordinary differential equations, or difference equations (Lazar, 2006). They can express continuous and discrete-time dynamics, jump discontinuity and transitions driven by logical commands. Hybrid systems can be composed by several operating modes, allowing them to show complex dynamical behaviors simultaneously. An ordinary general form of a hybrid model, describing a parameter concentrated model, can be represented by:

$$\frac{dx}{dt} = \psi_1\big(x(t), q^+(k), u(k)\big),$$
$$\psi_2\big(x(t), q^+(k), u(k)\big) \leq 0, \tag{1}$$

where, $x(t)$ is a continuous real-valued set, $q^+(k)$, and $u(k)$ are discrete real-valued variables.

Hybrid systems are a natural approach to formulate computer-controlled systems since they involve a physical process and a computer (controller). The classical approach to deal with computer-controlled systems has been through the use of sampled-data theory. However, this scheme can be encoded using a hybrid model. The hybrid model also captures a more general formulation, which is sometimes closer to real-time implementations in embedded control systems.

Hybrid systems are capable to retain an appreciable variety of dynamic behaviors. However, a hybrid system general formulation introduces high complexity for analysis purposes, application of techniques to design of controllers, simulation and validation. This fact led the research in hybrid systems to follow a particular path: instead of a general formalism, specific classes of systems are adopted to represent the process of interest.

To represent hybrid systems, one needs a description that is able to capture different types of continuous and discrete dynamics, is capable of modeling different ways in which discrete evolution affects and is affected by continuous evolution. In addition, it has to be composable and abstractable (Lygeros, 2004). Different description place more emphasis on different aspects, depending on the applications and problems they are designed to address. In this work, several classes of hybrid systems description are discussed and evaluated in chemical processes:

(a) *Piecewise affine systems*, PWA (Sontag, 1981, Johansson and Rantzer, 1998, Ferrari-Trecate et al., 2002) are described by:

$$\begin{cases} x(k+1) = A_i x(k) + B_i u(k) + f_i \\ y(k) = C_i x(k) + D_i u(k) + g_i \end{cases} \quad \text{for} \quad \begin{bmatrix} x(k) \in \mathbb{R}^m \\ u(k) \in \mathbb{R}^n \end{bmatrix} \in \Omega_i \tag{2}$$

where Ω_i are convex polyhedral in the input-state space, the variables: $y(k) \in \mathbb{R}^p$ denotes the output, $x(k)$ is the state, and $u(k)$ is the input set, at time k. PWA is a commonly used extension of linear systems to represent nonlinear and non-smooth processes.

(b) *Mixed-logical dynamical systems*, MLD. (Bemporad, 2004). This class of model was introduced by Bemporad and Morari (1999).

Mixed-logical dynamical systems are described by:

$$x\ k+1 = A_k x\ k + B_{1k} u\ k + B_{2k}\delta\ k + B_{3k} z\ k \quad (2.a)$$

$$y\ k = C_k x\ k + D_{1k} u\ k + D_{2k}\delta\ k + D_{3k} z\ k \quad (2.b)$$

$$E_{2k}\delta\ k + E_{3k} z\ k \geq E_{1k} u\ k + E_{4k} x\ k + E_{5k} \quad (2.c)$$

where $k \in \mathbb{Z}$; A_k, $B_{1k,2k,3k}$, C_k, $D_{1k,2k,3k}$, $E_{1k,2k,3k,4k,5k}$ are matrixes with appropriate dimensions. The domains are defined by:

$$\begin{aligned} x &= [\,x_c\ x_l\,]^T, x_c \in \mathbb{R}^{n_c}, x_l \in \{0,1\}^{n_l}, n = n_c + n_l \\ y &= [\,y_c\ y_l\,]^T, y_c \in \mathbb{R}^{p_c}, y_l \in \{0,1\}^{p_l}, p = p_c + p_l \\ u &= [\,u_c\ u_l\,]^T, u_c \in \mathbb{R}^{m_c}, u_l \in \{0,1\}^{m_l}, m = m_c + m_l \end{aligned} \quad (3)$$

Vector **x** denotes the states, which are composed with continuous states ("c" index), and logical binary states, 0-1 ("l" index). The vector **y** represents the model output data. The input vector **u** encloses continuous commands ("c" index) and binary commands, 0-1 ("l" index). The MLD system is said to be well posed if **δ**(k) and **z**(k) are uniquely defined in its domain, once **x**(k) and **u**(k) are specified. An interesting advantage is, in certain conditions, MLD systems can be equivalent to other model classes. This fact means that systems modeled by other classes can be converted into MLD systems that possess well definite techniques that favor its use for optimal and predictive control purposes.

In addition, in the MLD framework, logical propositions can be rewritten in the format of linear inequalities by means of the transformation of Boolean variables in binary integer variables. Previous operational knowledge of a process can also be added systematically under the form of inequalities, which form the set of constraints of an optimization problem that is base for optimal control problem synthesis (Krilavicius, 2006).

(c) *Linear complementarity systems*, LC (van der Schaft and Schumacher, 1998; Heemels et al., 2000).

In discrete time, this description uses a set of nonnegative orthogonal vectors (complementarity variables), $v(k), w(k) \in \mathbb{R}^q;\ 0 \leq v(k) \perp w(k) \geq 0$ and is given by:

$$\begin{aligned} x(k+1) &= Ax(k) + B_1 u(k) + B_2 w(k), \\ y(k) &= Cx(k) + D_1 u(k) + D_2 w(k), \\ v(k) &= E_1 x(k) + E_2 u(k) + E_3 w(k) + \xi \end{aligned} \quad (4)$$

(d) Extended *linear complementarity systems*, ELC (de Moor, 1999).

In this description one has an auxiliary variable $\varpi(k) \in \mathbb{R}^r$, and,

$$\begin{aligned} &x(k+1) = Ax(k) + B_1 u(k) + B_2 \varpi(k), \\ &y(k) = Cx(k) + D_1 u(k) + D_2 \varpi(k), \\ &E_1 x(k) + E_2 u(k) + E_3 \varpi(k) \leq \xi \\ &\prod_{j \in \phi_i} [\xi - E_1 x(k) - E_2 u(k) - E_3 \varpi(k)]_j = 0 \ \text{ for } \forall i \in \{1,2,\dots,p\} \end{aligned} \quad (5)$$

(e) *Max-min-plus-scaling* systems (MMPS) (De Schutter and van den Boom, 2001).

In this class of model, discrete event systems can be modeled by

$$x(k+1) = \mathcal{M}_x\big(x(k), u(k), \varpi(k)\big),$$
$$y(k) = \mathcal{M}_y\big(x(k), u(k), \varpi(k)\big), \quad (6)$$
$$\mathcal{M}_c\big(x(k), u(k), \varpi(k)\big) \leq c$$

where $\mathcal{M}_x, \mathcal{M}_y$ and $\mathcal{M}_c$ are MMPS expressions (i.e., expressions defined by the composition of max and min functions, sum, and multiplication by a scalar) in terms of the state *x*(k), the input *u*(k), and the auxiliary variables $\varpi(k)$, which are all real-valued.

Figure 1 shows the relationship between the model classes introduced up to this point in the work.

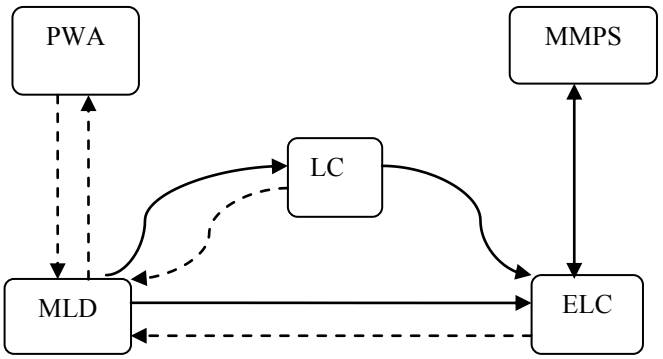


Figure 1 – Representation of links between a few classes of hybrid system descriptions. The arrow indicates that the class or origin is a subset of the model target class. Dashed line only indicates that special conditions need to be imposed to have that condition satisfied.

(f) *Switched Systems* (Liberzon, 2003)

The idea of switching system consists of a number of modes represented by differential equations and a set of switching rules. These systems arise in processes with variable structure and may represent networked control systems, in which information shared by continuous subsystems is updated in a discrete fashion, and embedded systems, in which computer software interacts with physical devices. A switched system is a particular kind of hybrid system that consists of several subsystems and a switching logic among them. The feature that distinguishes a switched system from a general hybrid system is that its continuous state does not exhibit jumps at the switching instants. Figure 2 shows the usual trajectory for a switched system with three switching instants.

Most approaches for representing hybrid systems use combinatorial or finite techniques such as graphs or automata extended by the capability to handle continuous elements. The continuous elements are *embedded* into a discrete dynamic structure. Examples of these descriptions are: hybrid automata, hybrid Petri Nets, and differential-algebraic models.

(d) *Hybrid automata* (Branicky et al., 1998).

Hybrid automata can model a large class of hybrid systems as they consider a discrete event system where the continuous dynamics in each discrete state are modeled by an arbitrary differential (difference) equation. A hybrid automaton is a dynamical system representation that is a collection of a set of discrete states, a set of continuous states, a vector field, a set of initial states, a domain, a set of edges, a guard condition, and a reset map. A usual way of representing this model is a directed graph.

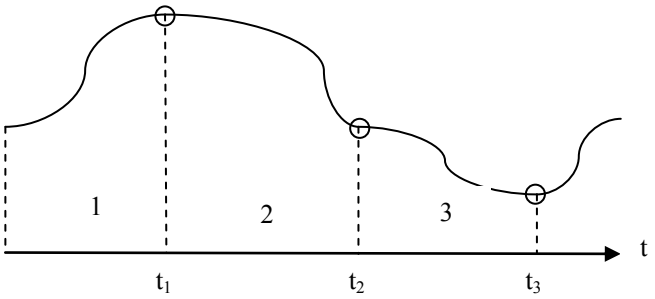


Figure 2 – Trajectory of a switched system. The indicated time instants are the discrete instant there is a change in the switching signal.

(e) Hybrid Petri Nets (Ghomria and Allab, 2007)

Petri nets were originally developed for modeling sequences of finite automata and they have been found to be useful for simulation and analysis of resource allocation problems of discrete event dynamic systems with both sequential and parallel elements. There is a number of extensions to the original concept of Petri nets in the literature including timed, hierarchical, multi-valued logical as well as stochastic elements. Figure 3 shows a simple Petri net description of a control valve.

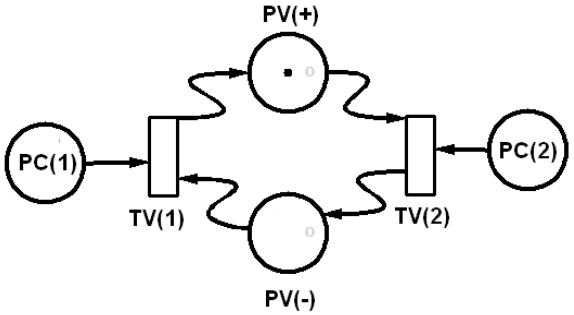


Figure 3 – Petri nets for input/output relationship in a process valve.

All these classes of hybrid systems introduced herein will be presented and discussed with examples of Chemical Engineering in the keynote session. Modeling and simulation of industrial plants with a hybrid description are investigated, tolerant control and fault propagation analysis will be presented and evaluated in this context, process control of hybrid process including logical decisions will be also evaluated by hybrid model based control strategies.

References

A. Bemporad and M. Morari, 1999. Control of systems integrating logic, dynamics, and constraints. Automatica 35 (3), 407–427.

A. Bemporad, 2004. Efficient conversion of mixed logical dynamical systems into an equivalent piecewise affine form. IEEE Transactions on Automatic Control 49 (5), 832–838.

B. De Schutter and T. J. J. Van den Boom , 2001. Model predictive control for max-plus-linear discrete event systems. Automatica 37 (7), 1049–1056.

B. De Schutter and B. De Moor. 1999. The extended linear complementarity problem and the modeling and analysis of hybrid systems. In P. Antsaklis, W. Kohn, M. Lemmon, A. Nerode, & S. Sastry (Eds.), Hybrid systems V, Lecture Notes in Computer Science, Vol. 1567 (pp. 70-85). Berlin: Springer.

D. Liberzon, 2003. Switching in Systems and Control, Birkhauser, Boston, MA. Volume in series Systems and Control: Foundations and Applications.

E. D. Sontag, 1981. Nonlinear regulation: the piecewise linear approach. IEEE Transactions on Automatic Control 26 (2), 346–357.

G. Ferrari-Trecate, F. A. Cuzzola, D. Mignone, and M. Morari, 2002. Analysis of discrete-time piecewise affine and hybrid systems. Automatica 38 (12), 2139–2146.

W. Heemels, B. De Schutter, and A. Bemporad, 2001. Equivalence of hybrid dynamical models. Automatica 37 (7), 1085–1091.

W. Heemels, J. M. Schumacher, and S. Weiland, 2000. Linear complementarity systems. SIAM journal on applied mathematics 60 (4), 1234–1269.

L. Ghomria and H. Allab, 2007. Modeling and analysis using hybrid Petri nets. Nonlinear Analysis: Hybrid Systems , 141–153.

M. Lazar, 2006. Model Predictive Control of Hybrid Systems: Stability and Robustness. Ph.D Thesis. Technische Universiteit Eindhoven, Netherlands

M. Johansson, A. Rantzer, 1998. Computation of piecewise quadratic Lyapunov functions for hybrid systems. IEEE Transactions on Automatic Control 43 (4), 555–559.

P. Bhowal, D. Sarkar, S. Mukhopadhyay e A. Basu, (2007) Fault diagnosis in discrete time hybrid systems – A case study. Information Sciences 177,1290–1308.

R. Levins, 1968. Evolution in changing environments. Princeton University Press.

T. Krilavicius, 2003. Hybrid Techniques for Hybrid Systems. PhD Thesis. University of Twente, Netherlands.

Y. Jin, 2000. Fuzzy modeling of high-dimensional systems: Complexity reduction and interpretability improvement. IEEE Transactions on Fuzzy Systems, 8(2), 212-221.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Integrating Refining to Petrochemical

Marcus Vinicius de Oliveira Magalhães

PETROBRAS, Av. República do Chile, 65, room 904, Rio de Janeiro, 20720-350, RJ, Brazil

Abstract

By definition petrochemical plants are integrated to refineries, since the most important feedstock for a petrochemical complex (i.e. naphtha) is produced at an oil refinery. Even when the petrochemical plant is natural gas based, a refinery must be in place to separate ethane and propane (and eventually butane), which are the actual raw materials for the steam cracking furnaces of the natural-gas-fed petrochemical plant. Although integration is an implied requirement between the two businesses, because one is downstream of the other, further integration is desirable for both of them. This work describes the potential benefits and challenges faced by this further level of integration.

Keywords: refinery, petrochemical, integration, benefits, challenges

1. Overview

The integration of refinery operations and production of petrochemicals can lead to a framework that adds value for both businesses. Substantial savings can be achieved, either for existing plants or new ones, by properly designing and/or operating a refinery and a petrochemical plant. These savings are derived from sharing common infrastructure (tank farm, utilities, terminals, pipelines …), leading to reduction in capital investments, or reduced inventories and energy/utility requirements if both sites are close enough to benefit from an integrated operation.

Low value components/products from one side, which are burned or sold at low prices, can be valuable raw material or blend stock component at the other side. This shipment of products entitles the increase of the aggregate margin due to the sourcing and security of feedstock supply and optimization of refined and petrochemical products slate, capturing synergies throughout the value chain.

Petrochemical industry often operates under cycles of profitability, with these cycles repeating every seven to ten years. Petrochemicals margins follow closely these cycles according to the operating rates. When the supply capacity is about to be met by the demand, operating rates are high and so are margins. Conversely, when operating rates are low, most likely due to the buildup of additional capacity, margins drop. Refinery profitability follows a different time and severity pattern, pretty much dependent on crude oil price. Hence, another beneficial aspect of refinery petrochemical integration is the potential enhanced profitability stability, which may be achieved with an integrated complex.

2. Benefits

The benefits of integrating refinery operations and petrochemicals production will be analyzed under the following standpoints: capital investments, operating costs, additional savings and environmental benefits.

2.1. Capital Investments

When setting up a new business unit, except for the case where processing steps can be eliminated with the integrated approach, it is not expected any major savings for the ISBL (Inside Battery Limits) class of costs. So, integration would not lead to any substantial savings at ISBL, when compared to the standalone situation.

On the other hand, integration can offer significant capital investment savings for the OSBL (Outside Battery Limit) class of costs. Typically, fewer tanks will be required, as opposed to the standalone situation, as the two plants may operate in a coordinated manner, implying that feedstock or byproduct will be pumped directly from one facility to the other. A direct consequence of a configuration with fewer tanks is the reduced inventory levels, thus enabling reduced working capital.

Refineries are typically large producers of utilities (e.g. cooling water, steam, electricity, fuel gas, flare, water and waste treatments etc). If some spare capacity is found for any of these utilities at the refinery side, substantial savings can be achieved by avoiding the construction of such a utility generator, when building a new petrochemical business unit integrated to a refinery.

Finally, additional savings can be made if the new business unit can operate sharing with the existing one several buildings, such as, administrative buildings, laboratory, control room, warehouses, workshops, and so on.

2.2. Operating Costs

Since raw material inventory to a downstream unit (or recycle to an upstream unit) can be drastically reduced or even eliminated, inventory costs can be substantially lowered if these products can be pumped directly between the integrated plants.

If the plants are in close proximity, some heat integration may also lead to potential energy savings, reducing utilities requirements and costs in both sides, and potentially, reducing CO_2 emissions (contributing to reduce impact on global warming). However, a high level of integration would require careful coordination from the control to the planning levels.

The two sites must plan their production on a collaborative manner in order to avoid shortage or surplus of feedstock and products on a short term basis, and also capture benefits from normal operations and spot opportunities. That implies the two business units must share sensitive information related to their operations to run their businesses smoothly and profitably under unplanned market disturbances, of either the supply of crude oil or products' (refinery fuels or petrochemical) demands.

Integration can also lead to significant reduction in maintenance costs. However, maintenance needs to be planned with care to minimize its possible greater impact on the profitability of integrated plants compared with standalone facilities. If the

shutdowns for maintenance of the two plants are not properly scheduled, upstream (downstream) unit's throughput may face reduction, or even worse, the unit may have to shut down, because the downstream (upstream) unit is not operating and the available storage space (inventory) may not be able to cope with the planned feed flowrate for the maintenance period.

Moreover, if both plants rely on the same labor resources to carry out the turnaround, it is likely that the maintenance costs will be higher, if the local market does not have the capacity to perform larger scope of maintenance tasks in parallel.

Finally, if staff can be shared, significant savings can be realized in administrative, operational (mainly OSBL) and technical services.

2.3. Additional Savings

Though sometimes not easy to quantify, integrated facilities, as opposed to stand-alone ones, may benefit from sharing common:

- Best practices related to engineering, operations, safety, environmental, reliability, automation, training and management;
- Fire fighting and medical crews, and;
- Purchase of goods and services.

2.4. Environmental Benefits

In addition, according to a study from AMEC assessing the benefits of industrial integration for facilities in Alberta, Canada, the environmental impact of integrated plants is significantly smaller than that of stand-alone counterparts, as listed below:

- 1.5 to 5.1 times lower air emissions (SO_2, NO_x, CO, CO_2 and PM);
- 1.4 to 2.2 times lower the ecological footprint;
- 1.1 to 1.6 times lower water consumption.

3. Enabling Raw Materials and Technologies

Some low value chemicals/streams available at one side may be of high value and interest at the other side. Two classical examples are cat-cracker off-gas and propylene produced at a refinery FCC (Fluid Catalytic Cracking). The off-gas is typically sent to the refinery fuel gas ring and burned at boilers and/or heating/reaction furnaces. The ethylene and ethane contents of this gas are around 15% to 20% vol. each. Ethylene is a petrochemical product itself and ethane a good quality feedstock for a steam cracking furnace, as it enables high ethylene yield. To extract both ethylene and ethane from the refinery off-gas there must be in place a gas plant (with cryogenic circuit) similar to the ones found downstream of pyrolysis furnaces.

Propylene, also produced at a refinery FCC, normally leaves the refinery in the LPG stream However propylene can be used in the manufacture of many petrochemical products (e.g. polypropylene, propylene oxide, cumene, acrylic acid, and so on). To extract the propylene out of the LPG one needs to rely on a sequence of distillation columns (splitters).

On the way back, from the petrochemical to the refinery, one can easily find other two classical examples of products, Hydrogen and Pygas (Pyrolysis gasoline) which are

valuable at the refinery side and not that much at a petrochemical plant. Pygas, if it is low benzene and olefins contents, is normally a very good component for gasoline blending that improves octane number.

In today's world refiners are facing tight environmental regulations imposing very low sulphur content in its fuels products. In order to cope with this never ending trend to make cleaner fuels, refineries must rely on hydrotreating processes, which require hydrogen as raw material, to reduce sulphur content down to the allowed limits. Although refineries normally have hydrogen generation capabilities in place, the trend is a growing demand for this ingredient. So, even if the hydrogen available at the petrochemical plant needs some further processing to meet purity requirements for use at refinery, it is clear that integration can offer a good opportunity to utilize the hydrogen, often burned at the petrochemical site.

In Brazil, the hydrogen consumption in five years time is estimated to exceed the current one by more than four times to meet the strict environmental specifications of the next generation of automotive fuels, mainly diesel and gasoline. Hence, the increased demand for hydrogen is likely to become a very attractive integration opportunity providing refineries with a supply framework of consuming excess hydrogen, if at low cost, from petrochemical plants combined with in-house production from hydrogen generation units.

Many other chemicals can play a more valuable role at the other side of fence, thus improving profitability of the integrated business. In general, olefins and aromatics, normally produced at a refinery FCC or Reforming unit, respectively, are of interest in a petrochemical plant. Petrochemical FCC, a specialized FCC suited to the production of petrochemicals, operating at higher severity and with a catalyst with higher ZSM-5 content than a conventional FCC, is becoming very popular technology these days as it is geared towards producing more petrochemicals at the expense of gasoline. Any technologies devised to increase the yield of light olefins and aromatics, and extract them, are more than welcome for both petrochemical and refinery businesses.

4. Challenges

As clearly indicated in the previous section, the integration of refinery and petrochemical facilities introduces a fertile terrain for the search of solutions which are good for both businesses. The purpose of this section is to highlight a few challenges of such integration.

The first challenge is related to devising a proper sequence of available technologies to produce petrochemicals from a C4 stream. The next two challenges are related to the development of cost-effective process technologies to recover valuable petrochemical building blocks (CO and ammonia) from refinery waste streams. The last challenge, that involves no actual molecules, has to do with the development of a mathematical tool to support the complex optimization of an integrated refinery petrochemical portfolio.

4.1. C4 Challenge

The C4 stream produced at an FCC encompasses a rich potential for petrochemical application. The table below shows typical composition and possible use of each component as a raw material.

Component	Application	Typical Composition (%)
Butadiene 1-3	PBR, ABS, SBR	1
Butene 1	Hexene 1	13
Butene 2	Propylene	26
Isobutene	ETBE, Isooctane	15
Isobutane	Alkylate	33
N-Butane	Maleic Anhydride	12

Although butadiene is a valuable petrochemical building block, as it plays the role of monomer or co-monomer of important polymers, its content in a typical FCC C4 stream is very low. The remaining components are found in significant composition and if they can be efficiently extracted from the bulk of the C4 they can be used as valuable raw material for a suitable process.

Given the very narrow range of the boiling points (from +3 ºC to -11 ºC) of the C4 components, separating them using a straightforward sequence of distillation columns is just not feasible in practical terms, since it is not economically viable. The challenge posed here is to design a suitable sequence of processes to recover, extract or transform each of the C4 components in a profitable manner, considering that for some processes, only one of the components is important, whereas the others are inert.

4.2. CO Challenge

Carbon monoxide is also a very important petrochemical building block. It can be used to produce many important chemicals, such as methanol, acetic or formic acid and n-butanol. CO can be obtained in a cost-effective manner, from an SMR (Steam Methane Reforming) unit, if a cheap source of natural gas is available. The search for cheap CO sources will ever be a key business driver for competitive production of chemical products derived from CO.

The exhaust gas from an FCC regenerator contains CO in low concentration, although the overall flowrate is large. This gas is normally destined to a CO boiler, where the CO is further oxidized to CO_2, producing additional steam, before finally vented to atmosphere. It sounds economically very promising to use this stream as a source of CO. However, the stream is typically found at very high temperature around (700 ºC) and very low pressure (slightly above atmospheric), and. it contains particulates and is also rich in N_2 and CO_2. Hence, recovering CO from this cheap source in a cheap manner is another challenge to refinery and petrochemical integration.

This particular challenge, once overcome, may be environmental beneficial, because CO_2 emissions will be reduced at this part of the plant.

4.3. Ammonia Challenge

Likewise, the development of cost effective process to recover ammonia from sour waters, largely produced in all refineries, will enable a cheap source of this important raw material for the petrochemical and fertilizer industry.

4.4. Modeling Challenge

The integration imposes many difficulties to deal with different cultures, requiring high degree of coordination and communication. Also, start-ups, shut-downs and normal operations are all more difficult compared with stand-alone facilities, and require more complex planning, scheduling, optimization and process control tools defined over much larger envelopes.

Due to the extensive range of petrochemical products and the very nature of the oil products and refinery processes, to devise an integrated business in an optimum manner will require a powerful modeling tool to support the selection process of the strategic opportunities in the long term horizon.

5. Conclusions

Refining and petrochemicals are faced with increasingly stringent environmental and regulatory constraints and fierce competition that have impacts on the profitability and sustainability of both industries. Integration across the supply chain will be a key factor to remain competitive and profitable, and to overcome the threats posed by the ever challenging market environment. In the new integrated context, the two industries will need to encourage the exchange of products, raw materials and utilities in an optimal way to make the best of all opportunities from both industries.

Acknowledgements

The author would like to express his gratitude to Dr. Montserrat Motas Carbonell and Dr. Panagiotis Karamertzanis for their suggestions and the review of this work.

References

Nexant, 2007, Benefits of Refinery/Petrochemical Integration, PERP Report, 06/07S7

C. Marcilly, 2001, Evolution of Refining and Petrochemicals: What is the Place of Zeolites, Oil & Gas Science and Technology – Rev. IFP, Vol. 56, No. 5, pp. 499-514

AMEC Earth and Environmental, 2007, A Study of Environmental benefits of Industrial Integration

P. Santos, 2008, Recentes Impactos da Indústria Petroquímica sobre o Parque de Refino, M.Sc. Thesis UFRJ/EQ, Rio de Janeiro.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Using systems engineering to reconstruct, analyze and redesign metabolism

Patrick F. Suthers[a], Vinay Satish Kumar[b], Alireza Zomorrodi[a], Sridhar Ranganathan[c] and Costas D. Maranas[a]

[a]Department of Chemical Engineering, [b]Department of Industrial Engineering, [c]Huck Institute of Life Sciences, The Pennsylvania State University, Department of Chemical Engineering, University Park, PA 16802-440, USA

Systems engineering computational techniques and in particular mathematical programming are increasingly being used to not only optimize metabolism for targeted overproductions but also for streamlining the process of metabolic model creation and curation. Key challenges in strain optimization include the selection of genes to eliminate (i.e., knock-out), add (i.e., knock-in) or modulate their expression levels in pursue of an overproduction target (e.g., a biofuel compound). Similarly, in metabolic model reconstruction and curation we seek to make model modifications so as agreement with experimental data is maximized. In both cases, this leads to the formulation of large mixed-integer, and in some case bilevel, formulations not very different from the ones encountered in process systems engineering.

Currently, there exists about twenty different microbial and eukaryotic metabolic reconstructions (e.g., *Escherichia coli, Saccharomyces cerevisiae*, *Bacillus subtilis*) with many more under development. All of these reconstructions are inherently incomplete with some functionalities missing or erroneously added due to the lack of experimental and/or homology information. Our group has introduced optimization-based procedures (i.e., GapFind and GapFill) to identify and eliminate network gaps in these reconstructions (Satish Kumar et al., 2007). These gaps are manifested as metabolites that are lacking either production and/or consumption pathways. Under steady-state conditions these metabolites form isolated parts of the model that cannot carry any metabolic flow. We repair these inconsistencies by allowing more reactions to operate reversibly and/or by expanding the list of cytosolic/transport reactions in the model. We demonstrated this procedure for the genome-scale reconstructions of *Escherichia coli* (Reed et al., 2003) *a*nd *Saccharomyces cerevisiae* (Duarte et al., 2004).

In addition to connectivity, genome-scale metabolic reconstructions are generally validated by comparing *in silico* growth predictions for different single gene deletion mutants under different carbon substrates with *in vivo* growth data. These comparisons results in two types of model-prediction inconsistencies; either the model predicts growth when no growth is observed in the experiment (GNG inconsistencies) or the model predicts no growth when the experiment reveals growth (NGG inconsistencies). We recently introduced the optimization-based framework, GrowMatch (Kumar and Maranas, 2009), to automatically reconcile GNG predictions (by suppressing functionalities in the model) and NGG predictions (by adding functionalities to the model). We used GrowMatch to resolve inconsistencies between the predictions of the latest *in silico Escherichia coli* (*i*AF1260) model (Feist et al., 2007) and the *in vivo* data available in the Keio collection (Baba et al., 2006). This led to an improvement in the

consistency of *in silico* vs. *in vivo* predictions from 90.6% to 96.7%. Specifically, we were able to suggest consistency-restoring hypotheses for 56/72 GNG mutants and 13/38 NGG mutants. For many cases, GrowMatch identified fairly non-intuitive model modification hypotheses that would have been difficult to pinpoint through inspection alone.

Armed with GapFill and GrowMatch, we explored their application during the generation phase of a new genome-scale model, as opposed to *a posteriori* improving existing ones. We applied these frameworks during the reconstruction of the genome-scale metabolic model of *Mycoplasma genitalium*. *i*PS189 (Suthers et al., 2009). *M. genitalium* has received extensive attention in the literature as it is widely used as a model organism in many minimal genome studies. Key challenges for *M. genitalium* included estimation of biomass composition, handling of enzymes with broad specificities and the lack of a defined medium. Our computational tools were employed to identify and resolve connectivity gaps in the model (*i.e.*, GapFind and GapFill) as well as growth prediction inconsistencies with gene essentiality experimental data (*i.e.*, GrowMatch). The resulting curated model, *M. genitalium* *i*PS189 (262 reactions, 274 metabolites) was 87% accurate in recapitulating *in vivo* gene essentiality results for *M. genitalium*. Specifically, we found that the model correctly identified 149 out of a total of 171 essential genes (i.e., specificity of 87%) and 16 out of a total of 18 non-essential genes (i.e., sensitivity of 89%). This level of agreement not only meets, but exceeds thresholds for metabolic model quality put forth in the literature. Additional *in vivo* gene essentiality studies using a fully defined medium could usher a more accurate elucidation of the true metabolic capabilities of *M. genitalium*, as well as suggest improvements to the model.

Motivated by the paucity of single-gene deletion data for metabolic model curation we explored the generation of model predictions for multiple-gene deletion mutants. Note that essential genes are defined as genes whose deletion is lethal. By analogy, synthetic lethals refer to pairs of non-essential genes whose simultaneous deletion is lethal (i.e., no biomass formation). One can extend the concept of lethality by considering gene groups of increasing size where only the simultaneous elimination of all genes is lethal while individual gene deletions are not. We developed an optimization-based framework for the targeted enumeration of multi-gene (and by extension multi-reaction) lethals for genome-scale metabolic models (Zomorrodi et al., 2009). Specifically, these frameworks are applied to the latest *i*AF1260 model of *E. coli* (Feist et al., 2007) leading to the exhaustive identification of all double and triple lethals as well as targeted identification of higher-order ones. Graph representations of these synthetic lethals revealed a variety of motifs ranging from hub-like to highly connected sub-graphs. These results provided a birds-eye view of the redundant mechanisms available for redirecting metabolism and revealed complex patterns of gene utilization and interdependence. The procedure also opened the door for the use of falsely predicted synthetic lethals as guidelines for metabolic model testing and curation. By analyzing the functional classifications of the genes involved in synthetic lethals (such as carbohydrate or nucleotide metabolism) we uncovered trends in connectivity within and between classifications.

Recently, we exploited the same mathematical concepts used to exhaustively explore multi-gene lethality, to classify all fluxes in the metabolic model depending upon whether or not they must increase, decrease, or become equal to zero to meet a pre-

specified overproduction target (Ranganathan et al., 2009). This classification is not limited to individual reactions but is extended for pairs, triplets, etc. while making use of all flux data available for the strain before engineering though metabolic flux analysis (MFA). The final output of this analysis is a map that spans all possible sets of engineering interventions capable of meeting a pre-specified overproduction target. The developed methodology was tested by exhaustively identifying all engineering interventions for a variety of bioproducts ranging from hub metabolites (e.g., succinate) to biofuels (e.g., isobutanol) and many others. The method recapitulated known engineering targets but also revealed new non-intuitive ones that boost productivity by performing coordinated changes on distal pathways

References

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., Mori, H., 2006. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol. 2**,** 2006 0008.

Duarte, N. C., Herrgard, M. J., Palsson, B. O., 2004. Reconstruction and validation of Saccharomyces cerevisiae iND750, a fully compartmentalized genome-scale metabolic model. Genome Res. 14**,** 1298-309.

Feist, A. M., Henry, C. S., Reed, J. L., Krummenacker, M., Joyce, A. R., Karp, P. D., Broadbelt, L. J., Hatzimanikatis, V., Palsson, B. O., 2007. A genome-scale metabolic reconstruction for Escherichia coli K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. Mol Syst Biol. 3**,** 121.

Kumar, V. S., Maranas, C. D., 2009. GrowMatch: an automated method for reconciling in silico/in vivo growth predictions. PLoS Comput Biol. 5**,** e1000308.

Ranganathan, S., Suthers, P. F., Maranas, C. D., 2009. OptFlux: An Optimization Procedure to Exhaustively Identify all the Engineering Interventions Leading to Targeted Overproductions. in preparation.

Reed, J. L., Vo, T. D., Schilling, C. H., Palsson, B. O., 2003. An expanded genome-scale model of Escherichia coli K-12 (iJR904 GSM/GPR). Genome Biol. 4**,** R54.

Satish Kumar, V., Dasika, M. S., Maranas, C. D., 2007. Optimization based automated curation of metabolic reconstructions. BMC Bioinformatics. 8**,** 212.

Suthers, P. F., Dasika, M. S., Kumar, V. S., Denisov, G., Glass, J. I., Maranas, C. D., 2009. A genome-scale metabolic reconstruction of Mycoplasma genitalium, iPS189. PLoS Comput Biol. 5**,** e1000285.

Zomorrodi, A., Suthers, P. F., Maranas, C. D., 2009. Genome-scale Gene/Reaction Essentiality and Synthetic Lethality Analysis. submitted.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Integration of Production Planning and Scheduling

Christos T. Maravelias

Univeristy of Wisconsin, Madison, WI 53706, USA

Abstract

We start our presentation with an overview of supply chain management, the associated planning problems, and a motivation for the integration of medium-term production planning and short-term scheduling decisions (Maravelias and Sung, 2008). We briefly review existing modeling and solution methods and then present four new approaches to this integrated problem.

First, we discuss how lot-sizing problems can be extended to address features often found in process networks (Sung and Maravelias, 2008). Specifically, we discuss the problem where a) multiple items can be produced in each planning period, b) sequence-independent setups can *carry over* from previous periods, c) setups can *cross over* planning period boundaries, and d) setups can be longer than one period. We develop a mixed-integer programming (MIP) formulation for this generalized problem and derive a number of optimality properties that can be used to tighten the formulation. We also discuss a network representation and the resulting MIP formulation. We present examples that illustrate how the proposed model can be used to address medium –size single- and multi-unit problems.

Second, we develop polyhedral results for a discrete-time state-task network (STN) formulation for the production planning of multi-stage continuous processes (Maravelias and Papalamprou, 2009). Specifically, we express the feasible region of the LP-relaxation as the intersection of two sets. We show that the constraints describing the first set yield the convex hull of its integer points. For the second set, we show that for integral data the constraint matrix is κ-regular, and that the corresponding polyhedron is integral if the length of the planning period is selected appropriately. We use this result to prove that for rational data, integer variables assume integral values at the vertices of the polyhedron. We also discuss special cases of this problem, for which we derive stronger results. We illustrate how these results can be used to effectively address large-scale production planning problems.

Third, we present a projection-based method where a detailed MIP scheduling model is used off-line to obtain a convex approximation of feasible production levels and a convex underestimation of total production cost as a function of production levels (Sung and Maravelias, 2007). These approximations involve only planning variables yet provide all the relevant scheduling information necessary to solve the integrated problem with high quality. We then develop two MIP formulations for the representation of nonconvex regions: the first expresses non-convex regions as unions of polytopes, while the second as difference of polytopes. We use these formulations to generate a non-convex projection of the scheduling feasible region onto the space of production level variables (Sung and Maravelias, 2009). We discuss how this method can be used in a rolling-horizon manner to address hard planning-scheduling problems, as well as a process design tool.

Finally, we discuss methods for the decomposition of planning-scheduling problems in multi-stage batch processes (Ferris et al., 2009). We generate subproblems based on fixing: a) batching decisions, b) batch-unit assignment variables in a bottlenecking

stage, c) the last batch in the sequence on each unit of the bottlenecking stage, and d) strong branching. We discuss how the above methods can be combined and carried out in a dynamic fashion exploiting the natural hierarchy of decisions (batching → unit-task assignment → sequencing). We also discuss the implementation of this dynamic decomposition algorithm within GAMS using the master-worker paradigm: the master processor generates and spawns all the tasks to worker processors, collects the results, and communicates information among workers; worker processors are part of a computational grid. We present computational results on a collection of six instances of varying difficulty.

Keywords: Production planning; scheduling; mixed-integer programming.

References

M. C. Ferris, C. T. Maravelias, A. Sundaramoorthy. (2009). Simultaneous Batching and Scheduling Using Dynamic Decomposition on a Grid. *INFORMS Journal on Computing*, in press.

C. T. Maravelias, C. Sung. (2008). Integration of Production Planning and Scheduling: Overview, Challenges and Opportunities. In Proceedings: *Foundations of Computer-aided Process Operations*, Cambridge, MA.

C. T. Maravelias, K. Papalamprou. (2009). Polyhedral Results for Discrete-time MIP Formulations for Scheduling and Production Planning. Submitted for publication.

C. Sung, C. T. Maravelias. (2007). An Attainable Region Approach for Effective Production Planning of Multi-product Processes. *AIChE J.*, 53 (5), 1298-1315.

C. Sung, C. T. Maravelias. (2008). A Mixed-Integer Programming Formulation for the General Capacitated Lot-sizing Problem. *Computers and Chemical Engineering*, 32(1), 244-259.

C. Sung, C. T. Maravelias. (2009). A Projection-Based Method for Production Planning of Multi-product Facilities. *AIChE J.*, in press.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Robust Integration of RTO and MPC

Darci Odloak

University of São Paulo - Department of Chemical Engineering, Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil. Brazil.
E-mail: odloak@usp.br

Abstract

Here, it is studied the stable integration of real time optimization (RTO) and model predictive control (MPC) for the case where the controller has a two layer structure. Stability is also obtained when model uncertainty is considered in both the target calculation and dynamic layers.

Keywords: Real Time Optimization, Model Predictive Control, Robust Stability.

1. Introduction

The usual approach of integrating real time optimization (RTO) and model predictive control (MPC) is through a three layer structure (Ying & Joseph, 1999) where in the upper layer we have the real time optimization, which is based on a rigorous nonlinear steady state model of the process. The RTO layer solves an optimization problem that defines the optimal operating conditions subject to equipment constraints and product specifications. As represented in Figure 1, the optimal inputs and outputs of the system are then passed to an intermediary layer where, based on the estimation of the disturbances $d(k)$, feasible targets for the inputs and outputs are re-computed through the solution of a QP based on the linear static model of the system.

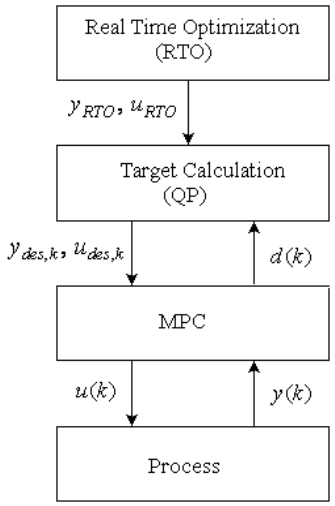


Figure 1- Control structure with RTO and the two-layer MPC.

Finally, in the MPC level, the optimizing targets are dynamically implemented. Since the optimization problem that is solved in the RTO layer is executed with a time period of hours and the target calculation and MPC problem are executed with a time period of minutes, one can consider that the RTO layer is dynamically decoupled from the other two layers. Consequently, the stability of the RTO and MPC integration will depend on the interaction between the target calculation and dynamic layers. Thus, the scope of this work is to propose a QP/MPC strategy that is stable to large excursions of the optimal operating conditions of the process system. Along the trajectory followed by the process operating point, one may expect a significant change in the parameters of the linear model on which the target calculation and dynamic layers are based. Here, it is proposed stable QP/MPC strategies for the case where the model of the process system is well known and for the case where we have a certain class of uncertainty in the process model.

2. The stable QP/MPC for the nominal system

Suppose that we have a stable system with *nu* inputs and *ny* outputs and consider the structure represented in Figure 1, where the solution of the RTO problem produces optimal values for the system outputs $y_{RTO,i=1,\ldots,ny_t}$ and for the system inputs $u_{RTO,i=1,\ldots,nu_t}$. In this system ny_t is the number of outputs that are selected to have optimizing targets and nu_t is the number of inputs that have optimizing targets. It is also assumed that $ny_t + nu_t = nu$, which is the number of available degrees of freedom of the system. Then, at time step *k*, assuming that the linear model considered in the QP and MPC layers is perfectly known, the target calculation layer of the structure represented in Figure 1 solves the following problem:

$$\min_{y_{des}(k),u_{des}(k),\varepsilon(k)} V_{k,a} = \left(y_{des}(k) - y_{RTO}\right)^T C_y \left(y_{des}(k) - y_{RTO}\right) + \left(u_{des}(k) - u_{RTO}\right)^T C_u \left(u_{des}(k) - u_{RTO}\right) + \varepsilon^T(k) C_\varepsilon \varepsilon(k) \tag{1}$$

subject to

$$y_{des}(k) - x^s(k) = D^0\left(u_{des}(k) - u(k-1)\right) \tag{2}$$

$$y_{\min} - \varepsilon(k) \le y_{des}(k) \le y_{\max} + \varepsilon(k), \quad \varepsilon(k) \ge 0 \tag{3}$$

$$-m\,\Delta u_{\max} \le u_{des}(k) - u(k-1) \le m\,\Delta u_{\max} \tag{4}$$

$$u_{\min} \le u_{des}(k) \le u_{\max} \tag{5}$$

where

C_y, C_u and C_ε are positive definite weighting matrices; $x^s(k)$ represents the predicted output of the process system at steady state as defined in the dynamic layer of the control structure; $u(k-1)$ is the control action applied to the real system at the previous time step and *m* is the control horizon of the controller corresponding to the dynamic layer of the control structure represented in Figure 1.

The optimal solution to the problem defined in (1) is then passed to the dynamic layer where a constrained infinite horizon MPC is implemented. This controller is based on the works of Odloak (2004) and González and Odloak (2009), but with some modifications to accommodate the integration of the controller with RTO and the target

calculation stage. This controller, which has the same sampling period as the QP stage, solves the following optimization problem:

$$\min_{\Delta u_k, y_{sp}(k), \delta_y(k)} V_{k,b} = \sum_{j=0}^{\infty} \left(y(k+j|k) - y_{sp,k} - \delta_{y,k} \right)^T Q_y \left(y(k+j|k) - y_{sp,k} - \delta_{y,k} \right)$$
$$+ \sum_{j=0}^{\infty} \left(u(k+j|k) - u_{des,k} \right)^T Q_u \left(u(k+j|k) - u_{des,k} \right) + \sum_{j=0}^{m-1} \Delta u(k+j|k)^T R \Delta u(k+j|k) + \delta_y^T(k) S_y \delta_y(k)$$

(6)

subject to

$$-\Delta u_{\max} \le \Delta u(k+j|k) \le \Delta u_{\max}; \qquad u_{\min} \le u(k+j|k) \le u_{\max}, \quad j = 0,...,m-1$$

$$x^s(k) + \tilde{D}^0 \Delta u_k - y_{sp}(k) - \delta_y(k) = 0$$

$$\begin{cases} y_{sp}(k) = y_{des}(k) \\ \delta_y(k) = \varepsilon(k) \end{cases} \text{ (only for those outputs that have optimizing targets)}$$

$$y_{\min} \le y_{sp}(k) \le y_{\max}$$

$$u(k-1) + \tilde{I} \Delta u_k - u_{des}(k) = 0 \text{ (only for those inputs that have optimizing targets)} \qquad (7)$$

where

Q_y, Q_u, R, S_y and S_u are weighting matrices of appropriate dimensions

$\tilde{D}^0 = \left[\underbrace{D^0 \quad \cdots \quad D^0}_{m} \right]$, $\tilde{I} = \left[\underbrace{I_{nu} \quad \cdots \quad I_{nu}}_{m} \right]$, I_{nu} is the identity matrix with dimension *nu*, D^0 corresponds to the process gain and

$$\Delta u_k = \left[\Delta u(k)^T \quad \Delta u(k+1)^T \quad \cdots \quad \Delta u(k+m-1)^T \right]^T$$

Following the same steps as in Odloak (2004), we can prove that the controller produced by the solution of problem (6) at successive time steps is asymptotically stable in the sense that if $y_{des}(k)$ and $u_{des}(k)$ remain fixed, then the outputs and inputs of the system that have targets will converge to the targets and the remaining outputs and inputs of the system will converge to steady state values inside their respective bounds.

Remark 1

The problem defined in (6) is based on a dynamic model of the process system. In the nominal case, it is assumed that a single model is able to represent adequately the system near the operating point being considered. For a system with *nu* inputs and *ny* outputs, a state space model that is suitable to the implementation of an offset free MPC can be represented in the following form (Odloak, 2004):

$$\begin{bmatrix} x^s(k+1) \\ x^d(k+1) \end{bmatrix} = \begin{bmatrix} I_{ny} & 0 \\ 0 & F \end{bmatrix} \begin{bmatrix} x^s(k) \\ x^d(k) \end{bmatrix} + \begin{bmatrix} D^0 \\ D^d \end{bmatrix} \Delta u(k) \qquad (8)$$

$$y(k) = \begin{bmatrix} I_{ny} & \Psi \end{bmatrix} \begin{bmatrix} x^s(k) \\ x^d(k) \end{bmatrix}$$

where

$x^s = \begin{bmatrix} x_1 & \cdots & x_{ny} \end{bmatrix}^T$, $x^s \in \Re^{ny}$, $x^d = \begin{bmatrix} x_{ny+1} & x_{ny+2} & \cdots & x_{ny+nd} \end{bmatrix}^T$, $x^d \in \Re^{nd}$,

$F \in \Re^{nd \times nd}$, $\Delta u(k) = u(k) - u(k-1)$

In the state space model defined in (8), the state component x^s corresponds to the integrating poles produced by the incremental form of the model, and x^d corresponds to the system modes. For stable systems, it is easy to show that when the system approaches steady state, component x^d tends to zero. F is a diagonal matrix with components corresponding to the poles of the system. The system has *nd* stable poles. Matrix D^0 is the static gain of the system.
If the model defined in (8) is used to represent the system, there is no need to include the estimation of the unmeasured disturbances (represented as *d* in Fig.1) in order to eliminate offset in the system outputs.

Remark 2
If the nominal undisturbed system starts from a steady state and, at a given time *k*, the RTO layer produces a step change in y_{RTO} and u_{RTO}, and the new steady-state defined by (y_{RTO}, u_{RTO}) is reachable, then the sequential solution of the problems defined in (1) and (6) results in the convergence of $y_{des}(k)$ to y_{RTO} and $u_{des}(k)$ to u_{RTO}.
To prove this property, suppose that the problem defined in (1) is solved at time *k* and the optimal solution corresponds to $\left(y^*_{des}(k), u^*_{des}(k), \varepsilon^*(k)\right)$ that is passed to the lower level of the control structure where the problem defined in (6) should be solved at the same time step considering the new targets. Also, let the corresponding optimal value of the objective function of problem (1) be designated $V^*_{k,a}$. Now, observe that we can find a feasible solution to problem (6) by adopting $\Delta u_k = 0$, $y_{sp}(k) = y^*_{des}(k)$ (for those outputs that have targets), $y_{sp}(k) = 0$ and $\delta_y(k) = x^s(k) - y_{sp}(k)$ (for the other outputs). Note also that constraint (7) is certainly feasible because of (4). So, this problem is always feasible and from its solution we obtain the vector of optimal control sequence Δu^*_k. Then, the first component of Δu^*_k is injected in the true process and we wait until time step *k*+1 where problem (1) is solved again. Now, we can show that a feasible solution to problem (1) at time *k*+1 is given by $\left(\tilde{y}_{des,k+1}, \tilde{u}_{des,k+1}, \tilde{\varepsilon}_{k+1}\right) = \left(y^*_{des,k}, u^*_{des,k}, \varepsilon^*_k\right)$. It is easy to see that this solution satisfies constraints (3-5) and we need only to show that it also satisfies (2). For this purpose let us write the left hand side of (2) for the proposed solution as follows

$$LHS = \tilde{y}_{des,k+1} - x^s(k+1) = y^*_{des,k} - x^s(k) - D^0 \Delta u^*(k)$$

Now using (2) at time *k*, we have

$$LHS = D^0 \left(u_{des,k} - u(k-1)\right) - D^0 \Delta u(k)$$

Analogously, the right hand side of (2) for the proposed solution at time k+1 can be written as follows

$$RHS = D^0\left[\tilde{u}_{des,k+1} - u(k)\right] = D^0\left[u_{des,k} - u(k-1) - \Delta u(k)\right]$$

and it is clear that *LHS*=*RHS* and consequently (2) is satisfied by the proposed solution.

If the value of the objective function defined in (1) corresponding to the feasible solution $\left(\tilde{y}_{des,k+1}, \tilde{u}_{des,k+1}, \tilde{\varepsilon}_{k+1}\right)$ is designated $\tilde{V}_{k+1,a}$, then it is clear that $\tilde{V}_{k+1,a} = V^*_{k,a}$ and consequently $V^*_{k+1,a} \le V^*_{k,a}$, which means that the objective function of problem (1) is converging. Then, if the optimal steady state defined by $\left(y_{RTO}, u_{RTO}\right)$ is feasible, $V_{k,a}$ will converge to zero and the system inputs and outputs will converge to the their targets.

3. The robust QP/MPC for the uncertain system

Depending on the market conditions, product specifications, feedstock quality, etc, the RTO layer may define optimizing targets that correspond to quite different operating conditions, which means that a MPC controller, based on the nominal linear model, can result in poor performance or instability of the closed loop system. So, model uncertainty should be considered in the target calculation and dynamic layers of the MPC controller. For instance, suppose that the true gain matrix D^0_t of the process model, lies in a convex polytope defined by L vertices $D^0_1, D^0_2, ..., D^0_L$, which means that

$$D^0_t = \sum_{i=1}^{L} \lambda_i D^0_i, \;\; \sum_{i=1}^{L} \lambda_i = 1, \;\; \lambda_i \ge 0 \tag{9}$$

Now, we may define the objective function of the target calculation layer for each of the vertices of the polytope that characterizes the uncertain system defined in (9) as follows

$$V_{k,a}(D^0_i) = \left(y_{des,i}(k) - y_{RTO}\right)^T C_y \left(y_{des,i}(k) - y_{RTO}\right) + \left(u_{des}(k) - u_{RTO}\right)^T C_u \left(u_{des}(k) - u_{RTO}\right) + \varepsilon_i(k)^T C_\varepsilon \varepsilon_i(k) \qquad i = 1, ..., L \tag{10}$$

Then, the optimization problem to be solved in the QP layer to calculate the targets for the MPC controller is defined, for the uncertain system, as follows

$$\min_{\substack{y_{des,i}(k), u_{des}(k), \varepsilon_i(k) \\ i=1,...,L}} V_{k,a}(D^0_n) \tag{11}$$

subject to

(4), (5), and

$$V_{k,a}(D^0_i) \le \tilde{V}_{k,a}(D^0_i), \quad i = 1, ..., L \tag{12}$$

$$y_{\min} - \varepsilon_i(k) \le y_{des,i}(k) \le y_{\max} + \varepsilon_i(k), \quad \varepsilon_i(k) \ge 0, \quad i = 1, ..., L \tag{13}$$

$$y_{des,i}(k) - x^s(k) = D^0_i\left[u_{des}(k) - u(k-1)\right], \quad i = 1, ..., L \tag{14}$$

where

- D_n^0 is the gain matrix of the nominal or most probable model of the process system.
- $\tilde{V}_{k,a}(D_i^0)$ is computed with $\left(\tilde{y}_{des,i}(k), u_{des}^*(k-1), \varepsilon_i^*(k-1)\right)$ where $\tilde{y}_{des,i}(k-1)$ satisfies the equation: $\tilde{y}_{des,i}(k) - x^s(k) + D_i^0\,\Delta u(k-1) = D_i^0\left(u_{des}^*(k-1) - u(k-2)\right)$ (15)

Remark 3

Problem (11) is always feasible, because $\left(\tilde{y}_{des,i}(k), u_{des}^*(k-1), \varepsilon_i^*(k-1),\ i=1,...,L\right)$ is a feasible solution to the problem (11). Also, if at time *k*, $\left(y_{des,i}^*(k), u_{des}^*(k), \varepsilon_i^*(k),\ i=1,...,L\right)$ is the optimal solution to the problem defined in (11), then, the solution that corresponds to the true process model whose gain matrix is $D_t^0 = \sum_{i=1}^{L} \lambda_i D_i^0$ is given by

$$\left(y_{des,t}^*(k),\, u_{des}^*(k), \varepsilon_t^*(k)\right) = \left(\sum_{i=1}^{L} \lambda_i y_{des,i}^*(k),\, u_{des}^*(k),\, \sum_{i=1}^{L} \lambda_i \varepsilon_i^*(k)\right)$$

It is easy to show that this solution satisfies contraints (4-5) and (13-14). To show that it also satisfies constraint (12), we write this inequality explicitly for the optimal solution and for each vertex of the polytope that charcterizes the model uncertainty:

$$\begin{aligned}
&\left(y_{des,i}^*(k) - y_{RTO}\right)^T C_y \left(y_{des,i}^*(k) - y_{RTO}\right) + \left(u_{des}^*(k) - u_{RTO}\right)^T C_u \left(u_{des}^*(k) - u_{RTO}\right) + \varepsilon_i^*(k)^T C_\varepsilon \varepsilon_i^*(k) \\
&\le \left(\tilde{y}_{des,i}(k-1) - y_{RTO}\right)^T C_y \left(\tilde{y}_{des,i}(k-1) - y_{RTO}\right) + \left(u_{des}^*(k-1) - u_{RTO}\right)^T C_u \left(u_{des}^*(k-1) - u_{RTO}\right) \\
&\qquad + \varepsilon_i^*(k-1)^T C_\varepsilon \varepsilon_i^*(k-1), \qquad i = 1,...,L
\end{aligned} \tag{16}$$

Using the Schur complement, it is possible to show that (16) is equivalent to

$$\begin{bmatrix}
C_y^{-1} & 0 & 0 & 0 & \left(y_{des,i}^*(k) - y_{RTO}\right) \\
0 & C_\varepsilon^{-1} & 0 & 0 & \varepsilon_i^*(k)^T \\
0 & 0 & -C_y^{-1} & 0 & \left(\tilde{y}_{des,i}(k) - y_{RTO}\right) \\
0 & 0 & 0 & -C_\varepsilon^{-1} & \varepsilon_i^*(k-1) \\
\left(y_{des,i}^*(k) - y_{RTO}\right)^T & \varepsilon_i^*(k)^T & \left(\tilde{y}_{des,i}(k) - y_{RTO}\right)^T & \varepsilon_i^*(k-1)^T & \begin{array}{l} -\left(u_{des}^*(k) - u_{RTO}\right)^T C_u \left(u_{des}^*(k) - u_{RTO}\right) \\ +\left(u_{des}^*(k-1) - u_{RTO}\right)^T C_u \left(u_{des}^*(k-1) - u_{RTO}\right) \end{array}
\end{bmatrix} \ge 0$$

$$i = 1,...,L \tag{17}$$

Now, since (17) is a linear matrix inequality, any linear combination of the corresponding process models will produce a similar combination of the inequalities represented in (17), and consequently (12) is satisfied for the true model. So, $V_k(D_t^0)$ is decreasing and will converge to zero if the steady state corresponding to $\left(y_{RTO}, u_{RTO}\right)$ is reachable by the true plant.

Before defining the optimization problem that is solved in the dynamic layer of the control structure defined in Figure 1, for the case where there is model uncertainty, let

us characterize the type of uncertainty of the dynamic model considered in the MPC. Let any dynamic model be represented by the following set of parameters: $\theta = \left[D^0, D^d, F \right]$. Here, we assume that the true model of the process system is represented as $\theta_t = \left[D_t^0, D_t^d, F_t \right]$, where $\left(D_t^0, D_t^d \right) = \sum_{i=1}^{L} \lambda_i (D_i^0, D_i^d)$, $\sum_{i=1}^{L} \lambda_i = 1$, $\lambda_i \geq 0$ and $F_t = F_{j=1,\ldots,L}$. We should observe that the uncertainty is not polytopic in terms of the parameters related to the dynamics of the system. This means that the gain of the true system is a convex combination of a finite of possible gains, while the dynamics of the system is defined by one element of a finite set of possible dynamics.

Based on the works of Odloak (2004) and González et al. (2007) that propose robust MPC controllers for the type of uncertainty described above, the problem that is solved in the dynamic layer of the control structure of the uncertain system is the following

$$\min_{\Delta u_k, y_{sp,i}(k), \delta_{y,i}(k)} V_{b,k}(\theta_n) =$$

$$\sum_{j=0}^{\infty} \left[y_n(k+j|k) - y_{sp,n}(k) - \delta_{y,n}(k) \right]^T Q_y \left[y_n(k+j|k) - y_{sp,n}(k) - \delta_{y,n}(k) \right]$$

$$+ \sum_{j=0}^{\infty} \left[u(k+j|k) - u_{des}(k) \right]^T Q_u \left[u(k+j|k) - u_{des}(k) \right] + \sum_{j=0}^{m-1} \Delta u(k+j|k)^T R \, \Delta u(k+j|k)$$

$$+ \delta_{y,1}^T(k) S_y \delta_{y,1}(k) + \cdots + \delta_{y,L}^T(k) S_y \delta_{y,L}(k) \tag{18}$$

subject to

$-\Delta u_{\max} \leq \Delta u(k+j|k) \leq \Delta u_{\max}, \quad j = 0, \ldots, m-1;$

$u_{\min} \leq u(k+j|k) \leq u_{\max}, \quad j = 0, \ldots, m-1$

$y_{sp,i}(k) = y_{des,i}; \quad i = 1, \ldots, L$ (for those outputs that have optimizing targets)

$y_{\min} \leq y_{sp,i}(k) \leq y_{\max}, \quad i = 1, \ldots, L$

$x^s(k) + \tilde{D}_i^0 \Delta u_k - y_{sp,i}(k) - \delta_{y,i}(k) = 0; \; i = 1, \ldots, L$

$u(k-1) + \tilde{I} \Delta u_k - u_{des}(k) = 0$ (only for those inputs that have optimizing targets)

$V\left(\Delta u_k, \delta_{y,i}(k), y_{sp,i}(k) \right) \leq V\left(\Delta \tilde{u}_k, \tilde{\delta}_{y,i}(k), \tilde{y}_{sp,i}(k) \right), \; i = 1, \ldots, L$

where

$$V\left(\Delta u_k, \delta_{y,i}(k), y_{sp,i}(k) \right) = \sum_{j=0}^{\infty} \left[y_i(k+j|k) - y_{sp,i}(k) - \delta_{y,i}(k) \right]^T Q_y$$

$$\times \left[y_i(k+j|k) - y_{sp,i}(k) - \delta_{y,i}(k) \right] + \sum_{j=0}^{\infty} \left[u(k+j|k) - u_{des}(k) \right]^T Q_u \left[u(k+j|k) - u_{des}(k) \right]$$

$$+ \sum_{j=0}^{m-1} \Delta u(k+j|k)^T R \, \Delta u(k+j|k) + \delta_{y,i}^T(k) S_y \delta_{y,i}(k)$$

$\Delta \tilde{u}_k = \left[\Delta u^*(k|k-1)^T \quad \cdots \quad \Delta u^*(k+m-2|k-1)^T \quad 0 \right]^T$; $\tilde{y}_{sp,i}(k) = y_{sp,i}^*(k-1)$

$\tilde{\delta}_{y,i}(k)$ is such that $x^s(k) + \tilde{D}_i^0 \, \Delta \tilde{u}_k - y_{sp,i}(k) - \tilde{\delta}_{y,i}(k) = 0, \; i = 1, \ldots, L$

Note that the problem defined in (18) is always feasible because $\left(\Delta\tilde{u}_k, \tilde{y}_{sp,i}(k), \tilde{\delta}_{y,i}(k), i=1,...,L\right)$ is a feasible solution to this problem.

Assuming that the problem defined in (11) has converged, that is $y_{des,i}(k) = y_{RTO,i}$ and $u_{des}(k) = u_{RTO}$ for a finite *k*, then, it can be shown that $\left(\Delta\tilde{u}_k, \sum_{i=1}^{L}\lambda_i \tilde{y}_{sp,i}(k), \sum_{i=1}^{L}\lambda_i \tilde{\delta}_{y,i}(k)\right)$ is a feasible solution to problem (18) for the true system and, corresponding to this solution, we have $V_{b,k+1}\left(\Delta\tilde{u}_{k+1}, \tilde{\delta}_{y,t}(k+1), \tilde{y}_{sp,t}(k+1)\right) = V_{b,k}\left(\Delta u_k^*, \delta_{y,t}^*(k), y_{sp,t}^*(k)\right)$ and consequently $V_{b,k}(\theta_t)$ is decreasing and the closed loop system is stable.

4. Conclusion

Here, it was proposed a two step algorithm to the robust stable integration of real time optimization and multivariable control of a process system in the control structure of the process plant. The first step consists of a target calculation layer based on an optimization problem where convergence of the computed targets to the set points defined by the RTO layer is assured. The second step of the algorithm is based on an extended infinite horizon MPC where the errors on the inputs that have targets are penalized and the outputs are controlled at the desired targets or inside zones. An extension of the method is presented to the case where we have polytopic uncertainty in the model parameters associated to the gains of the process system. Uncertainty in the dynamic parameters can be included by considering a finite set of possible dynamics.

References

Ying C.-M. & Joseph B. (1999). Performance and stability analysis of LP-MPC and QP-MPC cascade control systems, *AIChE J.* 45 (7), pp. 1521–1534.

Odloak D. (2004)**.** Extended robust model predictive control. *AIChE Journal* 50 (8), pp. 1824-1836.

González A. H., Marchetti J. L. & Odloak D. (2007)**.** Extended robust model predictive control of integrating systems. *AIChE Journal* 53 (7), pp. 1758-1769.

González A. H. & Odloak D. (2009**).** A stable MPC with zone control. *Journal of Process Control* 19 (1), pp. 110-122.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Sustainable Supply Chains: Key Challenges

Ana Paula Barbosa-Póvoa

Centro de Estudos de Gestão do-IST, CEG-IST, UTL, Av. Rovisco Pais 1049–001 Lisboa, Portugal, apovoa@ist.utl.ptl

Abstract

In the new economic context the long term success of any organization is built not only on profits and profitability but also on its contribution to the future of people and the future of the planet. Being supply chains a key cornerstone in any organization the consideration of sustainability at the supply chain level is recognized as an emerging area that needs to be studied in a systematic way. The concept of sustainable supply chain is raising both in the industry and academic communities and studies have been reported on this new supply chain management concept. An overview of some of these works is performed along this paper and some major issues that should form a future research agenda on the area are identified.

Keywords: Sustainability, Supply Chains, Energy Efficiency, Life Cycle Analysis.

1. Introduction

Supply Chain Management as a concept was introduced back in the early 1980s (Oliver and Webber, 1982) but received most attention from 1990 onwards. Nowadays, supply chains are the foundation of any organization that wants to act in the current global competitive economy. Conventionally, supply chains have been seen as operational structures responsible for the production and distribution of products to different markets disperse around the globe, whose main objective was the satisfaction of costumers at a minimum cost. Supply chain research communities tended to focus on sub-sets of the supply chain management decisions instead of looking in a cross-functional and co-coordinated way to the different existent levels of decisions (Varma et al., 2007). This paradigm is however changing and advances in global competition with a parallel exhaustion of natural resources and an increased society awareness towards environment, created a new way of thinking when managing supply chains.

The need of seriously exploring the concept of sustainable supply chains within a collaborative perspective is seen as a goal to improve companies' revenue growth and costumer's recognition (Kleindorfer, et al., 2005). As stated by the New Zealand World Business Council for Sustainable Development, supply chains need to be seen as operational structures that manage the raw materials and services from suppliers to manufactures/service providers and back with improvement of the social and environmental impacts explicitly considered (NZBCSD, 2003). To respond to this challenge enterprises´ must invest on the design, planning and operation of their logistic structures while considering the minimization of their global energy consumption (Grossmann, 2004). As reported by Côte et al. (2008) supply chain management can significantly influence environmental performance, both positively and negatively. Substantial opportunities thus exist, along the supply chain, to reduce company's energy consumption. The extension of the traditional supply chain to consider the reserve flows where waste management is incorporated and a better rationalization of resources is

achieved, reduction of greenhouse gases emissions through the use of less polluted production and transportation modes; substitution of chemicals that might reduce the generation and management of hazardous waste or the reduction of the packaging waste are only some of the options that companies should consider.
Along this paper sustainable supply chains are being addressed. Some major works considering this new supply chain paradigm are analyzed and some of the key challenges that should to be considered in a future research agenda on sustainable supply chains are identified.

2. Sustainability and Supply Chains

Sustainable Supply Chains Management has its roots in both Sustainability and Supply Chain Management literature and involves a broadened approach of the supply chain management. Sustainability as defined by the World Commission on Environment and Development (WCED, 1987) is the use of resources to meet needs of the present without comprising the ability of future generations to meet their own needs. On the other hand, Supply Chain Management can be viewed as a systems approach to manage the entire flow of information, materials, and services from raw materials suppliers through factories and warehouses to customers and back. A focus on supply chains is therefore a step towards the adoption of a growth on sustainability. As stated by Linton et al. (2007) sustainability must integrate issues and flows that go beyond the core of the supply chain management such as: product design, production, distribution, end of life products and recovery processes.

Product design: should look into methodologies to assess environmental implications at the product design level. These can assist in the determination on how to design a product so as to minimize its environmental impact over its usable life afterwards. The integration between the product design and the processing phase should be explored so as to diminish environmental impacts and guarantee the long-term sustainability.

Production: research at the process level should concentrate efforts in the development of cleaner process aiming at reducing by-products that might damage the environment and human health. Process technologies should be designed and operated to easily incorporated reusable products aligning in this way with the need of decreasing the use of resources.

Distribution: transportation appears as one of the main contributors to the supply chain footprint. The reduction of greenhouse gases through the use of cleaner transportation modes coupled with an optimal definition of the supply chains structures will contribute positively towards sustainability.

End-of-life products and Recovery processes: how to effectively recycle end-of-life products by using a reserve chain is an area that has become a field of rapidly growing importance. The initial product design has a great influence on to each degree a product can be re-used, remanufactured, recycled, incinerated or disposed of. Work on how to capture the value remaining in products at the end of its life should be extended.

The in-breath in each one of these areas is out of the scope of the supply chain management but their study is definitely related to the supply chains systems. Efforts on both sides will support the goal of responsible care as stated by Grossmann (2004). Process Systems Engineering play an important role in any of the above areas and its contribution towards a global life cycle assessment will contribute to the achievement of

such goal. The development of systematic methods and tools that guarantee the design of environmental benign products and processes coupled with sustainable supply chains is therefore the goal to accomplish.

3. Sustainable Supply Chain Management

Adding the Sustainability component to supply chain management implies considering the influence and relationships between supply chain management and the environment as well as the social aspects of the business practices. As a major challenge the research on the area needs to develop adequate measures for the environment supply chains impacts and should taken as basis the engagement of all the supply chain activities in the triple bottom line (3BL), the threes P´s, people, profit and planet (Kleindorfer, et al., 2005). Such concern has been identified by different authors during the present decade and different approaches have been followed. Some looked only to the forward supply chains while others explored the reverse logistics as a way to deal with the environmental and social pressures. The closed-loop supply chain concept emerged where both flows, forward and reserve, are considered simultaneously. As referred by Guide and Van Wassenhove (2002) the companies that have been most successful with their reverse supply chains are those that closely coordinate them with the forward supply chains, creating the closed-loop supply chain. These aim to contribute towards the sustainable development by defining and operating the right systems that guarantee a reduction of resources by recycling and recovery end of life products back into the chain. As stated by Salema et al. (2007) the literature on closed-loop supply chains, has been slowly building, but remains sparse.

Jayaraman et al. (1999) developed a mixed integer programming formulation to model a closed-loop supply chain. The model was tested on a set of problems based on the parameters of an existing electronic equipments remanufacture firm. Fleischmann et al. (2001) proposed a model for the location of logistic facilities. Their work was the first to propose a general model formulation (Mixed Integer Linear Programming Model, MILP) that simultaneously optimizes the reverse and the forward networks. Two case studies were used to explore the model application. Krikke et al. (2003) also proposed a MILP model for the design of a closed-loop supply chain where both location-allocation decisions and product design are considered. The objective function includes both supply chain costs and environmental impacts using a performance indicator (based on LCA approach). As product design is involved, an assembly and disassembly of products is explicitly modelled. Beamon and Fernandes (2004), presented a model for a single product closed-loop supply chain design. Fandel and Stammen (2004) propose a strategic model for the supply chain design. Salema et al. (2006, 2007, 2008) study the problem of designing simultaneously the forward and reverse networks. In these works the models evolved along time and different aspects of the closed-loop supply chains have been considered. The proposed models are fairly general as they incorporate facility capacity limits, multi-product flows and uncertainty under a dynamic context. Strategic versus tactical decisions were considered. The final networks structures as well as the planning of materials along the chain are obtained. Quarasigui Frota Neto et al. (2008), developed a framework for the design of logistics structures in which profitability and environmental impacts are balanced. A multi-objective model was developed. An European pulp and paper industry case-study was solved.

All the above works focused essentially on the strategic and at most on the planning contexts. At the operational level fewer works appeared that dealt with the closed-loop supply chains. Amaro and Barbosa-Póvoa (2008a) looked into the detailed scheduling

of supply chains with reverse flows where different product recovery policies are analysed. A real case-study of the pharmaceutical industry has been solved. Later on, the same authors (Amaro and Barbosa-Póvoa, 2008b) presented a generic approach where the integration between the planning and the scheduling level was proposed. The model also contemplated the reserve flows where management of end-of-life products was considered.

On the reserve logistic structures several works have also appeared. A review, on the characteristics of the research on reserve logistics, is presented by Rubio et al (2008). Ammons et al. (2000) developed a MILP model for the design and planning of reverse production systems. The State-Task Network (STN) was used as representation methodology. The model was applied to the recycling of a network router from an Original Equipment Manufacturer. The same problem was addressed by Realff et al. (2000) through a different approach where a robust optimization framework was used. Duque et al. (2007), using the maximal State-Task Network proposed a optimization model for the design and operation of a recovery route for residual industrial products with uncertain demands. The final model is able to suggest the optimal processing and transport routes, while optimizing a given objective function and meeting design and environmental constraints. This work was extended in Duque et al. (2009) where the eco-indicator 99 was used to quantify the environmental impact of the chain. Quarasiguasi Frita Neto et al. (2009) studied the eco-efficiency methodology and proposed a multi-objective linear problem with three objectives: minimize costs, cumulative energy demand and waste in a reserve logistics network.

On the forward chains the sustainability concept has been mainly associated with the quantification of the environmental impacts. As referred by Papageorgiou (2008), on its review on supply chain optimization for the process industries, Zhou et al. (2000) developed a goal programming model to account for sustainability aspects on the supply chains of continuous processes. Hagelaar et al. (2002) explore the concept of Life Cycle Analysis (LCA) when applied to a supply chain context and concluded that no guidelines exist for an integration of these two strategies. Turkay at al. (2004) explored the problem of multi-company collaborative supply chain management where not only economical goals where considered but also environmental aspects where incorporated. Later on, Hugo and Pistikopoulos (2005) looked into the supply chains planning problem where environmental concerns were accounted for. A mathematical formulation was developed which was applied to a bulk chemicals supply chain. In 2006, Soylu et al. (2006) analysed the synergies that may exist on a collaborative supply chain of energy systems. Matos and Hall (2007) analysed the integration of sustainable development concepts into the supply chains and build up a framework that should guide the utilization of LCA methods into the supply chains. Two case studies were studied, oil sand refining and agriculture biotechnology.

4. Conclusions and Key Challenges

As presented in the reviews published by Kleindorfer, et al. (2005), Srivastava (2007) and Seuring and Muller (2008) and through the analysis presented above it can be said that although a growing interest is observed on the incorporation of sustainability into the supply chain systems, new advances and developments are still required. The footprint of all the supply chain activities needs to be contemplated from the development of new products till the production, storage, transportation and reuse of the current ones. Investment on sustainable technologies, operations and supply chains must exist and an integrate decision over the life cycle of products should characterize

the research on the field. To achieve such target the research community should look into the following key challenges:
1 . What are the good measures for sustainability within the sustainable supply chains? Further work needs to be done on how to measure properly the environmental impacts and its consequences in the social well being of the population.
2. How to defined or redesign the supply chains to facilitate collection, refurshiment, recycling or disposal or returned products? The optimization of the supply chains structures to easily incorporate the end-of-life products is required. Although, several studies appeared in the closed-loop supply chain, general models are still scarce and its application to a large number of supply chains has not yet been done.
3. How to deal with uncertainty? An uncertainty analysis is required. Uncertainty influences the sustainable supply chains both at the supply chain structure definition as well as at the associated planning and scheduling activities. Risk modeling aspects should also be considered. Different decisions lead to different levels of risk that need to be defined so as to take the right decision.
4. How to trade-off the different issues of the 3BL into the supply chains management, profit, people and planet? Multi-objectives approaches need to be explored, which encompass with all the above aspects.

5. Acknowledgments

The author gratefully acknowledge the financial support provided by Fundação para a Ciência e Tecnlogia (FCT), Portugal.

References

Amaro, A. C. & A.P.F.D. Barbosa-Póvoa (2008), Optimal Supply Chain Management with Detailed Scheduling, Ind. Eng. Chem. Res., 47 (1), 116 -132

Amaro, A. C. & A.P.F.D. Barbosa-Póvoa (2008), Planning and Scheduling of Industrial Supply Chains with Reverse Flows: A real pharmaceutical case-study, Computers & Chemical Engineering, 32, 6006-2635

Ammons, J. C., M. Realff, & D. Newton. (2000). Decision models for reverse production system design. Handbook on Environmentally Conscious Manufacturing. Kluwer, Bonton, 341-362

Beamon, B. M., & C. Fernandes. (2004). Supply-chain network configuration for product recovery. Production Planning & Control 15:270-281.

Côte, R.P., J. Lopez, S. Marche, G. Perron, & R. Wright (2008). Influences, practices and opportunities for environmental supply chain management in Nova Scotia SMEs, Journal of Cleaner Production, 16, 15, 1561-1570.

Fandel, G., & M. Stammen. (2004). A general model for extended strategic supply chain management with emphasis on product life cycles including development and recycling. International Journal of Production Economics 89, 293-308.

Fleischmann, M., H. R. Krikke, R. Dekker, & S. D. P. Flapper. (2000). A characterisation of logistics networks for product recovery. Omega, 28, 653-666.

Duque, J., A.P.F.D. Barbosa-Póvoa & A.Q. Novais (2007), Synthesis and Optimisation of the Recovery Route for Residual Products with Uncertain Product Demands, Computers and Operations Research, 34, 1463-1490.

Duque, J., A.P.F.D. Barbosa-Póvoa & A.Q. Novais (2009), Effect of life cycle impact assessment on the design and scheduling of a recovery network for industrial polluted waste, European Symposium of Computer Aided Process Engineering, Carcow, Poland.

Linton, J. D., R. Klassen & V. Jayaramann (2007). Sustainable supply chains: an introduction, Journal of Operations Management (JOM), 25, 1075-1082.

Kleindorfer, P.R., K. Singhal & L.N. Van Wassenhove (2005). Sustainable Operations Management, Production and Operations Management, 14, 4, 482–492.

Krikke, H., J. Bloemhof-Ruwaard &L. N. Van Wassenhove. (2003). Concurrent product and closed-loop supply chain design with an application to refrigerators. International Journal of Production Research 41, 3689-3719.

Guide, D. R. & L.N. Van Wassenhowe (2002). The Reverse Supply Chain, Harvard Business Review, 80,2,25-26.

Grossmann, I. E. (2004). Changes in the new millennium: product discovery and design, enterprise and supply chain optimization, global life cycle assessment, Computers and Chemical Engineering, 29, 29-39.

Hagelaae G. & J.G.A. van der Vorst (2002). Environmental supply chain management: using life cycle assessment to structure suppy chains, International Food and Agribusiness, 4, 399-412.

Hugo, A., & E. N. Pistikopoulos (2005). Environmentally conscious long-range planning and design of supply chain networks. Journal of Cleaner Production, 13, 1471-1491.

Jayaraman, V., V. D. R. Guide & R. Srivastava (1999). A closed-loop logistics model for remanufacturing. Journal of the Operational Research Society (JORS), 50, 497-508.

Quarisiguai Frota Neto , J.Q., J.M. Bloemhof-Ruwaard, J.A.E.E.van Nunen & E. Heck(2008). Designing and evaluating sustainable logistics networks, International Journal of Production Economics,111, 195-208.

Quarisiguai Frota Neto , J.Q., G. Walther, J. Bloemhof, J.A.E.E. van Nunen, E. Heck & T. Spengler (2009). A methodology for assesssing eco-efficiency in logistics networks, European Journal of Operational Research (EJOR), 193, 670-682.

Matos, S. & J. Hall J (2007). Integrating sustainable development in the supply chain: the case of life cycle assessment in oil and gas and agricultural biothechnology, JOM, 25,6, 1083-1102.

NZBCSD, (2003). Business Guide to a Sustainable Supply Chain, New Zealand Business Council for Sustainable Development.

Oliver, R. & M. Webber (1982), Supply Chain management: logistics catches up with the strategy in Christopher, M. (Edt.), Logistics: The Strategic Issues, Chapmann & Hall, London, 63-75.

Papageorgiou, L. (2008). Supply Chain Optimization for the process industries: Advances and Opportunities. FOCAPO 2008, USA.

Realff, M. J., J. C. Ammons, & D. Newton (2000). Strategic design of reverse production systems. Computers & Chemical Engineering 24,991-996.

Rubio, S., A.Chamorro & F. Miranda (2008), Characteristics of the research on reverse logistics, International Journal of Production Research, 46, 4, 1099-1120.

Salema, I., A. P. F. D. Barbosa-Póvoa, & A. Q. Novais (2006), A Warehouse based Design Model for Reverse Logistics, Journal of Operational Research Society,57,6, 615-629.

Salema, I., A. P. F. D. Barbosa-Póvoa, & A. Q. Novais (2007), An Optimization Model for the Design for a Capacitated Multi-product Reverse Logistics Network with Uncertainty, European Journal of Operations Research, 179,1063-1077.

Salema, I.G., A. P. F. D. Barbosa-Póvoa, & A. Q. Novais (2008), A strategic and tactical model for closed-loop supply chains, ORSpectrum, DOI 10.1007/s00291-008-0160-5.

Soylu, A., Oruc, C., M. Turkay, K. Fujita, & T. Asakura (2006). Synergy analysis of collaborative supply chain management in energy systems using multi-period MILP. EJOR, 174, 387-403.

Srivastava, S.K, (2007). Green supply-chain managemnt: a state of the art literature review. International Journal of Management Reviews,9,1, 53-80.

Turkay, M., C. Oruc, K. Fujita & T. Asakura (2004). Multi-company collaborative supply chain management with economical and environmental considerations. Computers & Chemical Engineering, 28, 985-992.

Varma, V. A., G. V. Rekalitis, , G.E. Blau & J. Pekny (2007). Enterprise-wide modelling and optimization – An overview of emerging research challenges and opportunities, Computers & Chemical Engineering 31, 692-711.

Zhou, Z., S. Cheng & B. Hua (2000). Supply chain optimization of continuous process industries with sustainability considerations. Computers & Chemical Engineering, 24,1151-1158.

WCED, World Commission on Environmental and Development, (1987). Our common future, Oxford University Press, New York.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Real Time Industrial Process Systems: Experiences from the Field

Carlos A. Ruiz

Soteica LLC, Ideas & Technology
16010 Barker's Point Lane, Suite 580, Houston, TX 77079, USA
Phone: +1 (281) 829 3322, Fax: +1 (281) 966 1710
e-mail: carlos.ruiz@soteica.com, http://www.soteica.com

Abstract

Over a period of more than 20 years of experience, deploying industrial projects related to advanced process control and real time optimization in processing industries around the world, many times the following or similar issues have been found:

- A lot of applications of Process Systems Engineering related technologies are extremely successful when used offline by Engineers, in their offices, but only a few of them succeed when transferred to the Operators to be routinely used in real time, tied to online field data, under unsupervised automatic execution.
- The use of certain technologies, deployed under well engineered, successful real time projects, received enthusiastically by Engineers and Operators are not used more than for just a few months after the project completion.
- Real time systems with pretty good up-time figures degrade and finally unused when the Engineer in charge of them is promoted to a new position as a consequence of being so successful with them.
- Very well proven, sound technologies, fail when applied at a given Site while they succeed in a different location owned by the same Company.
- After a failure with one advanced technology, the Site or the Company becomes "immunized" against similar technologies and only after several years, with the renewal of the Engineering staff and young blood that has not been exposed to the past failure, does the technology have a chance to be considered again.

This paper will address some of the root causes for those issues and discuss about the best practices on how to avoid project failures. The range of projects from where experience was acquired range from individual pieces of equipment and isolated Units to whole Areas or Site wide systems (for example, the utilities system comprising the steam, fuels, boiler feed water and electricity networks). All of them were simulated, optimized and/or controlled with real time, online, industrially well established technologies.

Main industrial project steps are explained and critical details to be taken into account when dealing with real time data and online automatic executions. The objective is to help achieve a successful technology deployment avoiding project pitfalls and ensuring the continuous use of the applications based on the proper technology transfer and maintenance.

Keywords: process systems engineering projects, real time optimization, online data

1. Plant Information Systems: An Enabling Technology

DCSs (Distributed Control Systems (DCS) and Plant information (PI) systems are the most widely available data sources in the current industrial context. PI systems usually acquire data from the DCSs and store it in a unique repository. The long term, facility wide PI system based historians constitute what is known as an *enabling technology*, because they became the cornerstone on top of which many other applications can be built. Centralized real-time databases provide access to massive actual and historical process, laboratory and financial data.

Modern PI systems filled the gap from the isolated, customized and proprietary DCS databases and proprietary network protocols to real time applications, by providing a centralized data repository. Moreover, modern open standards allow data to be accessed easily by every application that has a need for it.

Only a few years ago, the interfacing work was a big chunk of any industrial project. Many hours of special programming and integration, requiring detailed knowledge about a lot of proprietary Application Programming Interfaces (APIs) was needed. Nowadays, almost all of the DCS and PI system providers adhere to widely adopted open standards like OPC (OLE for Process Control). Originally built around the Microsoft DCOM technology, OPC standard protocols are nowadays being updated to provide easier ways to deal with data transfer and security issues. Details can be obtained from the OPC Foundation web page (OPC Foundation, 2009).

2. Real Time Process Systems

Control rooms worldwide are facing a shortage of people. Operators are concentrating more and more on the Units under his/her responsibility. DCS screens flourished and hundreds or thousands of tags are available per Unit and a multitude of process diagrams and trends can be projected to the walls and even the ceilings. But only a few employees are available to take care about the process operability and economics. Industrial Sites are becoming increasingly complex and inter-related, not only between Units but also with the power and utilities systems.

Tighter and increasingly restrictive regulations related to emissions are also imposing constraints and adding complexity to their proper management. Deregulated feedstock, electric and fuels markets with varying contracted prices (seasonal or daily) also introduce additional challenges to Operating personnel. Industrial facilities like Refineries and Petrochemicals are becoming increasingly aware that process and utilities systems need to be optimally operated and managed together.

Process systems engineering applications came to the rescue. They have always attempted to produce tools that improve the way industrial production systems were designed, controlled and managed. In the early days, offline applications where the only possibility but the idea was always around about using them in real time.

As explained above, the evolution from having plant information scattered throughout many islands of automation to a unified and centralized Plant Information System was a clear enabling layer for such work.

However, to successfully implement a real time process engineering system there are still several issues and questions that need to be addressed and recommended implementation practices useful to be taken into account.

2.1. Proper Technology Selection

Select the technology that fits with the problem solution rather than force the problem to be solved with an a priori selected technology. For example, if a simple linear regression with two input variables is sufficient enough to infer the needed variable, do

not force the use of a several inputs neural network. But always high rank systems based on sound process engineering principles (i.e., mass and energy balances and thermodynamics) than those based on correlations or non-physical sense solutions (Friedman, 1999).

2.2. Avoid "Computerization"

To solve a given problem, engineers need to apply the ***EBC*** approach suggested by Richalet, 2000, rather than the straight use of ***C***. Follow the ***EBC*** procedure implies:

- First, use your ***E***ye to inspect trends, data, process diagrams, DCS screens...
- Second, use your ***B***rain to carefully understand and design the real time system and select or design the proper software tool.
- Finally, only at the end, use the ***C***omputer to solve the problem.

The enthusiastic but inexperienced engineers tend to go straight to the ***C***omputer, forgetting or minimizing to analyze and perfectly understand the problem to be solved, their real handles and constraints.

2.3. Balance Model Complexity Against Expected Results

Model complexity must be balanced against expected benefits. Remember: "the simplest is generally the best". Engineering judgment and experience plays a key role in this task. Operators and Engineers will appreciate any efforts towards simplification, especially for future re-training and system maintenance.

2.4. Real Time Data Seldom Can be Used "as is"

Carefully designed and engineered "safety envelops" must be built around real time online data, in order to validate all the signals, in and out of the applications. Minimum and maximum limits for each variable, controller windup identification and associated valve openings and controllers statuses must be combined to properly identify and isolate measurement problems. Steady state detection procedures must also be considered to warn Operators or directly skip an optimization run if the Process is unstable.

What should a system do if bad data is detected? If the bad data feeds "critical" variables of the system, then its execution should be suspended and Operators warned. If bad data is related to "noncritical" variables then some kind of backup or default values could be used but such an action needs to be defined on each individual variable. The implementation of data validation procedures is a key activity for every real time project.

2.5. Pricing Information Must be Updated

When economics is the main driver to apply advanced process engineering real time systems, prices used for the optimization or objective function calculations must be diligently updated. When possible, manual data entering should be avoided. Automatic pricing information must be implemented acquiring the prices directly from process information systems or web pages (some examples we have used in the past include the sites of ERCOT in Texas, USA and OMEL in Spain, for electricity market spot prices).

2.6. Have Operators Always in Mind

From the very beginning try to involve Operators, inform them about the project objectives and solve their concerns. A common question is about being replaced by the new system and fired after completion. Second in rank is their worries regarding if the new system will increase their burden. First question is usually outside our control but we never have seen a single Operator being fired due to our projects, most of them very successful. With respect to the second concern, an effort needs to be done to not increase Operator's burden, implementing automatic data updates or procedures to

minimize his/her manual intervention.
Clear and easy operating interfaces need also to be built. Operators must be properly trained and assisted right by the DCS console after being left alone with the new system. It is always a good idea to spend many hours at the DCS console with the Operators to discuss the proposed solutions and technology, explaining to them what is needed so they will accept the introduction of the new technology. Properly designed Operating manuals must be provided and updated as frequently as needed.

2.7. Do not "Prototype" but Deploy Well Engineered Systems

Do not expose Operators nor final users to "interim or prototyping versions" but only to finished, very well built models, including the Operating interface and true handles limits and constraints. Operator's general attitude is to closely adhere to their first impression. There is a single "silver bullet" to use with them: they will never forget if early disappointed or frustrated when trying to use the system for the first time. Moreover, many of them will became "vaccinated" against the technology and refuse to use it, subtle or frontally, any more.

2.8. Document the Application and Maintain it "Forever Green"

Engineering documentation must be provided and updated as soon as changes are implemented in the model or to the Operating interface. The application needs to be maintained "forever green", reflecting continuously the current Process handles and constraints. Changes on the process side must be immediately reflected in the model. Any process change needs to be immediately taken into account because they may impact the way the online variables are considered into the models.
Local resources need to be properly trained and technology providers need to assist and remotely maintain the applications.
As a sidenote it is important to mention the huge impact of remote access capabilities on systems performance. Only a few years ago just a minor model adjustment would take a lot of time for the new information to be exchanged back and forth. Nowadays, VPN or Webex like secure remote access can be easily implemented and solutions can be provided in minutes, instead of hours or even days.

3. Industrial Projects Execution

To cite only a few examples with which the author has been exposed, the following technologies that are widely used in the Process Industry rely heavily on real time data:

- Model Based Predictive Control (MPBC), for example A. Sanz et al., 2005
- Process Monitoring, for example A. Heins et al., 2007
- Real Time Optimization, for example M. Kihn et al. for open loop, 2008 and D. Ruiz et al., 2008 for closed loop applications
- Abnormal Situation Management, for example D. Ruiz et al., 2002

All of them include software where very efficient and robust numerical methods must deal with real time, online data and cope with the unavoidable uncertainties and errors always present in them. But even if the piece of software is implemented soundly this is not a guarantee of success. Only a well executed engineering project, carried on by experienced Application Engineers will guarantee the success.

3.1. Collect the Required Information

A document is submitted to the Site with all the informational requirements for the real time project. It is very important to define a single project owner from the Site side, which will act as a single interface to provide the needed information and coordinate all the project steps.

3.2. Kick-Off Meeting
Prior to the Kick-Off Meeting, the provided information will be reviewed to have a better understanding of the Site process, main constraints and economic position. Additional questions or clarifications are sent to the Site regarding particular issues, as required. The operating procedures, optimization and control strategies should also be discussed. Operators must be introduced to the project goals and to the technology to be deployed.

3.3. Software Installation and Real Time Data Gathering
The software is configured and licensed on the Application server PC. It will also be connected to the OPC data server. Remote access to the model also needs to be configured at this time and needs to be available throughout the rest of the project, including maintenance and sustainability periods.

3.4. Functional Design Specification
Based on the information obtained during the Kick-Off meeting, a Functional Design Specification document is prepared and approved by the Site. This document becomes the blueprint for the entire project.

3.5. Model Building and Configuration
During this stage, the model and the report are built working remotely on the Application server. The model is built and grows with continuous access to online real time data. Eventually, a second trip to the facility would occur during this stage and would be used for mid-term review of the model and optimization.
The real time Application will run routinely, but optimization recommendations will still not be implemented but reviewed by the Engineering staff.
During this stage, a few experienced Operators can help to review the system results.

3.6. Operators Training and Commissioning
Site engineers will train the Operations staff to start and stop the Application, understand issues and potential problems and to implement the recommendations, when applicable. At the end of this period, Engineering and Operations staff begin with the daily use of the Application, which is transferred to its final owners.
Project documentation will be provided and a benefits report should be prepared based on the current results. It should be submitted to the appropriate Management level, to provide the proper feedback about the implemented technology and project execution quality.

3.7. Support and Sustainability
After project completion usually a routine remote check of the health of the online system is performed with the purpose of ensuring the Site continues to benefit from its implementation over time. As a part of the program, a regular review regarding the status of the model and the optimization and simulation solutions is proactively done on a periodic basis. A regular report can then be generated and distributed to the appropriate focal point.

4. Conclusions

Current data availability in the Plant Information systems, widely available at the Process Industries, provides the foundation on top of which you can build Applications to better operate, control and optimize these processes.
Inherent problems of real time data need to be addressed by carefully building a safety net around them, in order to avoid the use of unreliable and bad information.

The success of an industrial real time project deployment can be highly increased by following a set of simple rules and ensuring a proper knowledge transfer to both, Operators and Engineers.
Application maintenance and support helps to sustain the economic benefits after commissioning allowing them to be always in good shape and still used and appreciated by the owners many years after handing the project over to them.

References

OPC Foundation, 2009, http://www.opcfoundation.org/

Y. Z. Friedman, 1999, Avoid Advanced Control Project Mistakes, Chapter of the book Advanced Process Control and Information Systems for the Process Industries, L. Kane (editor) , ISBN 0-88415-239-1, Gulf Publishing Co., Houston, Texas, Pages 42-48

J. Richalet, 2000, Personal communication

A. Sanz, A. Cardete, R. Lucio, R. Martínez, S. Muñoz, D. Ruiz, C. Ruiz, O. Gerbi and J. Papon, 2005, Model Based Predictive Control Increases Batch Reactor Production, , Hydrocarbon Processing, Vol 84, Nro. 5, Pages 61-65

A. Heins, R. Veiga, C. Ruiz, A. Riera, 2007, Fouling Monitoring and Cleaning Optimisation in a Heat Exchanger Network of a Crude Distillation Unit, Congress: Heat Exchanger Fouling and Cleaning VII, Tomar, Portugal, July

M. Kihn, D. Ruiz, C. Ruiz, A. García Nogales, 2008, Online Energy Costs Optimizer at Petrochemical Plant, Hydrocarbon Engineering, May, Pages 119-123

D. Ruiz, C. Ruiz, 2008, Closed loop energy real time optimizers,, European Refining Technology Conference (ERTC) Annual Meeting Energy Workshop, Vienna, Austria, November

D. Ruiz, Ch. Benqlilou, J. Nougués, L. Puigjaner and C. Ruiz, 2002, Proposal to Speed Up the Implementation of an Abnormal Situation Management in the Chemical Process Industry, Ind. Eng. Chem. Res., Vol 41, Pages 817-824

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Current trends in parallel computation and the implications for modeling and optimization

John D. Siirola

Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM, 87185

Extended Abstract

Process Systems Engineering (PSE) is built on the application of computational tools to the solution of physical engineering problems. Over the course of its nearly five decade history, advances in PSE have relied roughly equally on advancements in desktop computing technology and developments of new tools and approaches for representing and solving problems (Westerberg, 2004). Just as desktop computing development over that period focused on increasing the net serial instruction rate, tool development in PSE has emphasized creating faster general-purpose *serial* algorithms. However, in recent years the increase in net serial instruction rate has slowed dramatically, with processors first reaching an effective upper limit for clock speed and now approaching apparent limits for microarchitecture efficiency. Current trends in desktop processor development suggest that future performance gains will occur primarily through exploitation of parallelism. For PSE to continue to leverage the "free" advancements from desktop computing technology in the future, the PSE toolset will need to embrace the use of parallelization.

Unfortunately, "parallelization" is more than just identifying multiple things to do at once. Parallel algorithm design has two fundamental challenges: first, to match the characteristics of the parallelizable problem workload to the capabilities of the hardware platform, and second to properly balance parallel computation with the overhead of communication and synchronization on that platform. The performance of any parallel algorithm is thus a strong function of how well the characteristics of the problem and algorithm match those of the hardware platform on which it will run. This has led to a proliferation of highly specialized parallel hardware platforms, each designed around specific problems or problem classes.

While every platform has its own unique characteristics, we can group current approaches into six basic classes: symmetric multiprocessing (SMP), networks of workstations (NOW), massively parallel processing (MPP), specialized coprocessors, multi-threaded shared memory, and hybrids that combine components of the first five classes. Perhaps the most familiar of these is the SMP architecture, which forms the bulk of current the desktop and workstation market. These systems have multiple processing units (processors and/or cores) controlled by a single operating system image and sharing a single common shared memory space. While SMP systems provide only a modest level of parallelism (typically 2-16 processing units), the existence of shared memory and full-featured processing units makes them perhaps the most straightforward development platform. A challenge of SMP platforms is the discrepancy between the speed of the processor and the memory system: both latency and overall memory bandwidth limitations can lead to processors idling waiting for data. Clusters, a generic term for coordinated groups of independent computers (*nodes*) connected with high-speed networks, provide the opportunity for a radically different level of parallelism, with the largest clusters having over 25,000 nodes and 100,000

processing units. The challenge with clusters is memory is distributed across independent nodes. Communication and coordination among nodes must be explicitly managed and occurs over a relatively high latency network interconnect. Efficient use of this architecture requires applications that decompose into pseudo-independent components that run with high computation to communication ratios. The level to which systems utilize commodity components distinguishes the two main types of cluster architectures, with NOW nodes running commodity network interconnects and operating systems and MPP nodes using specialized or proprietary network layers or microkernels. Specialized coprocessors, including graphics processing units (GPU) and the Cell Broadband Engine (Cell), are gaining popularity as scientific computing platforms. These platforms employ non-general purpose dependent processing units to speed fine-grained, repetitive processing. Architecturally, they are reminiscent of vector computing, combining very fast access to a small amount of local memory with processing elements implementing either a single-instruction-multiple-data (SIMD) (GPU) or a pipelined (Cell) model. As application developers must explicitly manage both parallelism on the coprocessor and the movement of data to and from the coprocessor memory space, these architectures can be some of the most challenging to program. Finally, multi-threaded shared memory (MTSM) systems represent a fundamental departure from traditional distributed memory systems like NOW and MPP. Instead of a collection of independent nodes and memory spaces, an MTSM system runs a single system image across all nodes, combining all node memory into a single coherent shared memory space. To a developer, the MTSM appears to be a single very large SMP. However, unlike a SMP that uses caches to *reduce* the latency of a memory access, the MTSM *tolerates* latency by using a large number of concurrent threads. While this architecture lends itself to problems that are not readily decomposable, effective utilization of MTSM systems requires applications to run hundreds – or thousands – of concurrent threads.

The proliferation of specialized parallel computing architectures presents several significant challenges for developers of parallel modeling and optimization applications. Foremost is the challenge of selecting the "appropriate" platform to target when developing the application. While it is clear that architectural characteristics can significantly affect the performance of an algorithm, relatively few rules or heuristics exist for selecting a platform based solely on application characteristics. A contributing challenge is that different architectures employ fundamentally different programming paradigms, libraries, and tools. Knowledge and experience on one platform does not necessarily translate to other platforms. This also complicates the process of directly comparing platform performance, as applications are rarely portable: software designed for one platform rarely compiles on another without modification, and the modifications may require a redesign of the fundamental parallelization approach. A final challenge is effectively communicating parallel results. While the relatively homogenous environment of serial desktop computing facilitated extremely terse descriptions of a test platform, often limited to processor make and clock speed, reporting results for parallel architectures must include not only processor information, but depending on the architecture, also include operating system, network interconnect, coprocessor make, model, and interconnect, and node configuration.

There are numerous examples of algorithms and applications designed explicitly to leverage specific architectural features of parallel systems. While by no means comprehensive, three current representative efforts are the development of parallel branch and bound algorithms, distributed collaborative optimization algorithms, and multithreaded parallel discrete event simulation. PICO, the Parallel Integer and

Combinatorial Optimizer (Eckstein, et al., 2001), is a scalable parallel mixed-integer linear optimizer. Designed explicitly for cluster environments (both NOW and MPP), PICO leverages the synergy between the inherently decomposable branch and bound tree search and the independent nature of the nodes within a cluster by distributing the independent sub-problems for the tree search across the nodes of the cluster. In contrast, agent-based collaborative optimization (Siirola, et al., 2004, 2007) matches traditionally non-decomposable nonlinear programming algorithms to high-latency clusters (e.g. NOWs or Grids) by replicating serial search algorithms intact and unmodified across the independent nodes of the cluster. The system then enforces collaboration through sharing intermediate "solutions" to the common problem. This creates a decomposable artificial meta-algorithm with a high computation to communication ratio that can scale efficiently on large, high latency, low bandwidth cluster environments. For modeling applications, efficiently parallelizing discrete event simulations has presented a longstanding challenge, with several decades of study and literature (Perumalla, 2006). The central challenge in parallelizing discrete event simulations on traditional distributed memory clusters is efficiently synchronizing the simulation time across the processing nodes during a simulation. A promising alternative approach leverages the Cray XMT (formerly called Eldorado; Feo, et al. 2005). The XMT implements an MTSM architecture and provides a single shared memory space across all nodes, greatly simplifying the time synchronization challenge. Further, the fine-grained parallelism provided by the architecture opens new opportunities for additional parallelism beyond simple event parallelization, for example, parallelizing the event queue management.

While these three examples are a small subset of current parallel algorithm design, they demonstrate the impact that parallel architectures have had and will continue to have on future developments for modeling and optimization in PSE.

Keywords: parallel computing, optimization, discrete event simulation

Acknowledgements

This work was performed in part at Sandia National Laboratories. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

J. Eckstein, C. A. Phillips and W. E. Hart, PICO: An object-oriented framework for parallel branch-and-bound, In *Proc Inherently Parallel Algorithms in Feasibility and Optimization and Their Applications*, Elsevier Scientific Series on Studies in Computational Mathematics, 219-265, 2001.

J. Feo, D. Harper, S. Kahan, and P. Konecny. Eldorado. In *Proc. 2nd Conf. Computing Frontiers*, 28–34, 2005.

K. S. Perumalla. Parallel and distributed simulation: traditional techniques and recent advances. In *Proc. 2006 Winter Simulation Conference*. 84-95, 2006.

J. D. Siirola, S. Hauan, and A.W. Westerberg. Computing Pareto fronts using distributed agents. *Computers & Chemical Engineering*, 29(1):113-126, 2004.

J. D. Siirola and S. Hauan. Polymorphic Optimization. *Computers and Chemical Engineering* 31(10):1312–1325, 2007.

A. W. Westerberg. A retrospective on design and process synthesis. *Computers & Chemical Engineering*, 28(4):447-458, 2004.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Waste-to-Energy Systems as Examples of Efficient Process Systems Engineering Approach

Petr Stehlík

Institute of Process and Environmental Engineering (VUT-UPEI), Faculty of Mechanical Engineering, Brno University of Technology, Technická 2, 616 69, Brno, Czech Republic

Abstract

Waste-to-energy systems may serve as a relatively simple, but all the more instructive example of industrial processing plant. Their design, operation and control are the subject of significant amount of research efforts worldwide. The present contribution briefly summarizes several recent advances in these areas in the context of respective computing-based methodologies. Practical and sophisticated approaches are demonstrated through industrial examples.

Energy production from contaminated biomass is a subset of waste-to-energy systems, but it is still better viewed as a separate application area. Due to the present rocketing importance of renewable energy sources, contaminated biomass becomes an interesting distributed energy source. Therefore we dedicate part of the paper to computational support in the design of a biomass combustion process.

Keywords: waste to energy, simulation, process design.

1. Introduction

Basic systems of the unit for the thermal processing of waste – incineration itself, heat recovery system, off-gas cleaning system – are briefly described in the first section as well as control and maintenance. A system for overall performance evaluation of waste-to-energy processes is the subject of the next section. After that, the design of the waste treatment technologies (namely decision support and optimization) is discussed with focus on available computational tools. Finally, computational support in the design of contaminated biomass combustor is presented.

2. Operation control and maintenance

Some types of waste (e.g. municipal solid waste in developed countries) have a high energy potential and it thus seems more appropriate to consider them rather as alternative fuels than waste (Bebar et al. 2005). The waste energy potential may be exploited for example in thermal waste treatment units (municipal and hazardous waste incinerators, see Figure 1), where heat produced by waste combustion is converted to utilizable forms of energy (e.g. steam, electricity, heating of utilities etc.) (European IPPC Bureau 2005). In the design phase of such plants it is advantageous to utilise software tools such as TDW (Puchyr et al. 1998) to make material and energy balances necessary for the selection of appropriate technologies and determination of other important design parameters. Incineration is however a complicated process, which besides useful energy generates also a number of polluting substances that have to be collected and neutralised (Dvorak et al. 2009). These include namely solid particulate matter, nitrogen oxides (NO_x), carbon monoxide, sulphur oxides, hydrogen chloride and

fluoride, volatile organic compounds (VOCs), heavy metals (Cd, Hg, Cr, Zn, Cu, Pb) and also persistent organic pollutants (POPs) like PAH, PCB and PCDD/F.

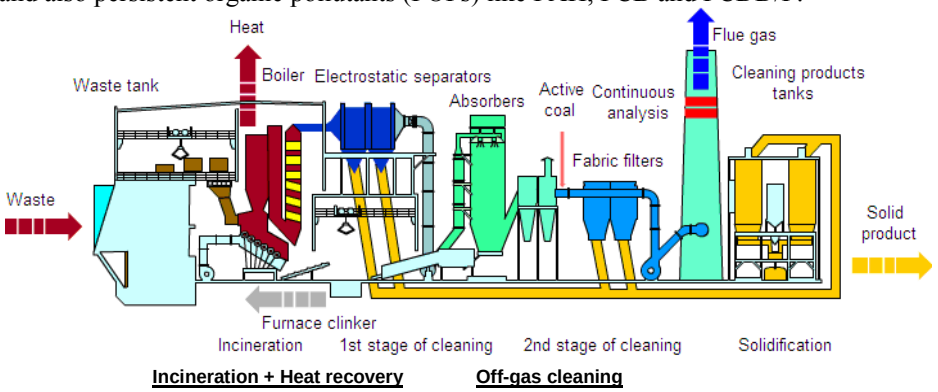


Figure 1 Schematic layout of typical unit for the thermal processing of waste

During the combustion process itself it is necessary mainly to keep temperature in the primary combustion chamber above 850°C, while in the secondary combustion chamber (afterburner) the required temperature is 1100°C and a 2 s residence time (Council Directive 2000/76/EC). These measures are designed to reduce mainly carbon monoxide VOCs and POPs. After completing the thermal treatment, the hot flue gas is utilised for preheating of combustion air and for the production of steam in a waste heat boiler (heat recovery steam generator, HRSG). Cooled flue gas after that undergoes mechanical and chemical cleaning, which removes polluting substances down to levels required by the applicable legislation.

Provided that all these steps are performed properly, it is then possible to re-classify waste incineration from a waste disposal technology to waste-to-energy process, as it makes full use of the waste energy potential and generates utilizable forms of energy.

3. Waste-to-energy process simulation

Systems for thermal treatment of waste have developed into complex processes consisting of several subsystems such as incineration section, heat recovery section, etc., as described above. Simulation software systems based on mathematical modelling support the design of up to date units for the thermal processing of waste (Cimini et al. 2005, Grieco and Poggio 2009). However, in order to be able to easily implement innovative processes and equipment it may be beneficial to create customized in-house packages (Friedl et al. 2008, Stehlík et al. 2000).

The author's group has recently developed a so-called W2E (Waste-to-Energy) software, which features such a supporting tool for waste treatment process simulation (Pavlas et al. 2005). It is written in Java and provides user-friendly and intuitive interface. The principle of modelling and simulation is the same as in other similar systems (Felder and Rousseau 1999). The software easily incorporates emerging new technologies, provides computer aid in the research and development of thermal treatment technologies and enables presentation of results on a professional level. Flexibility of the code is provided by its specific architecture ensuring that model modifications consist of few steps and no advanced programming skills are needed. The development of the W2E program is still in progress.

4. Design decision support, optimization

Computing-based methodologies can substantially contribute to improved design and bring together researchers and practitioners. Let us demonstrate this statement through the following examples.

4.1. Hot gas database

Heat recovery is one of the basic systems of units for the thermal processing of waste. Heat recovery system can be generally realized through a wide range of heat exchangers from conventional ones to specialized types. However, in real applications of units for the thermal treatment of waste, special types of heat exchangers are frequently proposed and designed. The selection between conventional and special type of heat exchanger depends mainly on the specific purposes of the heat exchange application in units for the thermal treatment of waste. Energy contained in flue gas is utilized for air pre-heating, steam generation, water heating or technological purposes. Unless due respect is paid to the specific features of process fluids in these processes, serious problems during operation may follow, including excessive fouling, structural damage due to thermal expansion and leakages.

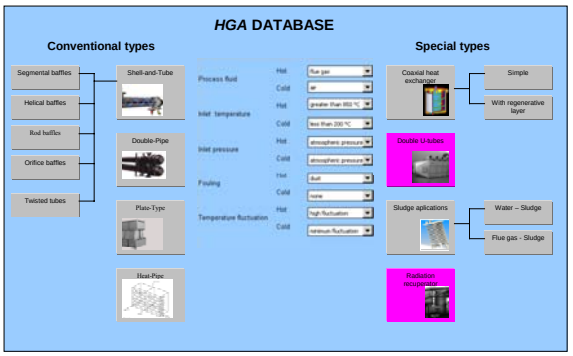


Figure 2 Main user interface window of HGA database

These were the key reasons that led to the development of a multi-purpose computational system equipped with a module featuring database for hot gas applications (HGA Database, see Figure 2), where data concerning both the common and specific types of heat exchangers are contained. HGA database (Kilkovský et al. 2007) is a decision support tool which helps to find optimum design solution as well as to avoid possible serious operating problems. Conventional heat exchangers are given preference in the selection strategy where possible, but in some cases there is no other choice than to select a novel type of heat exchanger (Stehlík 2005).

4.2. SNCR system design

Substantial efforts were recently reported to develop chemical kinetic mechanisms of the SNCR process that would be suitable for coupling with CFD codes e.g. by Cremer et al. 2000 or Saario and Oksanen 2008. Building on these results, successful simulations of industrial units e.g. by Shin et al. 2007 are the natural outcome and a proof that computational support in SNCR system design is both viable and reliable, in spite of the fact that most of such applications naturally do not appear in the open literature.

A comprehensive review of the SNCR technology covering both practical and modelling issues provides a paper by Tayyeb Javed et al. 2007. We may conclude this section by stating that computational analyses of SNCR systems in waste incinerators,

power plants (fired by pulverized coal) and biomass boilers are already applicable in the plant design phase.

4.3. Secondary combustion chamber design

Secondary combustion chambers or afterburners serve typically in hazardous waste incinerators to ensure complete decomposition of the products of the first stage reactor, which may be either a rotary kiln or a gasifier. Their design and operation optimization has been the subject of several experimental and CFD (computational fluid dynamics) studies. Most of them concentrated on the functionality (i.e. pollutant emissions), but some also discussed the possibility to optimize the design with respect to investment costs. There may also be observed a shift from experimental investigations of Watanabe et al. 1996 or Öberg 2003 to computational studies like those by Ficarella and Laforgia 2000 or Hájek et al. 2006.

The most notable difference between the experimental and computational works is that the latter typically examine several alternative design options, whereas all experimental works focused on operational parameters. Thanks to its cost-efficiency the computational approach may thus yield more fundamental improvements than experiments.

5. Design of contaminated biomass combustion unit

A new modern unit for energy production from contaminated biomass and/or alternative fuels, shown in Figure 3, is notable for its efficient integration of proven technological solutions using sophisticated computational support. The newly designed unit reflects present day's needs and the necessity to utilize waste biomass or biomass-based alternative fuels for the production of energy in ways regardful of the environment.

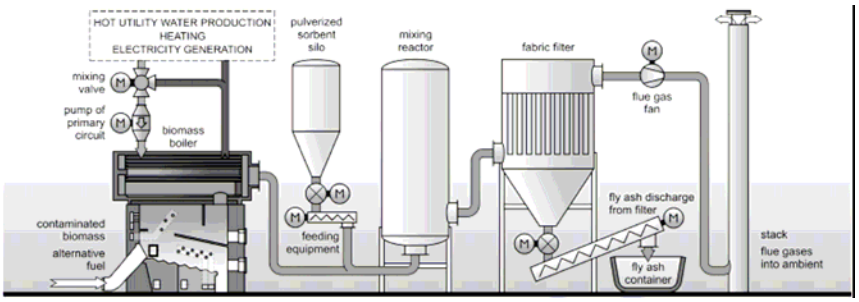


Figure 3 Schematic layout of the new unit for the utilization of contaminated biomass

Key features of the technology include optimized design, maximized heat utilization and minimized pollutant emissions. These features are discussed in some more detail in the following paragraphs.

5.1. Design optimization

During development of the unit, great attention has been paid to the combustion chamber. The locations and number of secondary air ports and flue gas recirculation were optimized using CFD (Computational Fluid Dynamics) modelling. An illustration of the CFD results is included in Figure 3. The issues addressed in the optimization process had a common target – optimum combustion conditions.

5.2. Heat utilization

Configuration of the heat utilization system (the boiler's main heat exchanger and utility extraction system to individual consumers or networks) is governed by specific needs in each individual application.

Condensing stage may be included in the recuperative unit. The heat exchanger is in this case made of stainless steel, with vertical tubes in which flows the flue gas and on the outside is the heated medium. Condensate collected at the outlet is sprayed back into the flue gas at the inlet section of the heat exchanger, thus washing the flue gas side of heat exchanging surfaces and preventing fouling.

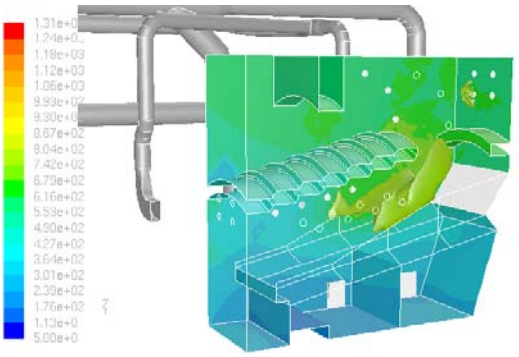


Figure 4 An illustrative display of temperature distribution on inner walls and iso-surface of 800°C in the combustion chamber (at 25% of the nominal firing capacity)

5.3. Flue gas cleaning

Combustion of contaminated biomass produces flue gas, which may contain various harmful compounds. Due to this fact, the unit must be equipped with appropriate flue gas cleaning equipment.

Formation of CO and volatile organic compounds is generally minimised by optimised design and operation of the combustion chamber, as these species are the products of incomplete combustion

Harmful compounds as for example nitrogen oxides, acid compounds (SO2, HCl, HF), heavy metals and fine particulate matter (fly ash) may also be produced by the combustion process. Due to this fact, the unit must be equipped with appropriate flue gas cleaning equipment designed specifically for the collection of those polluting substances that may appear in each individual application (e.g. injection of powder of sodium hydrogen carbonate (NaHCO3) for the removal of acid compounds, adsorption of traces of heavy metals on another pulverised adsorption agent, surface filtration in fabric filter, collecting both agents together with fly ash and products of neutralisation).

6. Conclusion

The present contribution provides a demonstration of up to date efficient and practically proven process system engineering approach in the field of environmental protection. Several examples are shown and novel features are highlighted. Due to limited space, references are provided instead of detailed discussions.

References

L. Bebar, P. Stehlík, L. Havlen, J. Oral, 2005, Analysis of Using Gasification and Incineration for Thermal Processing of Wastes, Applied Thermal Engineering, 25, pp. 953 – 1096

S. Cimini, M. Prisciandaro, D. Barba, 2005 Simulation of a waste incineration process with flue-gas cleaning and heat recovery sections using Aspen Plus, Waste Management, 25, pp. 171–175

Council Directive 2000/76/EC of the European parliament and of the council of 4 December 2000 on the incineration of waste

M.A. Cremer, C.J. Montgomery, D.H. Wang, M.P. heap, J.-Y. Chen, 2000, Development and implementation of reduced chemistry for computional fluid dynamics modeling of selective non-catalytic reduction, Symposium (International) on Combustion, 28(2), pp. 2427-2434

R. Dvorak, T. Parizek, L. Bebar, P. Stehlík, 2009, Incineration and gasification technologies completed with up-to-date off-gas cleaning system for meeting environmental limits, Clean Technologies and Environmental Policy, 11(1), p. 95

European IPPC Bureau, 2005, Reference Document on the Best Available Techniques for Waste Incineration, Brussles, available online at <http://eippcb.jrc.es>

R. M. Felder, R. W. Rousseau, 2005, Elementary Principles of Chemical Processes, Third Edition, John Wiley & Sons

A. Friedl, A. Miltner, P. Schausberger, 2008, Modelling and simulation of flue gas cleaning processes, 18th International Congress of Chemical and Process Engineering CHISA 2008, Prague, Czech Republic

A. Ficarella, D. Laforgia, 2000, Numerical simulation of flow-field and dioxins chemistry for incineration plants and experimental investigation, Waste Management, 20(1), pp. 27-49

E. Grieco, A. Poggio, 2009, Simulation of the influence of flue gas cleaning system on the energetic efficiency of a waste-to-energy plant, Applied Energy, doi:10.1016/j.apenergy.2008.12.035

Hájek J., 2008, Computational fluid dynamic simulations in thermal waste treatment technology - Design, optimisation and troubleshooting. Energy 33 (6), pp. 930-941

Hajek J., Petr P., Sarlej M., Piskovsky M., Parizek T., Bebar L., Stehlík P., 2006, Optimization of Afterburner Chamber Using CFD, 17th International Congress of Chemical and Process Engineering CHISA 2006, Prague,

Kilkovsky B., Pavlas M., Jegla Z. and Stehlík P.: DatabaseHGA: Common ans Specific Types of Heat Exchangers for Hot Gas Applications, 2007, In: Heat Transfer in Components and Systems for Sustainable Energy Technologies, Heat-SET 2007, pp. 18-20

T. Öberg, 2003, Optimization of an industrial afterburner, Journal of Chemometrics, 17(1), pp.5-8

Pavlas M., Stehlík P., Oral J., 2005, Computational approach for efficient energy utilization in process industry, 8th Conference on Process Integration, Modelling and Optimisation for Energy Saving and Pollution Reduction Press 05, Giardini di Naxos- Taormina, Italy

R. Puchyr, J. Oral, J. Richter, P. Stehlik, 1998, Software TDW for Simulation of Thermal Disposal of Wastes Processes, 13-th International Congress of Chemical and Process Engineering CHISA ´98, Prague, Czech Republic

A. Saario, A. Oksanen, 2008, Comparison of Global Ammonia Chemistry Mechanisms in Biomass Combustion and Selective Noncatalytic Reduction Process Conditions, Energy & Fuels, 22(1), pp. 297-305

M.S. Shin, H.S. Kim, D.S. Jang, 2007, Numerical study on the SNCR application of space-limited industrial boiler, Applied Thermal Engineering, 27(17-18), pp. 2850-2857

Stehlík, P. Heat Transfer as Important Subject in Waste to Energy Systems. 2005, In proceedings: Heat Transfer in Components and Systems for Sustainable Energy Technologies, Grenoble. pp. 11-21, ISBN: 2-9502555-0-7

P. Stehlík, R. Puchyr, J. Oral, 2000, Simulation of Processes for Thermal Treatment of Wastes, Waste Management, 20, pp. 435-442

M. Tayyeb Javed, N. Irfan, B.M. Gibbs, 2007, Control of combustion-generated nitrogen oxides by selective non-catalytic reduction, Journal of Environmental Management, 83(3), pp. 251-289

T. Watanabe, M. Asai, T. Kondo, M. Shimizu, Y. Takeuchi, H. Aramaki, M. Naito, 1996, An advanced fluidized-bed swirl incinerator for dioxin control during municipal waste disposal, Chemosphere, 31(1), pp. 177-187

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Process Systems Engineering: From Solvay to the 21st Century. A History of Development, Successes and Prospects for the Future

George Stephanopoulos

A.D. Little Professor of Chemical Engineering, Massachusetts Institute of Technology
77 Massachusetts Avenue, Cambridge MA 02139, USA

Abstract

The term Process Systems Engineering (PSE) is relatively recent, but the engineering of processing systems is as old as the beginning of the chemical industry, around the beginning of the 19th century. Initially, the practice of PSE was informal and as time went on it was formalized in progressively increasing degrees. Today, it is solidly founded on engineering sciences and an array of systems theoretical methodologies and approaches. This paper provides an overview of the history of Process Systems Engineering (PSE), i.e. its origin and evolution, a brief illustration of its tremendous accomplishments in the development of modern chemical industry, and an outline of the role it can play in addressing the societal problems that we face today; securing sustainable production of chemicals and materials for the human wellbeing, alternative energy sources, improving the quality of life and of our living environment. Finally, this paper proposes PSE as the foundational underpinning of modern chemical engineering; the only one that ensures the discipline's cohesiveness in the years to come.

Keywords: History of Process Systems Engineering, Process Development, Process Design, Process Control, Process Safety.

Introduction

PSE is the field that encompasses the activities involved in the *Engineering of Systems involving Physical, Chemical, Biological Processing Operations.* These systems, whose variety and purpose are very broad, include but are not limited to the following: Processing plants producing energy, chemicals and materials that societal needs require; manufacturing of a very broad variety of foods and devices that enhance quality of life; therapeutic products and processes to treat human diseases; and systems which ensure the quality of the environment in which humans, plants and animals live.

The approaches of today's formalized PSE are driven by two considerations: (a) The interest is in the "behaviour" of the system as a whole. (b) The emphasis is on studying how the components of the system and their interactions contribute to the overall "behaviour" of the system. The implications are clear: PSE is the integrative complement of all sub-areas of chemical engineering and the one that provides the significance and criticality, or absence thereof, of "local" engineering questions. For example, the operational and economic "behaviour" of a chemical process is the integrated effect of all its unit operations, whose criticality in determining the behaviour of the process can be systematically analyzed. Also, a drug is not just its active

component but a system with three interacting elements; active component, formulation additives, and delivery vehicle. Similar interactions one can easily recognize in the components of a product and associated fabrication lines (e.g. digital processors, LCD or OLED displays, batteries and fuel cells); the components of global production and transportation of energy systems; the constituents of energy and the environment policy-making systems, and others. It is fair to say that no manufacturing process or product reaches the market without the integrative scrutiny of systems engineering.

Mathematical modeling of the components' behaviour and their interactions is the essential tool of modern PSE activities, which along with the parallel development of computers, numerical methods for the solution of mathematical models, optimization theory, and dynamic systems and control theories constitute the formal framework of modern PSE. Consequently, the historical evolution of PSE is very closely related to the evolving needs of the chemical industry, the academic content of chemical engineering and a variety of supporting disciplines.

Today, PSE is well established as core capability in chemical engineering. Unfortunately, its educational foundation remains fragmented and a significant number of educational institutions do not recognize it as foundational component of the chemical engineering curriculum. As chemical engineering tries to redefine its intellectual core and disciplinary cohesiveness, and battles centrifugal forces to the interface with other disciplines, PSE retains its quintessential integrative character and provides the core glue to the development of chemical engineering in the 21st century.

1860s to 1920s: The Formative Period

The chamber and Leblanc processes for the production of sulfuric acid and sodium carbonate, respectively, in mid-18th century England are the most celebrated examples of the early chemical industry. Both were batch, dirty, and rather costly processes, and almost from their inception attracted the interest of businessmen, scientists, and engineers, who were challenged to develop processes with lower unit production cost, cleaner and safer. As people started looking into ways of improving the economics, safety and environmental impact of these processes, two things became clear very early on: (a) The manufacturing system had to be viewed as a system of interconnected operations. (b) The components of the manufacturing system, e.g. the chemistries at the core of the process and the separation units were determining the behavior of the system as a whole, i.e. economics, safety and environmental impact.

These observations set off the period of PSE that I will call the "Formative Period**",** and led to the initial informal delineation of the elements of what later on we would call "Process Development", the core element of PSE. It focused on the essential definition of the processing system, i.e. chemistries, reactors, separators, their integration into a productive whole, and the mechanical design of the equipment where the operations were taking place. The most celebrated example of this attitude is the *Solvay Process*, patented by the Belgian Ernest Solvay in 1861. Solvay may be seen as the first Process Systems Engineer. His 1872 ammonia-based soda production process was a breakthrough. He synthesized the process by interconnecting distinct operations of gas-liquid contacting, reaction with cooling, and separations; he invented new types of equipment for integrating these operations and carrying them out continuously on a large scale. He himself dealt with all aspects of an integrated chemical processing system: the chemistry, the materials handling, and the engineering of an integrated processing system, such as operating conditions and the design of specific equipment. The Solvay process may be seen the model process in which all aspects of PSE are explicitly delineated and directly addressed, in a fairly systematic and more importantly,

system-wide basis. Solvay went beyond the parts to the whole as the essence of the manufacturing system. The result was: a continuous process with careful integration of chemical and physical operations; use of recycles for improved yields and reduction of wasted raw materials; reduction of environmental pollution; and significant cost efficiency. Skip Scriven in his marvelous 1988 paper "On the Emergence and Evolution of Chemical Engineering" declared Solvay as the epitome of what we would later call "chemical engineer". It is therefore fair to argue that the original chemical engineers were in fact process systems engineers.

Over the following 50 years Solvay's essential guidelines for process development were adopted and further improved by many in England and America. Typical examples include: Rudolf Knietz's sulfuric acid process; Herbert Dow's inventive integration of electrochemistry and air-blowing processes that led to processes for bromide products, sodium hydroxide, chlorine, and many of their derivatives ; Charles Hall's production of aluminium; Comprehensive starch processes integrating chemical and fermentation operations; Carl Bosch's historic breakthrough process for ammonia, using Haber's catalytic chemistry; Dubbs' (Universal Oil) continuous thermal cracking process, etc.

By the late 1910s early 1920s, the process development experience of the previous 50 years has set the general scope of the PSE core. A chemical process is a system and its overall behavior is the coordinated effect of basic "Unit Operations"; a concept introduced by George Davis (Manchester Lectures, 1888) and a term coined by Arthur D. Little (MIT, 1915). A certain nascent standardization of the process development process is taking form: focus on cost, environment and safety; evolve from chemistries to processing systems, from batch to continuous processes; introduce material recycles and process system-wide management of energy; recognition that further improvements need better quantitative understanding of the process system at large.

Consequently, during the period 1860s to 1920s, the activities that today we characterize as Process Development (the core activity of PSE) define the central character of what is evolving to become Chemical Engineering. Thus, it is fair to say that during this period a Chemical Engineer is essentially a Process Systems Engineer.

1920s to 1950s: The Waiting Period

The identification of a chemical engineer as a process systems engineer becomes weaker and more tenuous during the ensuing period of 1920s to 1950s, which I call the "Waiting Period". During this period the focus is on the components of a process, i.e. Unit Operations, not the overall process as an integrated system. Continuous deepening in the understanding of the phenomena that take place in basic unit operations, leads to the invention of new unit operations and continuous improvements in the quantitative description of their behaviour. The foundational 1926 book "Principles of Chemical Engineering" by Walker, Lewis and McAdams organizes and provides elementary quantification for physical unit operations, but omits chemical reactors and, more to our subject, an overall view of a chemical process.

With respect to Process Systems Engineering the main developments during the Waiting Period are the following: (a) *Process analysis and design:* rudimentary material and energy balances are introduced and semi-empirical correlations form the basis for the design of unit operations. A gradually increasing quantitative sophistication in estimating reaction rates, transport rates, and thermophysical properties takes hold. (b) *Process Development:* The objective, as exemplified by the Dupont team (Chilton, Colburn) was fast and accurate process design and scale-up. Significant expansion takes place in the number, scope and complexity of new processes; high-

pressure, high-temperature, catalytic, multi-product processes. (c) The complexity of the new processes aggravates the trade-off problems that process designers must solve among cost, operational integrity, environmental emissions, and safety. The need for quantitative modeling of the processes becomes more acute.

In parallel and outside the disciplinary domain of chemical engineering, a series of important developments are taking place and will have a profound effect on the explosive growth of PSE from the late 1950s to today. These developments are: (i) Servo-regulator theory is introduced and addresses the instrumentation needs for monitoring and controlling chemical process operations, which become progressively more complex. This is the period of Bode, Nyquist, Ziegler, Nichols. Books on applied mathematics for engineers started appearing and creating an atmosphere conducive to the development and solution of better models for the description of physical and chemical operations: Sokolnikoff and Sokolnikoff, 1934; Reddick and Miller, 1938; Von Karma and Biot, 1940. Furthermore, the advent of still primitive computational machines is feeding developments on numerical methods for the solution of differential and algebraic equations, along with the introduction of academic courses to prepare the students. The works of Aitken (the δ^2-acceleration method; 1925), Steffensen (1933), and Wegstein (1958) will provide an essential tool for later steady state simulators. The foundations of Mathematical Programming are laid: Linear (Kantorovich,1939; Dantzig, 1947); Nonlinear (Karush, 1939; Kuhn, Tucker, 1951); and Integer (Gomory, 1958); all of which will play a profound role in the next phase of PSE developments.

1960s to Present: The Period of Explosive Expansion and Growth

By the early 1960s three factors are aligned to cause the onset of an explosive growth period for PSE: chemical industry had been growing rapidly, worldwide, for more than 10 years, exerting significant pressure for less costly and processes; a science-based description of the basic phenomena in unit operations is vigorously pursued, producing more reliable quantitative descriptions; the "computer" is entering industrial life in a very rapid determined way and affects all aspects of process systems engineering. The large jumps in energy and petrochemical raw materials, 10 years later, would further accentuate the need for better processes. From the 1960s to today we have witnessed a remarkable growth in the quality and number of PSE activities and practitioners, worldwide. In this article I will briefly overview the most influential developments.

Steady-State Simulators: By the early 60s chemical engineers can put together mathematical descriptions of reactions, their rates, thermophysical properties of pure materials and mixtures, equilibrium and rate processes, and do all of these as functions of equipment design parameters and operating conditions. Numerical methods are engaged to define procedures for the solution of these large sets of equations and computers are called to execute these procedures. The "Sequential Modular Steady State Simulators" come to life and spread very quickly; initially through in-house developments and subsequently supplied by external service companies. The impact of these simulators for the analysis, design and subsequent optimization and control of chemical processes cannot be overestimated. They are credited with 3-fold reductions in capital cost per unit of all commodity chemicals, over a period of 20-30 years.

Process Synthesis: As Dale Rudd observed, "Since World War II, engineering education has moved strongly toward analysis, with courses dealing with individual process operations and phenomena … greatly strengthening engineering education by

showing how things are and how they work. Unfortunately, there was not a parallel development in educational preparation dealing with process synthesis". Starting in late 60s, systematic Process Synthesis ideas allow process engineers to start with given chemistries and end up with quite inventive process flowsheets. Finally, after nearly 100 years the core activity of PSE, i.e. process development can be put into a rational and systematic framework. The impact of this development cannot be underestimated. Now we have a systematic way of "inventing" new processing systems, not simply analyzing existing ones. Eastman Chemical's new methyl acetate process is one of the most celebrated examples. It is a remarkable feat and its success depended on the "ingenious" integration of many elementary chemical engineering tasks.

Process Control: With increasing process complexity the limits of the servo-regulator theory became acutely apparent: strong interactions among control loops; convoluted logic and systems for controlling constraints; ineffective process optimization. In response, process control becomes more "process"-centric and with the following three important developments its practice was revolutionized: digital computers; model-predictive control; and plant-wide focus. Page Buckley's book in 1964 introduced the essential notions for a systematic plant-wide control approach, and was followed by extensive work from the 1970s on. The famous DMC controllers from Shell Development in the1970s took the industry by storm. They truly revolutionized how process control was going to be practiced from that point on: one consistent framework for set-point regulation; satisfaction of operating constraints; steady-state and more recently dynamic optimization of operations in conjunction with the planning and scheduling of operations. The economic impact of MPC was staggering: savings of more than 0.5 billion dollars from its application on 13 Dow Chemical ethylene plants; more than 4,500 deployed systems in major chemical plants over a 15-20 year period.

Optimization: The foundations of optimization theory were in good state at the beginning of the 1960s. During the subsequent 20 years, there was an explosion in the field and significant progress took place in developing effective algorithms for linear, nonlinear, and combinatorial optimization. Chemical engineers in increasing numbers become involved as practitioners and developers of new methodologies. As a result, the practice of optimization spreads very rapidly and is being applied in a large variety of PSE activities: Process Design; Process Synthesis; Process Control; Planning and Scheduling of Process Operations; Supply-chain Management; Development of Models from Data; Product Design; Metabolic Engineering; Oil and Gas Production.

Risk Management and Process Safety: In the 1970s and 1980s a series of environmental devastations (toxic leaks from Love Canal and the meltdown of reactor core in Chernobyl) and high-profile industrial accidents (Bhopal, 3-Mile Island, refineries) sharpened the sensitivities to risk management and process safety. PSE played a central role in addressing these issues. First, it pushed a new mindset in dealing with Process Integrity: from "React and Fix" to "Predict/Detect and Prevent". Second, it fostered the development of a comprehensive Process Safety Management approach with Inherently Safer Process Designs; Advanced Process Safety methodologies and tools; integration of Safety Instrumentation Systems with process control systems; and advanced methods for the design of LOPAs, on-line intelligent diagnostic systems, and others.

Current State and Future Prospects

The Systems-view has been incorporated into an ever increasing number of problems that chemical engineers have been addressing over the last 20 years: product and process design; multi-process and multi-plant integration; supply-chain management; development of pharmaceutical products and biomedical devices; technology and policy for the energy and the environment. Modeling, the core of all PSE activities has advanced to levels of very sophistication: atomic and molecular models; hybrid models with for continuous and discreet real-valued and integer variables. Simulation has followed an equation-oriented path opening up ever increasing possibilities: distributed processes in time and space; processing at the micro-scale; interfacing with statistical mechanics and computational fluid dynamics. Optimization theories and algorithms move in tandem with the increasing sophistication of the models and the size of the applications, while expanding continuously the scope of problems it addresses. Risk management and process safety continue to attract strong attention and support deployment of new technologies and methods. Monitoring, diagnosis, control, planning and scheduling of process operations are being integrated under common theoretical frameworks and computer-based real-time deployments. The results of these efforts have had a deep impact in industrial life and continue to feed new developments.

While efforts towards new improved methodologies and computer-aided tools will continue, the future of PSE activities rests with its ability to continuously renew itself by addressing new classes of problems that the economy faces. The author believes that the following features will characterize PSE activities over the next 10 years:

(a) *Continuous expansion into new areas*: Biology has been one and PSE practitioners have been heavily involved in "Systems Biology" and "Systematized Synthetic Biology", while we have seen the first PSE efforts in the area of materials. Both areas offer a rich variety of important problems to solve.

(b) *Continuous expansion of the boundary that defines the "System"*: Multi-plant sites define the new boundaries; production of different forms of energy, their transportation, and places of energy consumption determine ever expanding boundaries; technology and policy for energy and the environment brings into the system a series of "non-processing" entities. The process of development of pharmaceutical products and the national agenda of Smart Manufacturing require the definition of system whose boundaries include a multitude of other activities.

(c) *Modeling for Discovery and Technological Innovation*: Beyond simulation and design, modeling is at the heart of systematizing experimentation for scientific discoveries and accessing the "risk" in technological innovation.

(d) *From macro- to micro- to nano-scale PSE:* Constructing and operating systems with processing steps at the nano-scale presents a series of problems that present-day PSE has not addressed. Design and operational precision at the nano-scale, the role of shape and geometry, and self-regulation at all levels of the system, are some of the problems that would feed future activities in PSE.

(e) *From "Complicated" to "Complex" Systems Engineering:* The need to understand and model the behavior of self-assembled systems, biological processes based on

few molecules, or the ever expanding scope of classical systems with "humans" as some of their components, pushes PSE to go beyond the scope of well-defined but large and complicated systems, and develop new approaches and methodologies for the so-called "complex" systems.

Acknowledgement

Due to restrictions in the length of the manuscript, this paper is an abridged version of a longer paper currently under preparation with a full list of references. It is mportant though to acknowledge the input that many people have had, either in person or through their writings and presentations that generously shared with me. They are: Paul Barton; Larry Biegler; Jim Davis; Tom Edgar; Ignacio Grossmann; Christos Maravelias; Manfred Morari; Costas Pantelides; John Perkins; Jim Rawlings; Scip Scriven; Jeff Siirola; Jim Trainham; Venkat Venkatasubramanian.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

New Opportunities for Process Systems Engineering in Industrial Biotechnology

John M. Woodley

Center for BioProcess Engineering, Department of Chemical and Biochemical Engineering, Technical University of Denmark, 2800 Lyngby, Denmark

Abstract

As industrial biotechnology develops into a mature sector in the coming years, new demands will be placed on evaluating economics and design to retrofit existing plant as well as design new plant. Process systems engineering will also have a major role to play in design of suitable control systems on a rigorous basis. Some new problems will also be forthcoming due to the particular characteristics of biological processing and this too will provide exciting opportunities for process systems engineering.

Keywords: bioprocesses, biocatalysis, fermentation, modeling, simulation.

1. Introduction

Process systems engineering offers many tools for the chemical engineer. Modeling, simulation and process evaluation tools are routinely applied in the bulk chemicals and fuels sector where small process improvements yield significant economic returns. In recent years bioprocesses have become more common. In particular they have found application in the high value products such as pharmaceuticals (and their intermediates). The emphasis in these cases is on rapid process implementation. However in recent years bioprocesses have been applied to bigger volume products such as fine chemicals, bulk chemicals and fuels. In these cases process improvement is the emphasis since this will give significant returns. A range of questions need to be addressed. For example: When should a bioprocess, rather than a chemical process be implemented? If a bioprocess is to be implemented can existing infrastructure be used? How can process plant be adopted for different biomass in different global regions? What is the optimum biorefinery? What options exist for process integration? Quantitative tools are required to solve such problems and process systems engineering has many of these tools currently. In addition, in the case of bioprocesses an extra option available to the engineer is the improvement of the catalyst itself. Hence we are now at the point that process engineering tools need to be applied to the complete set of bioprocesses, including pharmaceuticals, fine chemicals, bulk chemicals and fuels. The main synthetic operations include fermentation, microbial catalysis and enzyme catalysis. Downstream options are dependent on the nature of the product (macromolecular or small molecule). Small molecules are frequently processed in a similar way to other chemical products, although dilute aqueous solutions bring specific problems which need to be addressed. Macromolecules require more specialist operations such as filtration or chromatography. In all cases the molecules are frequently sensitive to extremes of pH and temperature placing specific restrictions and constraints on processing methods. Catalyst recovery (sometimes for recycle) also necessitates filtration and centrifugation. In this paper the specific role of process systems engineering in the development, design and improvement of industrial bioprocesses will be discussed.

2. Scope

Biotechnology is an enormous sector of industry from high value, low volume products (such as pharmaceuticals) to low value, large volume products (such as bio-fuels). To date the majority of implemented processes have focused on the former group. The emphasis here has been on implementing processes effectively to meet the tough regulatory demands placed on such products. Process implementation, rather than optimization, has been the necessary focus of process engineering (e.g. Pollard and Woodley, 2007; Woodley, 2008). In this paper the future processes for biotechnology which fall into the latter category will be discussed. These are the new sectors of industrial (or also called 'white') biotechnology where new opportunities exist for alternative feed-stocks based on renewable resources such as biomass and clean processes with reduced solvent inventories, renewable catalysts and mild conditions for reaction and separation (e.g. Dale, 2003). However this new 'bio-economy' will require the development of a suitable infrastructure and, like the oil-based counterpart will demand very high yield processes meaning that process engineering for the future implementation and development of these process will have an increasingly important role. Given the maturity of the field of process systems engineering it is clear that many new opportunities will be forthcoming.

3. Industrial biotechnology processes

3.1. Fermentation processes

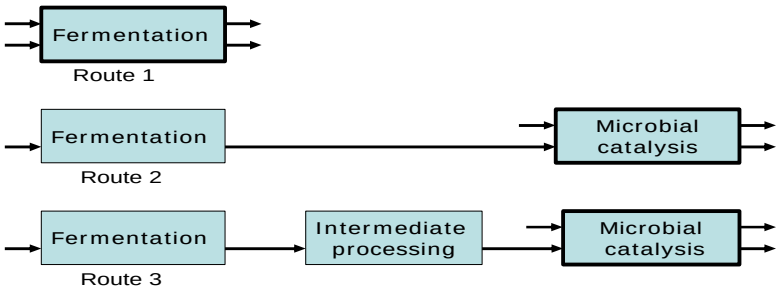
For a significant number of chemicals, the use of fermentation has become a standard alternative to fossil-based feed-stocks and technology. Nevertheless the possibility of growing microbial cells on a variety of sugars (derived from renewable biomass) has re-invigorated interest in this area. The consequence is that fermentation at a large-scale will become more common in the future of the chemical industry. Many fermentation processes can take place in similar process plant which is a significant advantage. The plant is relatively simple and the challenges lie in adequate mixing (sometimes with materials having complex rheology), suitable oxygen input (for aerobic processes) and process control. Downstream, the separation process depends on the product but will nearly always need to avoid high temperatures and extremes of pH. The solvent is water, meaning that the dilute product stream combined with the presence of many other products presents a significant process engineering challenge. Both large molecular weight and low molecular weight products can be made by fermentation. Most industrial biotechnology processes (excluding pharmaceutical processes) focus of low molecular weight products which can subsequently be used as platform chemicals or fuels or enzymes (for application in a range of industries, including detergents, textiles and food ingredients).

3.2. Microbial catalysis

In fermentation processes, by definition, the catalyst is growing during the process. This means that some of the reactant (or substrate) will inevitably be diverted from the product towards the catalyst, lowering the yield. An alternative (for no growth associated products) is to grow the cells first and subsequently carry out the reaction to increase the yield. This also enables the possibility of growth and reaction on different substrates (reactants) or under different conditions in each stage (such as temperature). Likewise the optimal cell concentration for conversion can be selected (Woodley, 2006) and suitable media for effective product recovery chosen. For processes requiring oxygen it can be highly important to select the optimal cell concentration in order to

avoid mass transfer limitations. Several tools are now available for evaluating the oxygen supply issues in such reactions (Baldwin et al, 2006; Law et al, 2006). The three potential routes are shown in Figure 1.

Figure 1: Alternative process scenarios for use of microbial cells for non growth associated biocatalysis. (Route 1 - Combined fermentation and microbial catalysis; Route 2 – Fermentation separated from microbial catalysis; Route 3 – Fermentation separated from microbial catalysis with intermediate processing to change catalyst concentration).



3.3. Immobilized enzyme processes

The presence of so many products is a consequence of the complexity of cells where many enzymes catalyze reactions giving a spectrum of products as well as decreasing the yield of the desired product on the reactant. An alternative, for short pathways, is to isolate the enzymes and then immobilize them on a solid support or behind a membrane or via aggregation, such that they are large enough and have the right properties to be recycled (like a heterogeneous catalyst). In this way a yield of product on the catalyst of around 5 – 10 tonnes / Kg catalyst can be achieved, which enables economic industrial operation. Such an approach has been widely used to assist in the synthesis of high value compounds such as pharmaceuticals and a limited number of well known lower value products such as high fructose corn syrup (HFCS). Many of these processes have been modeled (e.g. Chen *et al*, 2007).

4. The role of process systems engineering

4.1. Evaluation of process options

For some higher value products a bioprocess may in some cases be the only route to a given product (to ensure correct folding of a therapeutic protein for example or the synthesis of an optically pure pharmaceutical intermediate). However, the more usual situation is that there are other competing routes to the same product. Therefore, for now, biotechnology is just one of a number of options for the production of chemicals and fuels. The economic drivers for implementation depend on existing infrastructure, feedstock costs, feedstock availability as well as the efficiency of the relevant (bio)catalyst and (bio)process technology. This presents a fascinating set of alternative

routes and technologies from a given feedstock and/or to a given product(s). Process systems engineering has a particular role to enable such evaluations on a quantitative basis. Process systems engineering also bring the advantages of rapid computational methods such simulations enable alternatives to be quickly evaluated. The answer in a specific case to the problem formulated here will in addition depend on regional factors. Feedstock availability and cost is highly dependent on geographical location. A parallel set of evaluations concerning the need to retrofit existing plant, or build new plant, is also required. Objective functions to be optimized will need to be based not exclusively on economics but increasingly based also on sustainability metrics (e.g. Henderson *et al*, 2008) and integrated with life cycle analysis (to include evaluation of feed-stocks and products as well as processes).

4.2. Evaluation of platform chemicals

While the increasing cost of oil is driving particular interest in the production of new fuels from biomass there is little doubt that today of equal importance is the production of chemicals from biomass. Indeed for the supply of fuels in the future there are many potential sources aside from biomass. In a world with limited (or very expensive) oil it is less clear where the chemicals of the future will originate. There is currently an existing infrastructure based on the use of the 7 established platform chemicals (toluene; benzene; xylene; 1,3-butadiene; propylene; ethene; methane). In the short term one could consider if we can use the same infrastructure and just create the 7 chemicals from alternative sources. However in the longer term it will be necessary to devise new processes based on a different set of platform chemicals. One group will be based around glucose (the hydrolytic product of starch and cellulose and therefore readily available from biomass). In a biorefinery it will be necessary to develop a structure which can manage a range of feed-stocks, a range of technologies and a range of products. This presents a considerable challenge for design and optimization as well as process integration. An interesting example which illustrates the complexity and the challenge that lies ahead is the use of glucose or fructose to produce 5-hydroxymethylfurfural (HMF) or 2,5-furandicarboxylic acid (FDA). Greatest value is obtained by going the whole way from glucose to FDA. However even in this small reaction pathway there are many alternative technologies. Some can be integrated together, some give the required yield and selectivity, some are difficult to implement and others are untested at scale. This illustrates very well the challenge that design engineers face.

4.3. Process integration

The solvent for all bioprocesses, with a few exceptions, is water. Consequently the downstream process is frequently difficult and this is emphasized by the need to carry out separations at moderate temperatures. Given the dilute nature of the streams it is frequently the case that the majority of the costs are therefore in the downstream process. In the case of transport fuels removal of water becomes an essential requirement to reduce costs and avoid transporting large amounts of water. For example, in the case of ethanol which forms an azeotrope, this can be a significant cost. In other cases the product may be integrated within a biorefinery although at some point water will need to be removed. Consequently the integration of water use and reuse via recycle is an essential part of the design of industrial bioprocess facilities. Existing tools such as pinch technology will have an important role. The issue of water use in a biorefinery is in many ways analogous to the issue of heat use in a conventional refinery.

4.4. Biorefinery design

Two major types of biorefinery have been identified for the future, based on lignocellulose biomass utilization to provide a arrange of sugars (for subsequent (bio)catalysis or fermentation) and oil-based material (from biomass). In each case the current research emphasis on biorefineries is to ensure all the fractions of a particular biomass in a given situation are fully exploited. Likewise the development of downstream products is now being explored. For example glycerol (as a byproduct of biodiesel production) can be used as a platform chemical (for example via fermentation to produce 1,3-propanediol). Another interesting example concerns the production of bio-ethanol. This is widely developed as a fuel although there is considerable economic incentive for developing a range of other products (e.g. acetic acid) from ethanol, in other words using it as a platform chemical (Rass-Hansen *et al*, 2003). It is clear that evaluating integration of product and feed streams in this way, as well as use of products as fuels or platform chemicals will be an important role for process systems engineering. There is little doubt that the increasing range of technologies and opportunities as a result of enzyme–based or fermentation-based catalysis will give complex integration problems, which may require new tools (as discussed also in 4.2).

4.5. Biocatalyst design

A particular feature of bioprocesses is the use of biocatalysts, which may exist in several forms as indicated earlier and where options exist for modification. At the simplest level as a protein (isolated enzyme), the options for swapping amino acids via protein engineering exist. New enzymes which have been modified may display new tolerance to reactor conditions such as temperature or pH and may also have improved selectivity of reactivity (activity) on a given (non-natural) substrate or reactant. Order-of-magnitude improvements have been found in a number of cases although understanding the most effective method of making changes to the enzyme depends on past precedent and, to some extent, structural knowledge (e.g. Hibbert *et al*, 2005). In the case of cellular catalysts, individual enzymes can be over-expressed (increasing reaction rate of a given cell) and the regulatory control scheme fixed to direct the carbon to give improved rates and yields (via metabolic engineering). Some start has also been made to the development of pathways where enzymes coming from a variety of sources are cloned into single host to make a new pathway via a combination of genetic engineering and *de-novo* pathway engineering (e.g. Chen *et al*, 2006). In all these areas it is clear that those involved in process systems engineering need to inform the biological engineers about what is required in a given case and set suitable targets. Philosophically it is interesting to note that process implementation may come via process improvements or alternatively via catalyst improvements. In many cases both will be required. Understanding the necessary balance between these areas, as well as their integration with each other will be important for the future development of the field.

5. Future outlook

The development of new bioprocesses as a complement to existing chemical and fuel production is an exciting endeavor that will occupy many process engineers in the future. There will be a particular role for process systems engineers in this developing sector with the advantages of quantitative decision-making tools and rapid simulation that this brings. In the future suitable models will inform developments at the infrastructural level (evaluation of biorefineries, feed-stocks and integration), the process level (evaluation of alternative technologies and process integration) and the

catalyst level (alternatives for protein and metabolic engineering). The further development of process systems engineering tools (including property prediction packages and the development of a database for bio-based molecules) will be required. In addition an increasing dialogue with biochemical engineers and biologists will be necessary to enable the vision of industrial biotechnology to be fully exploited.

References

C.V.F. Baldwin J.M. Woodley, 2006, On oxygen limitation in a whole-cell biocatalytic Baeyer-Villiger oxidation process, Biotech. Bioeng., 95, 362-369.

B.H. Chen, A. Sayar, U. Kaulmann, P.A. Dalby, J.M. Ward, J.M. Woodley. 2006, Reaction modelling and simulation to assess the integrated use of transketolase and ? -transaminase for the synthesis of an aminotriol, Biocat. Biotrans., 24, 449-457.

B.H. Chen, F. Baganz, J.M. Woodley, 2007,. Modelling and optimisation of a transketolase mediated carbon-carbon bond formation reaction, Chem. Eng. Sci., 62, 3178-3184.

B.E. Dale, 2003, 'Greening' the chemical industry: research and development priorities for biobased industrial products, J. Chem. Tech. Biotechnol., 78, 1093-1103.

R.K. Henderson, C. Jimenez.Gonzalez, C. Preston, D.J.C. Constable, J.M. Woodley, 2008, Comparison of biocatalytic and chemical synthesis: EHS and LCA comparison for 7-ACA synthesis, Ind. Biotechnol., 4, 180-192.

E.G. Hibbert, F. Baganz, H.C. Hailes, J.M. Ward, G.J. Lye, J.M. Woodley, P.A. Dalby, 2005, Directed evolution of biocatalytic processes, Biomolecular Engng., 22, 11-19.

H.E.M. Law, C.V.F. Baldwin, B.H. Chen J..M.Woodley, 2006, Process limitations in a whole-cell catalysed oxidation: sensitivity analysis, Chem. Eng. Sci., 61, 6646-6652.

D.J. Pollard, J.M. Woodley, 2007, Biocatalysis for pharmaceutical intermediates: the future is now, Trends Biotechnol., 25, 66-73.

J. Rass-Hansen, H. Falsig, B. Jørgensen, C.H. Christensen, 2007, Bioethanol:fuel or feedstock, J. Chem. Tech. Biotechnol., 82, 329-333.

J.M. Woodley, 2006, Microbial biocatalytic processes and their development, Adv. Appl. Microbiol., 60, 1-15.

J.M. Woodley, 2008, New opportunities for biocatalysis: making pharmaceutical processes greener, Trends Biotechnol., 26, 321-327

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

The Significance of Experiments on PSE

Günter Wozny, Harvey Arellano-Garcia

Chair of Process Dynamics and Operation, Sekr, KWT-9, Berlin Institute of Technology, Strasse des 17 Juni 135, D-10623, Berlin, Germany

Abstract

Chemical product design, manufacturing and its process imply a spectrum of chemical and physical phenomena over different scales in time and space. An increased use of process simulation in different phases of process life cycle and on different scale makes the information management related to the modeling procedure and practice more relevant. With more complex models also experiments are of increasingly importance for model validation and an improved process understanding. However, the experimental data and meta-data has to be linked simultaneously to the simulation results. Moreover, the requirements for more efficient models, customization, reuse and documentation increase too. The benefit of systematic approaches such as experimental design methods are not sufficiently exploited and should be introduced in the normal process development procedure. Concurrent process development leads to new demands for experimental work beside the process simulation, synthesis and process control. In education new modern and effective standardized and modularized experimental training stations are needed for better understanding and increased students interest. In this work, different aspects of the significance of experiments related to PSE will be presented. For this purpose, different examples are selected and discussed in detail.

Keywords: distillation, ethylene separation, energy saving, process simulation.

1. Introduction

The chemical engineer's work involves a wide range of scales: from molecule in nanometers to plants and enterprises in km and from molecular rotation in picoseconds to life cycle analyses of plants in years. In the last decade several aspects have been investigated in detail and described in literature. Chemical engineering is changing from process design to product design and the integration of both. In the late 19th century till the 1960s the "Unit Operation concepts" including dimensionless group correlation was leading in research and education. Starting form the early 1960s transport phenomena (Bird et al, 1960) became a more and more important branch. Transport phenomena in fluid flows and those related to turbulence, droplets, films, bubbles got more attractive because these methods have helped engineers to get a better insight into the processes. In comparison to Unit Operation scale, transport phenomena were based on a smaller scale. Today, simulation tools like CFD support these methods and a wide range of successful applications in industry and academia have been reported. In the opposite scale direction, aspects of plant wide simulation, plant wide control were considered and investigated starting from the 1980s till today. Nowadays, simulation tools in industrial applications such as Aspen, ChemStation, Simulation Science etc. support engineering work. In fact, PSE methods are currently used in the daily engineers work in process design, optimization and control. Is there anything missing or what can be improved? In

this work, the main focus is not on detailed improvements of any single PSE method. It rather discusses the significance of experiments to PSE. Is there any need to consider these aspects in PSE or can PSE be seen completely separate from experimental research? For an improved understanding of this question, it is helpful, as usual, to take a short look in the literature and discuss definitions and the basics of PSE. For instance, Grossmann and Westerberg (2000) wrote "Process Systems Engineering is concerned with the improvement of decision – making processes for the creation and operation of chemical supply chain. It deals with the discovery, design, manufacture, and distribution of chemical products in the context of many conflicting goals" They continued: "The PSE area can also have validation through experiments for physically existing systems, but, in most cases, PSE has only "virtual" systems in the form of models (e.g. in the conceptual phase of a design). One can test the models, ideally by proving and disproving theorems about them, but, more often, one can only prove mathematically some useful properties or empirically demonstrate that the proposed model works for a set of example problems. So, it seems that experiments are not so important for PSE and there would be only a weak connection. The topic of education is reflected by J. Perkins (2000), he postulated in his article: "Educators have been concerned to develop curricula and courses designed to help students to adopt a "systems approach" to engineering problem solving". He continued: "Design is not only something picked up by experience but is a formal procedure with its own rules which could not only learnt by students but could be taught in a rigorous manner. The Chemical engineer needs to know about a whole range of techniques outside the narrow ever more scientific approach of chemical engineering science." In the attached tables about curricula, the reader will find in the first and also in the second year only one lecture related to experiments from 15 topics like simulation, design. Is that enough? Also, Ziemlewki (2009) summarized aspects in the direction of "Designing the New Global Chemical Engineer". Important topics in this paper are applied mathematics, software implementation, modeling aspects. It is recommended to teach Soft skills and train students in teamwork, leadership aspects, communication, cultural diversity understanding. Nothing about experimental skills is recommended there. Which are the main reasons? Based on our personal experience it is the opposite. Experimental knowledge should be expanded. Students should do different work in laboratory during their studies. Group work should be expanded. Some of the reasons, why experimental work is reduced, are well known. Experiments are expensive and time consuming. How can PSE improve this situation?

Research is, simplified expressed, formulating a hypothesis, observing nature or the real behavior in the plant, in part of the plants, making conclusions and integrating them in a more global context. If we would know all, then no experiments would be necessary. I noticed, once more this is a personal view and experience, an increased uncritically believing what the computer creates by our young students. They like to do computer work by trial and error and stop working after any undesired result happens. They have apparently success very quickly, but often without deep understanding. Because simulation does not support the creation of a mental process model as it does the experiment, there is no critical reflection. How can we change this situation? As we know, methods of Process development are beside the typical PSE methods like simulation and optimization, searching in literature and patents, doing labor experiments (seldom direct usable for equipment design), investigating behavior in Mini-Plants (reducing risk, determination of technical parameters), investigations in pilot-plants (all Unit operations, recycle streams included, cost analysis). Already

Krekel and Siekmann (1985) described important aspects of the role of experiments in process development. Design data for the improvement of existing plants or for the development of new chemical processes are often determined by experiments in laboratories and pilot plants. These frequently expensive test-runs should lower the scale up risk and should ensure its long term functioning and its economy. In this stage, PSE with selected methods like optimal experimental design for an exact planning of experiments can minimize the cost of development. Process simulation, process optimization, control design methods should support the experimental work in parallel. The main aims of experiments are often: determination of reaction data, delivering of first balancing data, production of small charges of products for selected first customers, generation of scale-up data, providing economic data, ensuring long term functioning.

During this stage of process development PSE methods should be used to improve and speed up the design process. An efficient process development can be done by combining Process Simulation and Mini-Plant technique. Moreover, the current economic situation, the international competition and the product life cycles, which are becoming shorter, demand a significant reduction of time from the conception of a product to its market introduction. A significant contribution to rapid and cost effective process design is the combination of PSE methods and experimental methods in different scales. Despite this, it can be noticed very often that PSE methods are used partially or totally independent from experimental work. Often different persons and even different departments are involved solving the postulated problems either in the PSE direction or doing experimental work. A more promising way is to use PSE methods like simulation, optimization and considering control aspects parallel to the experimental investigation from the very beginning. The following selected applications show how some aspects of this approach influences the results.

2. Applications

Changes in process conditions can influence behavior by changing the value of process parameters or by changing the constraints of the model considered. Thus, the dynamic behavior of a process unit will proceed through a series of states, with each state being associated with a different mechanism or model. Although the models are a formalism, the information of such models may offer insight into the considered processes underlying the phenomena being examined, and may be useful in suggestions for new experiments. A common *modus operandi* is to examine the prediction of a number of models, and make out which models are particular good at describing the experimental data, by means of fitting procedures or statistical methods (Faber et al., 2007). Nevertheless, the quality of model predictions for the description of the underlying phenomena depends on the model structure and on the values of the contained parameters which have to be fitted for model validation using experimental data. These measurements are not only used to determine the state variables but also to improve the quality of the process models. Despite increasingly rigorous process modeling, most of the process models use process parameters which have to be determined based on measurement data. As most of these model parameters depend on the actual state of the process and the plant configuration, and therefore, are subject to changes during process operation, it is in many cases reasonable to adjust these parameters using data taken directly from the plant. However, after the estimation, it is important to asses the

reliability and adequacy of the model and of the new parameter estimates again following an iterative procedure. Many relevant processes in bio- and chemical engineering are still not well understood, because the underlying physics is complex. A representative class of processes involved in this contemplation corresponds specifically to catalytic reactions. In these processes several steps could influence the reaction rate (see Fig. 1, for the reaction A→B). First, component **A** has to diffuse from the bulk flow towards the boundary layer (a). Following this, it has to enter into the catalyst pores (b). Then it can be adsorbed on the catalytic surface (c), where the reaction takes place (d).

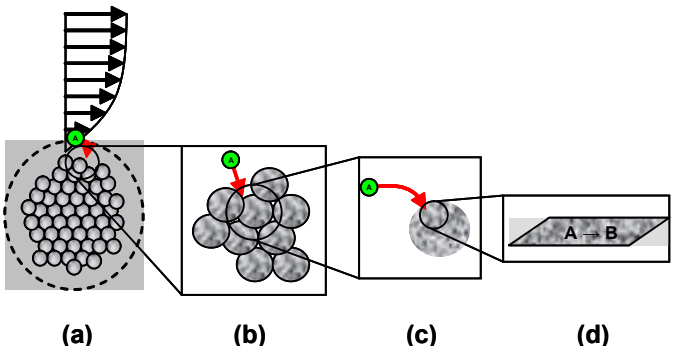


Figure 1 (a) External diffusion; (b) Internal diffusion; (c) Adsorption; (d) Reaction.

These steps are followed by desorption and transport of the product B to the bulk flow. At the steady state operation is though common to work with a rate limiting step (RLS) assuming that only one step dominates the reaction resistance. In most of the cases it is the reaction step (d). However, each step can be described with a set of model equations. The question is whether it is possible to identify process states, in which other steps dominate the process. This may occur, for example, due to catalyst decay. With this knowledge suitable counteractive measures can be initiated. A so-called Automated Mechanism Recognition (AMR) approach has been developed, which is based on the recognition of current dominant mechanisms (Arellano-Garcia et al., 2008).

Another application of interest on this issue represents the **fluidized bed reactor** for the OCM (oxidative coupling of methane) process. This suggestion is based on the fact that gas-phase reactions occurring in the bubbles are significantly slower than the catalytic reactions, which are very sensitive to the presence of oxygen. Oxygen is necessary for the reaction to occur, but unfortunately excess of oxygen or it's high availability to catalyst can cause combustion reactions to be dominating in comparison to selective – methane coupling reactions. Therefore so called "distributed feed of oxygen" is needed in order to control the combustion reactions and achieve higher selectivity.

This is possible with either **membrane reactors**, reactors with oxygen co-feed or with fluidized bed reactors for very fast catalytic reactions. The "co-feed" principle can only occur in fluidized bed reactors if reaction is significantly faster than mass-transport between bubbles and emulsion, and therefore controls the overall reaction rate. In this case, working conditions with excess of methane will ensure low oxygen concentration in the emulsion phase, and therefore increase the rate of selective reactions compared to combustion reactions. For getting a qualitative overview of the problem, model-based and experimental analysis of the reactor are needed.

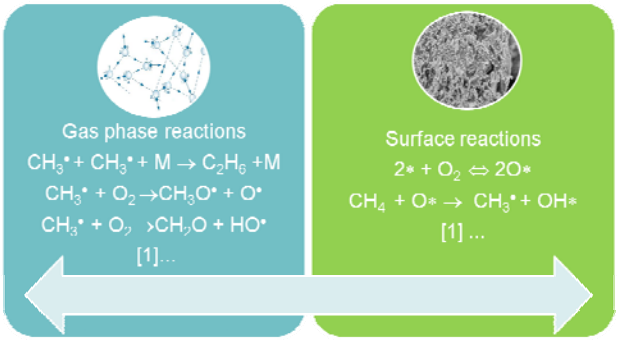


Figure 2 Gas phase and surface reactions in the OCM Process.

Another application in this work is given by a dynamic model for a **catalytic fixed bed reactor** (FBR) based on partial differential equations (PDE), which is discussed and used for the identification of reaction mechanisms taking place during the oxidation of sulphur dioxide over a vanadium pentoxid catalyst. The measured data is collected from a pilot plant, which uses commercial sized catalyst particles. In order to reduce the experimental effort a developed framework based on the methods of nonlinear optimal experimental design is applied using a steady state FBR model (Schöneberger et al., 2008, 2009). The systematic procedure is improved using a dynamic reactor model. This makes the time dependent measurement data valuable for the identification procedure.

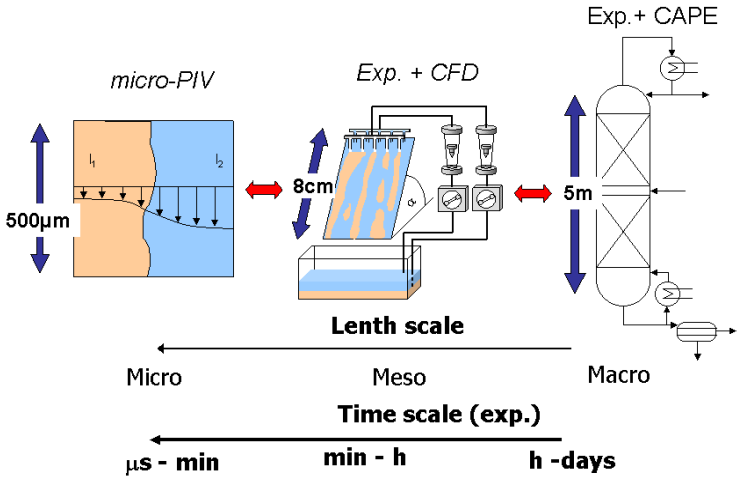


Figure 3 Combination of PSE method and experiments to understand flow behavior and separation efficiency of three phase distillation column.

In **three phase distillation** a counter current vapor flow influences the liquid flow additionally. Since the interfacial area between all phases is significant for mass and heat transfer, the determination of the interfacial area is necessary (Hoffmann et al., 2005). For this purpose, experimental setup for optical measurement methods, like particle tracking velocimetry (PTV), light induced fluorescence (LIF) and µ-PIV (Micro Particle induced velocimetry) are employed. To use the µ-PIV for investigating the flow velocity in packing expenditure is necessary. Normally a transparent plate is used and

the measurements are done from the back side. For this application the measurement technique has to be expanded to non transparent materials. With this modification usage of the experimental data for model validation is possible. Comparison of the experimental results with CFD simulation will be shown. The investigation of three phase distillations shows that the analysis of different scales from m up to μm and from minutes and hours up to μs is necessary for an improved understanding and description of the physical behavior of chemical processes (see Fig. 3). Using only one single PSE method (e.g. process simulation or parameter identification or CFD simulation) is not sufficient. On the track of understanding three phases (two liquid, one vapor phase) in one packed tower a detailed knowledge about the complex fluid dynamics on the packing surface is required. New measurements methods and the parallel use of different PSE methods as experimental design, parameter identification, CFD simulation would increase the efficiency and the speed of development.

3. Conclusions

For process design and development PSE methods and experimental investigations should be combined from the early state. By this means, risks of scale-up such as temperature distribution in the reactor, cooking of the catalyst, mechanical stability, different flow patterns leads to different yields and selectivity, foaming in columns, reboiler and so on can be reduced. Some of these problems can be solved by simulation. Parallel good planned experiments are necessary.

Acknowledgment
The authors acknowledge support from the Cluster of Excellence "Unifying Concepts in Catalysis" coordinated by the Berlin Institute of Technology and funded by the German Research Foundation – Deutsche Forschungsgemeinschaft.

References

2. Grossmann, I. E.; Westerberg, A.W.: Research Challenges in Process Systems Engineering, AIChE Journal September 2000, Vol. 46, No. 9, 1700.
3. Arellano-Garcia, H., Schöneberger, J., Wozny, G.: An Optimization Based Approach To Model Identification, Proceedings FOCAPO 2008, Boston, June 29 - July 2.
4. Faber, R. Arellano-Garcia, H.; Li, P.; Wozny, G.: An Optimization framework for parameter estimation of large scale systems, Chemical Engineering and Processing 46, 1085-1095, 2007.
5. Perkins, J.: Education in Process systems Engineering – Past, Present and Future; Computers and Chemical Engineering, 24(2000), 1367.
6. Schöneberger, J.; Arellano-Garcia, H.; Thielert, H.; Wozny, G.: Systematic Validation of Kinetic Models of Heterogeneous Catalyzed Gas Phase Reaction,Proceedings 4th International Conference on High Performance Scientific Computing,Modeling, Simulation and Optimization of Complex Processes, March 2-6, 2009, Hanoi.
7. Ziemlewki J.: Designing the New Global Chemical Engineer, CEP February 2009, 6-9.
8. Schöneberger, J.C., Arellano-Garcia, H., Thielert, H., Körkel, S., and Wozny, G.: Model based Experimental analysis of a fixed bed reactor for catalytic SO2 oxidation. Ind. Eng. Chem. Res., 2008, in press.
9. Hoffmann, A.; Ausner, I.; Repke, J.-U.; Wozny, G.: Fluid dynamics in multiphase distillation processes in packed towers, Comp. Chem. Eng. 29(6), 2005, pp.1433-1437.
10. Paschke, S., Repke, J., Wozny, G.: A way to analyze liquid film flows on opaque plate materials. 4th International Berlin Workshop on Transport Phenomena with Moving Boundaries, 27-28 September, 2007, Germany.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Review of Sustainable Energy – Recent Development and Future Prospects of Dimethyl Ether (DME)

En Sup Yoon, Chonghun Han

School of Chemical and Biological Engineering, Seoul National University
Shillim-dong, Gwanak-gu, Seoul, Republic of Korea.

Abstract

DME can be synthesized from natural gas, coal, biomass, and/or coal seam, and is a sulfur-free, near-zero aromatics synthetic fuel which is considered as an excellent substitute for conventional diesel and liquefied petroleum gas. Currently, various production technologies are developed and many commercial projects are actively being progressed. This paper presents recent development and future prospect of Dimethyl Ether (DME) as an alternative energy and simple review for biomass energy.

Keywords: dimethyl ether, sustainable energy, alternative fuel

1. Introduction

Sustainable energy resources are fundamental to economic stability and development. The threat of climate change, the erosion of energy security, and the growing energy needs of the developing world, all pose major challenges to energy decision makers. Innovation of new energy technologies and a better use of existing technologies will be required to meet these challenges. Environmental pollution and the greenhouse effect, resulting from high dependency on petroleum based energy resources, cause special concern and threaten energy security. As a result, greater efforts and attention are given to the development of non-petroleum based energy resources.

Among the candidates for the non-petroleum based energy resources, solar, geothermal, wind, and tidal energies have received much attention as green energy sources, while biomass energy, hydrogen and fuel cells, and gas-to-liquid coal gasification have been the subject of intense study as synthetic sources. Recently, synthetic energy sources have been in the lead among alternatives to petroleum based energy for their environmental friendliness, high energy density, and potential for commercialization. This paper deals with one of the synthetic energy challenges, dimethyl ether (DME). DME can be synthesized from natural gas, coal, biomass, and/or coal seam, and is a sulfur-free, near-zero aromatics synthetic fuel which is considered an excellent substitute for conventional diesel and liquefied petroleum gas (LPG). In this paper, we introduce recent studies and review the technological and economic aspects of DME manufacture and its prospective use as an alternative fuel. Also, we will review the biomass energy that can be used as feedstock for various synthetic energies as well as DME in the presentation.

2. Properties of DME

The properties of DME suggest that it could be a substitute for petroleum based energy with much better environmental performance. DME is a sweet nontoxic chemical with an ether-like odor and is liquefied at 6 atm or -25□. The characteristics of DME are

similar to liquefied petroleum gas (LPG), which can be used in replacement of LPG for domestic heating, clean combustion automobiles and for exploitation of existing land-based and ocean-based LPG infrastructures with minor modifications.
With a cetane number of 55-60, DME is also considered a substitute for diesel fuel (cetane number 55). The cetane number is a measurement of the combustion quality of diesel fuel during compression ignition. It is a significant expression of diesel fuel quality among a number of other measurements that determine overall diesel fuel quality. Notably, when compared with diesel fuel in CIDI engine tests, the combustion of DME produced much less pollutants than diesel (e.g., hydrocarbons, carbon monoxide, nitrogen oxides, and particulates).

3. Production Technology

3.1. Chemistry

DME is produced in a minimum of two steps. First, hydrocarbons from natural gas, coal, or biomass are converted into synthesis gas (syngas), a combination of carbon monoxide and hydrogen. Second, the syngas is then converted into DME in two different ways. One way is a conventional two-step process which consists of methanol synthesis and dehydration; the other is a one-step process which directly produces DME from syngas. The reactions are shown below.

Table 1. DME synthesis chemistry

Partial oxidation reforming	$CH4+1/2O2 \rightarrow CO+2H2$
Steam reforming	$CH4+H2O \rightarrow CO+3H2$
Gas/water shift reation	$CO+H2O \leftrightarrow CO2+H2$
Methanol synthesis	$2CO+4H2 \rightarrow 2CH3OH$
	$CO2+3H2 \rightarrow CH3OH$
Methanol dehydration	$2CH3OH \rightarrow CH3OCH3+H2O$
DME direct synthesis	$4H2+2CO \rightarrow CH3OCH3+H2O$

3.2. Process Technology

Regardless of the reaction route, the reforming technology for syngas production is fundamentally necessary for DME production but is not included in this review because it is a well-defined technology. The main process technologies for producing DME are classified according to reaction chemistry. Two step technologies, i.e. indirect synthesis, employ two reactors because the reactions must take place in separate reactors. Toyo, MGC, Lurgi and Udhe have their own indirect processes for DME production. On the other hand, one step technologies, without the intermediate production and purification of methanol, include Haldor Topsoe, JFE Holdings, Korea Gas Corporation. Air Products and Chemicals has also developed a different one-step process for direct synthesis in the liquid phase. Distinguishing features of the various technologies will be given in the following section.

3.2.1. Two Step Technology

The process licensed by Toyo has two distinguishing characteristics. First, the highly active steam-reforming catalyst developed by Toyo (ISOP catalyst) is applied to the reforming unit in the syngas generation process to make the facilities more compact. Second, the high performance MRF-Z™ reactor developed by Toyo, which has the features of multi-stage indirect cooling and radial flow to the methanol synthesis unit,

allows the construction of a unit upto the capacity of 6,000 ton/day in a single train. The process is composed of single train methanol synthesis and single train DME synthesis without using an oxygen generator even in a large capacity DME plant of 3,500 tons per day. Toyo successfully operated a fuel-use DME production plant in Sichuan Province in China in 2003. In the summer of 2008, the fifth fuel-use DME plant is to start operation at Shanxi Lanhua Clean Energy CO., Ltd. in Shanxi province in China.

3.2.2. One Step Technology

Korea Gas Corporation (KOGAS) has developed a process for DME direct synthesis from natural gas, CBM and bio gas. This process consists of reformer section, CO2 removal section, methanol regeneration section, DME reactor section and DME purification section. The tri-reforming process, using natural gas that contains 10 mol% CO2, was developed for the syngas process. A new catalyst was developed by KOGAS for DME synthesis. KOGAS demonstrated a 50~100-kg/day pilot plant in 2003 and a 10-ton/day pilot plant on 2008.

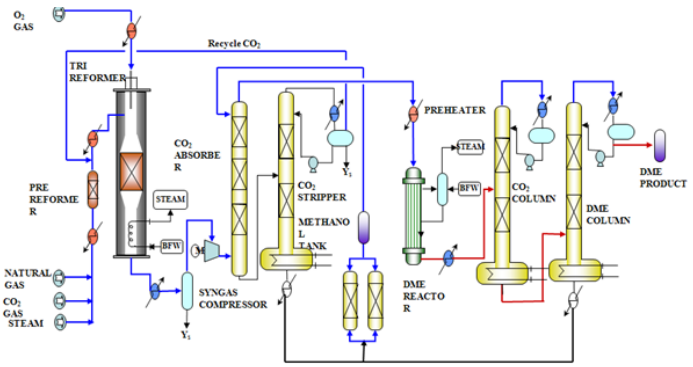


Fig. 1. The direct DME synthesis process developed by KOGAS

3.3. Related Research Works

Research on technologies for DME production has included improving catalyst activities for one-step and two-step reactions, studying reactor types such as fixed-bed reactor, slurry reactor, and fluidized-bed reactor for effectively removing reaction heat, studying the process operation, and improving optimization and scale up, and so on.

The fluidized-bed reactor is at the initial stage of laboratory testing, and its feasibility has not been yet established. Investigators are studying two different reactor types for the commercialization of DME synthesis: bubbling slurry reactor and fixed bed reactor. In a slurry reactor, fine catalyst particles are suspended in a solvent and the feed gas enters as bubbles from below. Controlling the reactor temperature is manageable, thanks to the large heat capacity of the solvent that is cooled by heat exchange tubes in the reactor. However the reactant should be transferred from gas bubbles to liquid phase solvent and then to catalyst particles for the synthesis reaction. Low diffusivity in the liquid phase can slow down the overall reaction rate. Loss of fine catalyst particles from the reactor could also be a problem. Much more data from further experiments would be needed for scale-up of the process. The fixed-bed DME process can potentially solve those problems. Therefore, our research group has developed a shell-and-tube type fixed-bed reactor with KOGAS. In a fixed bed reactor, operations are difficult, and the control of exothermic heat within the DME synthesis reactor is a primary problem.

Catalysts can be irreversibly deactivated when exposed to high temperatures (330^{o}C). Therefore, it is necessary to understand and predict reactor behaviors, including temperature profiles, under various conditions for design and scale-up. It is not feasible to gather all the data experimentally, however, and numerical simulations are highly valuable in the development of such a process.

4. Analysis of Economics

4.1. Markets

As described previously, DME presents many possibilities as a substitute of LPG, a source for power generation, and fuel for vehicle. For example, DME/LPG mixture is being used as fuel in China. We have also tested a DME/LPG mixture, rather than the complete replacement of LPG, as a possible energy source. Because the liquefaction conditions for DME are weaker than those for LPG (-42 or 7atm), DME mixed with 20% LPG can be used in the current facilities without any modifications as a residential and commercial fuel. Actual proof studies are being conducted by Korea Gas Safety Corporation, LPG Center of Japan, Jiutai Chemical of China and ENN Test results from the Korea Gas Safety Corporation show that all existing facilities (e.g. vessels, high-pressure hoses, pressure regulators, gas meters, fuse-cock, low-pressure hoses) can be used immediately at 10% price reduction.

DME as fuel for power generation produces less carbon dioxide emissions than coal or oil and has similar performances to gas. All experimental results from MHI, Hitachi and GE show that power generation using DME-containing fuel allows a safe and highly efficient operation of a dry low NOx combustion system while minimizing the generation of NOx and carbon monoxide. DME has a bright market prospect as a substitute of diesel used in automobiles. Further research and technical developments can lead to a rapid progress for DME.

Japan has developed nine types of prototype DME vehicles and installed 4 DME gas stations at various places. Japan is also trying test driving to prove actual performance. Shanghai Jiao Tong University in China has also set a goal to show DME vehicles at Shanghai in the 2010 world exhibition. In the EU, VOLVO, BUCKSHOW, and others have already finished the development and tests of DME trucks. Korea Institute of Energy Research has completed the development of DME fuel supply technology. At the present, they are operating test-runs of a 1.5-ton DME truck. They will soon complete the development of a 33-seater DME bus, and they plan to operate it in the middle of November, 2009.

DME is very valuable as an alternative to diesel thanks to its environmental factors. Measured against the enforced California ULEV (Ultra Low Emission Vehicle) Standard, DME vehicle emission satisfies numerical values of checklists (NOx, CO, HC, PM Etc.). CO and PM values for DME are remarkably lower than those required for ULEV values. Additionally, DME vehicle emission values checked against the 2009 Japan regulations showed satisfying results for the whole list (NOx, CO, HC, PM, etc).

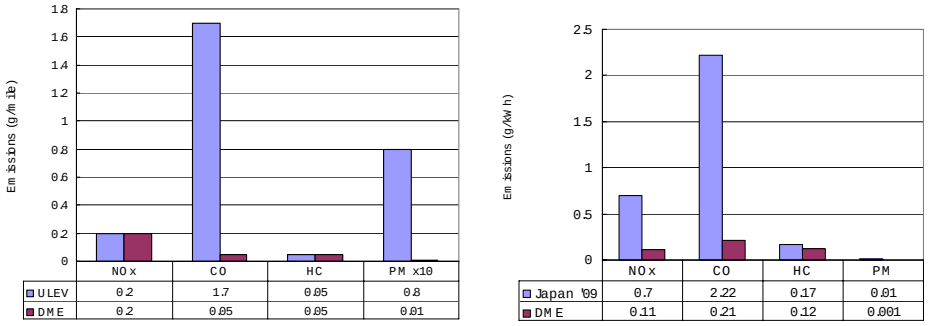


Fig. 2. A comparison of DME vehicle emission with environmentally enforced regulation.
(a) DME Truck using EGR and oxidation catalyst [source: AVL & TNO]
(b) DME Truck using EGR, de-NOx catalyst and oxidation catalyst
[source: National Traffic Safety and Environment Laboratory, Japan]

4.2. Economics Analysis

According to SRI's report in 2005, Under assumptions that DME production is 5,000MTPD (1,642,500tons/y) and feed gas is supplied at the price of 0.5 US$/MMBtu, total capital cost of a DME plant is about 590~840 Million US$ and the DME production cost is 4.45~5.03 US$/MMBtu, respectively

Table 2. DME production cost

Process	Total Capital Cost (Million US$)	Production Cost	
		US$/ton	US$/MMBtu
JFE	842.2	135.58	5.03
Haldor Topsoe	616.2	100.00	3.71
Toyo	590.5	119.93	4.45

[source: SRI Report (Sep, 2005)]

Major companies have reported that DME is economically more feasible than LPG. Toyo has reported in 2005 that DME CIF price in Japan is estimated 5.5~6.5 US$/MMBtu under the assumption of a 2 million tons/y production in Qatar and 1 US$/MMBtu for feed gas price. Toyo's report also shows that about 122 US$/MT for 2.4 million tons/y production in the Middle East can be achieved at 0.7US$/MMBtu and IRR 15 % before tax for feed gas. 122 US$/MT is equivalent to about 4.1 US$/MMBtu and LPG average price CIF Japan in the past five year is about 7.0 US$/MMBtu. Lurgi has also estimated 5~7 US$/MMBtu as DME production cost under the assumption of a 1.7 million tons/y production and feed gas price at 1 US$/MMBtu.
Recently, an economic study for a 1 million ton/year DME plant constructed in Southeast Asia for the Korean market, taking into account natural gas prices, was carried out by Korea Gas Corporation. Assuming that total capital cost is 715 Million US$ and feed gas price is 2 US$/MMBtu, DME CIF price in Korea is about 13 US$/MMBtu, 73% of average LPG CIF price in Korea based on an oil price of 90 US$. The production cost of DME is heavily sensitive on the feed gas cost rather than the product price. Therefore, in-depth considerations of feedstock, such as volatility in price, should be analyzed.

4.3. Related Research Works

There have been several economic analyses for individual GTL processes using natural gas as raw material. M.J. Gradassi and N.W. Green (1995) showed that the payout time of alternative technologies for natural gas conversion (12~91yrs) will be much longer than that of a conventional plant (7yrs). However Choi et al (1997) demonstrated that once-through natural gas FT plants can be competitive with crude oil plants, with the crude oil price at around $18.8/bbl. Patel (2005) showed that GTL diesel is competitive with existing fuel because its production cost is lower than that of refinery processes and its revenue ($2.42-3.15/MMBtu) is comparable with the returns of LNG processes ($2.64-3.08/MMBtu). Rahmim (2005) suggested that GTL processes would be viable and competitive to LNG processes when a gas price of $0.5/MMBtu or lower and capital cost of $20,000-30,000/bbl are available.

Morita (2001) compared the marketabilities of three major GTL products: FT-diesel, DME, and Methanol. Through CIF price studies, he concluded that the most promising choice for the substitute for diesel and LPG is DME when the feed gas price ranges $0.5 to 1.0/MMBtu in Japan. He showed that while FT-diesel is good for its environment-friendliness, it requires some premium to be comparable to petroleum-based diesel, and methanol is considered unsuitable because of its low power. However, there are few studies on the economics of GTL products, especially considering price variations of feedstock and product price. Along these lines, Lee (2009) compared the economic evaluations of three GTL products (FT-diesel, DME, and Methanol) to determine the optimal GTL product. In this study, comparative economic analyses were conducted using process design, cost estimation, and profitability evaluation, in that order. When the natural gas is supplied at the price of $3.00/MMbtu, constructing a DME plant is the most profitable.

5. Summary and Concluding Remarks

The exploitation of sustainable energy resources is indispensable for the development and the prosperity of human beings. Although the fuel utilization of green energy such as solar, geothermal energies, and wind or tidal power might be the ultimate goal for that, the development of synthetic energy is a currently high priority task in the aspects of environment friendliness, efficiency, energy density and possibility of commercialization. Among those, DME is one of most promising energy resources because it has better environmental performance as well as properties similar to traditional fuels.

DME can be manufactured directly from synthesis gas produced from many sources, including renewable materials (biomass, waste and agricultural products) and fossil fuels (natural gas and coal). Thus DME can increase the energy security by displacing petroleum derived fuels. Also, It has remarkable market potential as an automotive fuel, for electric power generation, and for home energy source such as heating and cooking, and is considered price competitive with LPG and diesel fuel.

Currently, various production technologies are developed and many commercial projects are actively being progressed by major industries especially in Asia. However, the profitability of DME is heavily dependent of the raw material cost, so in-depth studies of process technologies and economics with various feedstock (e.g., coal, biomass and unconventional gas such as CBM, LFG) should be preceded. Then, DME will play a major role as an alternative fuel.

References

Ali Hadipour and Morteza Sohrabi, 2007, Kinetic Parameters and Dynamic Modeling of a Reactor for Direct Conversion of Synthesis Gas to Dimethyl Ether, J. Ind. Eng. Chem., 13, 4, 558-565.

Bipin Patel, 2005, Gas Monetisation: A Techno-Economic Comparison Of Gas-To-Liquid and LNG, 7th World Congress of Chemical Engineering.

Chul-Jin Lee, Youngsub Lim, Ho Soo Kim, and Chonghun Han, 2009, Optimal Gas-To-Liquid Product Selection from Natural Gas under Uncertain Price Scenarios, Ind. Eng. Chem. Res., 48, 2, 794–800.

D. J. Wilhelm, D.R. Simbeck, A. D. Karp and R. L. Dickenson, 2001, Syngas production for gas-to-liquids applications: technologies, issues and outlook, Fuel Processing Technology, 71, 1-3, 139-148.

Daesung Song, Wonjun Cho, Dal Keun Park and En Sup Yoon, 2007, Comparison of the Performance of a Fixed Bed Reactor in the Two Cases, Mixture of Catalyst Pellets and a Hybrid Catalyst, for Dimetyl Ether Synthesis, J. Ind. Eng. Chem., 13, 5, 815-826.

Daesung Song, Wonjun Cho, Gibaek Lee, Dal Keun Park and En Sup Yoon, 2008, Numerical Analysis of a Pilot-Scale Fixed-Bed Reactor for Dimethyl Ether (DME) Synthesis, Ind. Eng. Chem. Res., 2008, 47, 13, 4553–4559.

Gerald N. Choi, Sheldon J. Kramer, Samuel S. Tam, 1997, Design/Economics of a Natural Gas Based Fischer-Tropsch Plant, Presentation at the AIChE Spring National Meeting.

Iraj Isaac Rahmim, 2005, Special Report: GTL Prospects, Oil & Gas Journal.

J. J. Lewnard, T. H. Hsiung, J. F. White, D. M. Brown, 1990, Single-step synthesis of dimethyl ether in a slurry reactor, Chemical Engineering Science, 45, 8, 2735-2741.

Kaoru Fujimoto, Kenji Asami, Tsutomu Shikada and Hiro-o Tominaga, 1984, Selective Synthesis Of Dimethyl Ether From Synthesis Gas, Chemistry Letters, 13, 12, 2051-2054.

Kohji Omata, Yuhsuke Watanabe, Tetsuo Umegaki, Gunji Ishiguro and Muneyoshi Yamada, 2002, Low-pressure DME synthesis with Cu-based hybrid catalysts using temperature-gradient reactor, Fuel, 81, 11-12, 1605-1609.

Makarand R. Gogate, Sunggyu Lee, Conrad J. Kulik, 1991, A Single-Stage, Liquid-Phase Dimethyl Ether Synthesis Process From Syngas I. Dual Catalytic Activity And Process Feasibility, Fuel science & technology international, 9, 6, 653-679.

Michael J. Gradass, N. Wayne Green, 1995, Economics of natural gas conversion processes, Fuel Processing Technology, 42, 65-83.

Shin Beom Lee, Wonjun Cho, Dal Keun Park and En Sup Yoon, 2006, Simulation of fixed bed reactor for dimethyl ether synthesis, Korean J. Chem. Eng., 23, 4, 522-530.

Sunggyu Lee, Makarand R. Gogate, Conrad J. Kulik, 1992, A novel single-step dimethyl ether (DME) synthesis in a three-phase slurry reactor from co-rich syngas, Chemical Engineering Science, 47, 13-14, 3769-3776.

Takashi Ogawa, Norio Inoue, Tutomu Shikada, Yotaro Ohno, 2003, Direct Dimethyl Ether Synthesis, Journal of Natural Gas Chemistry, 12, 4, 219-227.

Takashi Ogawa, Norio Inoue, Tutomu Shikada, Osamu Inokoshi, Yotaro Ohno, 2004, Direct Dimethyl Ether (DME) synthesis from natural gas, Studies in Surface Science and Catalysis, 147, 379-384.

Toshiyuki MII and Masayuki UCHIDA, 2005, Fuel DME Plant in East Asia, Proceedings of 15th Saudi-Japan Joint Symposium.

Troy A. Semelsberger, Rodney L. Borup, Howard L. Greene, 2006, Dimethyl ether (DME) as an alternative fuel, Journal of Power Sources, 156, 497–511.

Wen-Zhi Lu, Li-Hua Teng and Wen-De Xiao, 2003, Theoretical Analysis of Fluidized-Bed Reactor for Dimethyl Ether Synthesis from Syngas, International Journal of Chemical Reactor Engineering, 1, S2.

Wen-Zhi Lu, Li-Hua Teng and Wen-De Xiao, 2004, Simulation and experiment study of dimethyl ether synthesis from syngas in a fluidized-bed reactor, Chemical Engineering Science, 59, 22-23, 5455-5464.

Yuji Morita, 2001, Marketability of GTL from Natural Gas, The Institute of Electrical Engineers of Japan.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Global optimization of integer and mixed-integer bi-level programming problems via multi-parametric programming

Luis F. Domínguez and Efstatios N. Pistikopoulos

Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, SW7 2AZ, UK

Abstract

Bi-level programming problems (BLPPs) arise very often in areas of engineering, transportation control. A key feature of such problems from a mathematical viewpoint is that even for the simplest linear case, a global optimization approach is typically necessary. In this work, we present two multi-parametric programming based algorithms for the solution of integer and mixed-integer bi-level programming problems. The first algorithm addresses the mixed-integer case of the BLPP and employs a reformulation linearization technique (Sherali and Adams, 1990, 1994; Adams and Sherali, 2005) and continuous multi-parametric programming for the solution of the inner problem. The second algorithm addresses the integer case of the BLPP and approaches the inner problem using global multi-parametric mixed-integer programming (Dua et al. 2004). In both algorithms the solution of the inner problem is embedded in the outer problem to form a set of single-level optimization problems that can be solved to global optimality using a global optimization software.

Keywords: Bi-level Programming, Mixed-Integer Bi-level Programming, Reformulation Linearization Technique, Multi-parametric Programming.

1. Introduction

Originally introduced in the areas of game theory and multicriteria decision making, bi-level programming (BLP) has found important applications in areas of engineering, transportation and control. The main characteristic of bi-level programming problems (BLPP) is that the outer optimization problem is constrained by an inner optimization problem as follow:

$$\begin{aligned} &\min_{x,y} F(x,y) \\ &s.t.\ \ G(x,y) \\ &\quad \min\ f(x,y) \\ &\quad s.t.\ g(x,y) \le 0 \end{aligned} \tag{1}$$

where $x \in \Re^{n_1}$, $y \in \Re^{n_2}$ and functions $F(x,y)$, $f(x, y)$, $G(x, y)$ and $g(x, y)$ are assumed to be continuous and twice differentiable.

Considerable research effort has been made on devising algorithms for the mixed-integer non-linear case of Problem (1). Edmunds and Bard (1992) proposed a branch and bound (B&B) algorithm for the mixed-integer linear-quadratic BLPP and identified efficient branching rules to reduce computational burden along the difficulties associated with their B&B algorithm. Gümüs and Floudas (2005) introduced two approaches for the mixed-integer non-linear BLPP. The first approach addressed the case in which only continuous variables (linear or nonlinear) are present in the inner problem and mixed-integer variables (linear or nonlinear) are present in the outer problem. The second approach addressed the general case in which mixed-integer variables (linear or nonlinear) are present in both, the inner and outer problems. Whereas the first approach is an extension of their earlier work (Gümüs and Floudas, 2001), the second approach is based on using a reformulation linearization technique (Sherali and Adams, 1990, 1994; Adams and Sherali, 2005) to generate the continuous convex hull representation of the inner problem feasible space and transforming the reformulated inner problem using its KKT optimality conditions. The resulting single-level optimization problem is then solved to global optimality with the deterministic solver αBB. More recently, Faísca et al. (2007) proposed a parametric global optimization algorithm for various types of BLPPs. Among the classes that were covered in their work lie the linear-linear, linear-quadratic, mixed-integer linear and BLPPs with uncertainty. They used multi-parametric programming techniques developed at Imperial College (Pistikopoulos et al., 2007; Dua et al., 2002) to solve the inner problem parametrically. Rational reaction sets, obtained through the reformulation and solution via multi-parametric programming, were used to generate a set of single-linear, -quadratic or -mixed-integer programming problems which were subsequently solved to global optimality.

This work proposes two algorithms for the solution of integer and mixed-integer bi-level programming problems via multi-parametric programming. The first algorithm addresses the integer case of the BLPP and employs a reformulation linearization technique (RLT) to construct a parametric polyhedral convex hull representation of the inner problem. Because the RLT technique introduces nonlinearities during the inner problem reformulation a convexification procedure is employed in which nonlinear terms are -under and -over estimated and their convex envelopes constructed. Subsequently, continuous multi-parametric programming is employed to solve the reformulated convex inner problem. The second algorithm addresses the mixed-integer case of the BLPP and employs the same convexification procedure used in the first algorithm. In contrast to the first algorithm, which employs a continuous multi-parametric programming approach, the second algorithm makes use of multi-parametric mixed-integer programming to solve the inner problem. In both algorithms the solution obtained through multi-parametric programming is embedded in the outer problem to create a set of single-level optimization problems which are solved to global optimality using standard optimization techniques.

The paper is organized as follows. In Section 2 we provide an insight into the reformulation of the inner problem. In Section 3 & 4 we present two multi-parametric algorithms for the integer and mixed-integer cases of the BLP problem. Finally, concluding remarks are summarized in Section 5.

2. Inner Problem Reformulation

Using the property that any integer program can be substituted by its linear programming relaxation only if its constraint set can be represented by its vertex polyhedral convex hull (Wolsey, 1998), Gümüs and Floudas (2005) showed that, for integer and mixed-integer BLPPs, transformation of the inner problem using its KKT conditions produces an optimal integer solution only if a polyhedral convex hull representation of the inner constraint can be obtained. To obtain the explicit convex hull representation of the feasible solution space for the inner problem, Gümüs and Floudas (2005) employed a reformulation linearization technique (Sherali and Adams, 1990, 1994; Adams and Sherali, 2005). Due to the nonconvexities introduced by the RLT and the KKT transformation the resulting single-level optimization problem required of a global optimization algorithm for its solution.

Since the above property is also valid when the polyhedral convex hull is defined parametrically in the outer variables (Gümüs and Floudas, 2005), the idea of this work is to obtain the parametric polyhedral convex hull representation of the inner problem, reformulate the inner problem as a multi-parametric programming problem and then solve it using multi-parametric programming techniques.

Based upon the above insight, two algorithms for the solution of integer and mixed-integer BLPPs are presented next.

3. A Multi-parametric Programming Algorithm for Integer Bi-level Programming Problems

Consider the integer BLPP of the following form:

$$\begin{aligned} &\min_{x,y} F(x,y) \\ &s.t.\ G(x,y) \\ &\quad \min\ f(x,y) \\ &\quad s.t.\ g_l(x)y \le 0 \quad l = 1,...,L; \end{aligned} \tag{2}$$

where F and G are nonlinear nonconvex functions with $x \in X \subseteq [0,1]^{n_x}$ and $y \in Y \subseteq [0,1]^{n_y}$ and *f* and *g* are linear or polynomial functions.

Focusing on the inner problem, the convex hull representation of the inner constraint set is obtained by multiplying every constraint *l = 1, ...,L* in (2) by a polynomial factor $F_n(J_1, J_2)$ (Sherali and Adams, 1990, 1994; Adams and Sherali, 2005):

$$\begin{aligned} &\sum_{j=1}^{ny} g_l(x) y_j F_n(J_1, J_2) \le 0, \quad j = 1,..,L \\ &-F_n(J_1, J_2) \le 0, \quad \forall (J_1, J_2) \end{aligned} \tag{3}$$

After using the identity $y_j^2 = y_j$ and substituting $\omega_j = \prod_{j \in J_j} y_j$ and for all $j \in n_y$, a compact representation can be obtained as follows:

$$\begin{aligned} &\sum_{j=1}^{n_y} g_l(x) f_n(J_1, J_2) \le 0, \quad l = 1,..,L \\ &\text{-}f_n(J_1,\ J_2) \le 0,\ \forall (J_1,\ J_2) \end{aligned} \tag{4}$$

where $f_n(J_1, J_2)$ corresponds to the linearized version of $F_n(J_1, J_2)$.

From Equation (4) it is clear that multiplication of the inner constraint set by functional factors produces nonlinear terms $g_l(x) f_n(J_1, J_2)$ in the constraint set of the inner problem. This nonlinearity can be circumvented by constructing convex envelopes for the nonlinear terms generated by the RLT. By introducing convex estimators $\gamma_{i,j} = g_l(x) f_n(J_1, J_2)$ with corresponding envelopes $g_l^{conv}(x,y,\omega,\gamma)$, and grouping the optimization variables as:

$$y'(x) = [y(x),\ \omega(x),\ \gamma(x)]^T \tag{5}$$

the reformulated inner problem can be expressed as:

$$\begin{aligned} &\min_{y'} \ f(x, y') \\ &s.t. \sum_{j=1}^{n_y} \gamma_{l,j} \le 0, \quad l = 1, .., L \\ &\quad a'x + b'y' + c'\omega_{l,j} + d'\gamma_{l,j} \le 0,\ j = 1, ..., n_y \end{aligned} \tag{6}$$

where $a', ..., d'$ are coefficients derived from convex envelope g_l^{conv}.

Problem (6) corresponds to a multi-parametric programming problem, the solution of which can be obtained by using the algorithm of Dua et al. (2002). The solution of Equation (6) via multi-parametric programming allow us to obtain the inner variables as parametric functions of the outer variables and their regions of optimality:

$$y'(x) = \begin{cases} M^1 + N^1 x & \text{Valid in } \Pi^1 x \le \pi^1, \\ M^2 + N^2 x & \text{Valid in } \Pi^2 x \le \pi^2, \\ \quad \vdots & \quad \vdots \\ M^K + N^K x & \text{Valid in } \Pi^K x \le \pi^K, \end{cases} \tag{7}$$

where K is the number of computed parametric solutions, M^K, and π^K are constant vectors and N^K and Π^K are constant matrices obtained through the multi-parametric programming algorithm. The sets of affine functions thus obtained can be substituted into the outer problem to form K single-level I(NL)P optimization problems:

$$\begin{aligned} &\min_{x,y} F(x, y) \\ &s.t.\ G(x, y) \le 0 \\ &\quad \Pi^k x \le \pi^k, \quad k = 1, ...K \\ &\quad y'(x) = M^k + N^k x \end{aligned} \tag{8}$$

4. A Multi-parametric Programming Algorithm for Mixed-integer Bi-level Programming Problems

Consider the mixed-integer BLPP of the following form:

$$\begin{aligned} &\min_{x_1,x_2,y_1,y_2} F(x_1, x_2, y_1, y_2) \\ &s.t. \quad G(x_1, x_2, y_1, y_2) \le 0 \\ &\qquad \min_{y_1,y_2} f(x_1, x_2, y_1, y_2) \\ &\qquad s.t.\ g(x_1, x_2, y_1, y_2) \le 0 \end{aligned} \tag{9}$$

where F and G are nonlinear nonconvex functions with $x_1 \in X_1 \subseteq \Re^{n_{x_1}}$, $x_2 \in X_2 \subseteq [0,1]^{n_{x_2}}$, $y_1 \in Y_1 \subseteq \Re^{n_{y_1}}$, $y_2 \in Y_2 \subseteq [0,1]^{n_{y_2}}$ and f and g correspond to linear or polynomial functions.

Although a prametric convex hull representation of inner problem in (9) can also be obtained using a RLT, its construction is not necessary. The inner problem can be treated as the following multiparametric mixed-integer programing problem :

$$\begin{aligned} &\min_{y_1,y_2} f(y_1, y_2, \theta, \xi) \\ &s.t.\ g(y_1, y_2, \theta, \xi) \le 0 \\ &y_1 \in Y_1 \subseteq \Re^{n_{y_1}} \\ &y_2 \in Y_2[0,1]^{n_{y_2}} \\ &\theta \in \Theta \subseteq \Re^{p} \\ &\xi \in \Xi \subseteq [0,1]^{q} \end{aligned} \tag{10}$$

where the parameters θ and ξ are continuos and integer variables of the outer problem.

Problem (10) can be solved, for instance, with the algorithms of Dua and Pistikopoulos (2000), or Dua et al. (2004). By solving Problem (10) via multparametric mixed-integer programming a map of integer solutions with their continuous counterpart are obtained, each having regions of optimality of the form:

$$\{y_2, \xi\}^i \quad y_1^i(\theta) = \begin{cases} M_i^1 + N_i^1\theta & \text{Valid in } \Pi_i^1\theta \le \pi_i^1, \\ M_i^2 + N_i^2\theta & \text{Valid in } \Pi_i^2\theta \le \pi_i^2, \\ \quad \vdots & \qquad \vdots \\ M_i^K + N_i^K\theta & \text{Valid in } \Pi_i^K\theta \le \pi_i^K, \end{cases} \tag{11}$$

where $i = 1, .., I$ correspond to the number of integer feasible solutions.

Once the multi-parametric solutions for the inner problem have been found these can be embedded into the outer problem to form k single-level MI (N)LP problems as follows:

$$\begin{aligned} &\min_{x_1,x_2,y_1,y_2} F(x_1, x_2, y_1, y_2) \\ &s.t.\ G(x_1, x_2, y_1, y_2) \le 0 \\ &\quad \{y_2, x_2\}^i, y_2^i(x_1) = M_i^k + N_i^k x_1 \quad i = 1, ..., I \\ &\quad \Pi_i^k x_1 \le \pi_i^k, \quad k = 1, ..., K \end{aligned} \tag{12}$$

5. Concluding Remarks

Bilevel programming problems are inherently parametric and multi-parametric based algorithms take advantage of the parametric nature of such problems. For the solution of integer and mixed integer bi-level programming problems, we introduced two new algorithms based on continuous and mixed-integer multiparametric optimization. The proposed algorithms can provide not only the bi-level optimal solution but also any solution that is optimal for the inner and outer problems. This characteristic makes multi-parametric programming based algorithms suitable for problems involving optimistic (co-operative) formulations where there exist multiple optimal solutions and the leader can choose among them.

Acknowledgments

The authors gratefully acknowledge financial support from the Mexican Council for Science and Technology (CONACyT) and the King Abdullah University of Science and Technology (KAUST) for funding part of this research.

References

Adams, W.P. and H. D. Sherali (2005). A hierarchy of relaxations leading to the convex hull representation for general discrete optimization problems. Annals of Operations Research 140, 21–47.

Dua, V. and E. N. Pistikopoulos (2000). An algorithm for the solution of multi-parametric mixed integer linear programming problems. Annals of Operations Research 99, 123–139.

Dua, V. Papalexandri, K. and Pistikopoulos (2004). Global optimization issues in multi-parametric continuous and mixed-integer optimization problems. Journal of Global Optimization, 30, 59-89.

Dua, V., N. A. Bozinis and E. N. Pistikopoulos (2002). A multi-parametric programming approach for mixed-integer quadratic engineering problems. Computers and Chemical Engineering 26, 715–733.

Edmunds, T. and J. Bard (1992). An algorithm for the mixed-integer nonlinear bi-level programming problem. Annals of Operations Research 34(1), 149–162.

Faísca, N. P., V. Dua, B. Rustem, P. M. Saraiva and E. N. Pistikopoulos (2007). Parametric global optimisation for bi-level programming. J. Glob. Opt. 38, 609–623.

Güumüs, Z. and C. A. Floudas (2001). Global optimization of nonlinear bi-level programming problems. Journal of global 2, 1–31.

Gümüs, Z. and C. A. Floudas (2005). Global optimization of mixedinteger bi-level programming problems. Computational Management Science 2, 181–212.

Moore, J. T. and J. F. Bard (1990). The mixed integer linear bi-level programming problem. Operations Research 38(5), 911–921.

Pistikopoulos, E. N., M. C. Georgiadis and V. Dua (2007). Multi-parametric programming. Vol. 1. WILEY-VCH. Weinheim.

Sherali, H. D. and W. P. Adams (1994). A hierarchy of relaxations and convex hull characterizations for mixed-integer zero-one programming problems. Discrete Applied Mathematics 52(1), 83–106.

Sherali, H. D. and W. P. Adams (1990). A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. SIAM Journal on Discrete Mathematics 3(3), 411–430.

Wolsey, L. A. (1998). Integer programming. Wiley & Sons.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

On the optimal design of clinical tests for the identification of physiological models of type 1 diabetes mellitus

F. Galvanin[a], M. Barolo[a], S. Macchietto[b] and F. Bezzo[a]

[a]*DIPIC – Dipartimento di Principi e Impianti di Ingegneria Chimica, Università di Padova, via Marzolo 9, I-35131, Padova, Italy (federico.galvanin@unipd.it)*
[b]*Department of Chemical Engineering, Imperial College London, South Kensington Campus, SW7 2AZ London, UK*

Abstract

A model-based experiment design techniques is used to design improved test protocols for the identification of the parameters of a detailed physiological models of diabetes specific to single subjects. This paper considers the problem of parameter identification focusing on the impact of some decision variables such as sampling frequency and test duration on both the design effectiveness and the ability to meet safety critical constraints on the subject response. The proposed methodology permits to establish the minimal experimental budget required to achieve a satisfactory parameter estimation from the planned test without upsetting the subject excessively.

Keywords: model-based experiment design, parameter estimation, model identification, biomedical applications, diabetes mellitus.

1. Introduction

Type 1 diabetes mellitus (T1DM) is a metabolic disease of the glucoregulatory system affecting millions of people worldwide and causing the expenditure of millions of euros every year for health care. T1DM usually begins before the age of 40 and is characterized by a progressive destruction of the pancreatic islets and by an absolute deficiency of secreted insulin in the body. As a consequence, people with T1DM are not capable of maintaining the blood glucose concentration within a narrow normoglycaemic range and must rely on hexogenous insulin for survival. One of the most promising therapies for the disease derives from the use of an artificial pancreas, an external piece of equipment based on a control system maintaining the subject's normoglycaemia. An accurate dynamic simulation model of the glucose-insulin system can be useful to assist diabetes care and test glucose sensors, insulin infusion algorithms and decision support systems for diabetes. Several physiological models have been developed in the last four decades [1]. A specific difficulty is posed by the need to tailor a generic model to individual subjects, given the high variability of the responses, Standard clinical tests [2] are used to help diagnosing diabetes and to identify the parameters of simple models of glucose homeostasis, but, as long recognised [3], the optimal input perturbation to identify the metabolic parameters precisely could be different from that of a standard test. An "ideal" clinical test should be i) as quick and easy as possible to perform; ii) as informative as possible for parameter estimation purpose; iii) not invasive or harmful for the subject (i.e. oral tests are preferred to infusion tests) and iv) as undemanding as possible in terms of analytical resources.

Model-based design of experiments (MBDoE) [4] can be used to design a set of optimally informative clinical tests, allowing to estimate the model parameters in a statistically sound way, while at the same time fulfilling the required constraints of safety for the patient and an ease of conduction [5]. An MBDoE procedure is presented in this paper which includes, together with the definition and the assessment of a suitable optimality criterion, the definition of a set of constraints on both the experiment design variables and (more importantly here) on the dynamics of the predicted patient response. In particular, the frequency of the measurements and the duration of the test are determined, which together with other test design variables (e.g. time and size of insulin bolo uptake), affect both the optimality criterion and the dynamics of the glycaemia response in the subject. The focus of this work is to evaluate i) the impact of these design variables on the final parameter estimation, ii) the ability to meet critical safety constraints on the patient response within a specified degree of confidence and iii) the minimal experiment budget required to get a satisfactory parameter estimation from the planned test without upsetting the patient excessively. The study is carried out using a detailed physiological model of glucose homeostasis on a simulated subject.

2. MBDoE under constraints for physiological models

Several models have been proposed in the literature that are able to represent the dynamics of subcutaneous insulin infusion and of the rate of glucose appearance following a meal. In this study we adopt the Hovorka model with the insulin infusion sub-model proposed by Wilinska et al. [6] denoted as the Hovorka-Wilinska model (HWM). HWM belongs to the class of nonlinear dynamic models described by the following set of differential and algebraic equations (DAEs):

$$\begin{cases} \mathbf{f}\left(\dot{\mathbf{x}}(t), \mathbf{x}(t), \mathbf{u}(t), \mathbf{w}, \theta, t\right) = 0 \\ \hat{\mathbf{y}}(t) = \mathbf{g}\left(\mathbf{x}(t)\right) \end{cases}, \tag{1}$$

where $\mathbf{x}(t)$ is the N_x-dimensional vector of time-dependent state variables, $\mathbf{u}(t)$ and $\mathbf{w}$ are the time-dependent and time-invariant control variables (of dimensions N_u and N_w) , respectively, $\boldsymbol{\theta}$ is the N_θ-dimensional set of unknown model parameters to be estimated, and t is the time. The symbol ^ is used to indicate the estimate of a variable (or of a set of variables): thus, $\mathbf{y}(t)$ is the vector of measured values of the outputs, while $\hat{\mathbf{y}}$ is the vector of the corresponding values estimated by the model. In our case, the time-dependent manipulated inputs $\mathbf{u}(t)$ comprise the insulin subcutaneous infusion and the insulin subcutaneous bolus, whereas the glucose intake is represented as a time-invariant control variable w. There is only one measurable output $y(t)$, the blood glucose concentration G.

Model-based experiment design procedures aim at decreasing the model parameter uncertainty region predicted by model (1) as the solution of the optimisation problem

$$\boldsymbol{\varphi} = \arg\min\left\{\psi\left[\mathbf{V}_\theta\left(\boldsymbol{\theta}, \boldsymbol{\varphi}\right)\right]\right\} = \arg\min\left\{\psi\left[\mathbf{H}_\theta^{-1}\left(\boldsymbol{\theta}, \boldsymbol{\varphi}\right)\right]\right\} \tag{2}$$

subject to a set of constraints

$$\widetilde{\mathbf{C}} = z\left(\dot{\mathbf{x}}(t), \mathbf{x}(t), \dot{\mathbf{u}}(t), \mathbf{u}(t), \mathbf{w}, \boldsymbol{\theta}\right) \leq 0 \tag{3}$$

by acting on the experiment design vector $\boldsymbol{\varphi}$:

$$\boldsymbol{\varphi} = \left[\mathbf{y}_0, \mathbf{u}(t), \mathbf{w}, \mathbf{t}^{sp}, \tau\right]^T \tag{4}$$

which includes the N_y-dimensional set of initial conditions $\mathbf{y}_0$ on the measured variables, the duration of the experiment (τ), the continuously manipulated inputs $\mathbf{u}(t)$, usually approximated through a discrete (piecewise constant or piecewise linear) function and the time-constant manipulated inputs, $\mathbf{w}$. The set of time instants at which the output variables are sampled (blood sampling schedule) is also a design variable, and is expressed through the n_{sp}-dimensional vector $\mathbf{t}^{sp}$ of sampling times. $\mathbf{V}_\theta$ and $\mathbf{H}_\theta$ are the variance-covariance matrix of model parameters and the dynamic information matrix, respectively. The experiment is chosen to minimise a measurement function ψ of $\mathbf{V}_\theta$ (the particular form of the function represents the chosen design criterion [7]) which maximises the predicted information content of the experiment. The latter is expressed as the $N_\theta \times N_\theta$ dynamic information matrix

$$\mathbf{H}_\theta(\boldsymbol{\theta},\boldsymbol{\varphi}) = \sum_{k=1}^{n_{sp}} \sum_{i=1}^{N_y} \sum_{j=1}^{N_y} s_{ij} \mathbf{Q}_i^T \mathbf{Q}_j + \mathbf{H}_\theta^0 = \sum_{k=1}^{n_{sp}} \mathbf{M}_k + \mathbf{H}_\theta^0 \tag{5}$$

where $\mathbf{Q}_i$ is the $n_{sp} \times N_\theta$ dynamic sensitivity matrix of the i-th measured response, s_{ij} is the ij-th element of the $N_y \times N_y$ inverse matrix of measurements error, and $\mathbf{H}_\theta^0$ is the prior dynamic information matrix, taking into account the statistical information about the parametric system before each trial is carried out. The optimisation solution must fulfil an n_c-dimensional set of equality and inequality constraints in the form of eq. (3), where z is a complex function linking state, control and information variables. For the system under investigation there are important physical/physiological constraints that are strictly related to the physiology of the glucoregulatory system and should not be violated as they determine the maintenance of normoglycaemic levels during the test and the return to basal conditions at the end of the test.

The $\mathbf{M}_k$ matrix represents the contribution of the k-th sample to the overall predicted information content $\mathbf{H}_\theta$. This is deeply affected by the frequency of sampling (in terms of n_{sp}/τ ratio) and by the excitation pattern of the manipulated inputs; thus, the choice of sampling protocol has an impact on the effectiveness of test in terms of information availability. Discrete sampling and off-line analysis (typically, at rather low frequency) allow for the most precise analytical laboratory techniques to be used, albeit with a time delay; however, recent sampling techniques such as CGM (continuous glucose monitoring) could enrich the information content of the glucose test, allowing for a continuous recording of glucose level over a 24-hour period (high sampling frequency) [8]. In view of the above, it is worth assessing the effect of the sampling frequency in the design of a clinical test for parameter identification.

3. Case study

As in a recent study [5], the MBDoE approach is applied to a diabetic patient modelled using HWM. The purpose is to modify the usual protocol of standard tests such as PGT (postprandial glucose test) and OGTT (oral glucose tolerance test) so as to increase the overall information content of the test after a preliminary reference test. Here, the goal is to estimate precisely, with only one properly designed experiment, the set of metabolic parameters $\boldsymbol{\Theta}$ describing the insulin sensitivity of a specific subject (Θ_1, Θ_2, Θ_3) and his/her endogenous glucose production (Θ_4). Safe conditions must be guaranteed at all times during the clinical test; besides, the test should be completed within a specified time interval. It must be also guaranteed that after the test the subject may return to and remain at the basal settings. A D-optimal design procedure is carried out with the following constraints: interior constraints on the glycaemic curve to ensure

normoglycaemia at all times (60-170 mg/dL); an end point constraint on the glucose concentration (80 mg/dL), a constraint on the final derivative of the glucose concentration to ensure steady glycaemia at the end of the test, and a constraint on the test duration. The test formally ends with the last sampling point (which defines the duration of the experiment). The (simulated) glucose measurements are available with a constant relative variance of 0.033 and the elapsed time between two consecutive measurements cannot be shorter than 5 minutes.

The following protocols are proposed and assessed:

1. MPGT-n_{sp}: modified postprandial glucose test with n_{sp} = 5, 10, 20 (τ = 600 min);
2. MOGTT-n_{sp}: modified OGTT with n_{sp} = 5, 10, 20 (τ = 840 min);
3. MPGT-opt: modified postprandial glucose test with n_{sp} = 20 and τ optimised;
4. MOGTT-opt: modified OGTT with n_{sp} = 20 and τ optimised.

A variable number of samples is chosen in protocols 1. and 2. to evaluate the impact of the sampling frequency on the quality of the final estimate, while in 3. and 4. the goal is to assess whether a test duration could be shortened by increasing the sampling frequency. To improve numerical robustness, parameters are normalised with respect to the true values describing the subject.

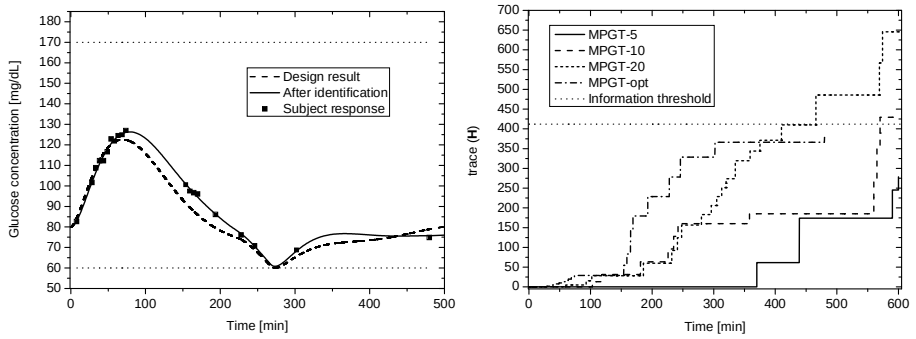
3.1. Modified Postprandial Glucose Test (MPGT)

The purpose is to identify the model set of parameters with a single experiment comprising two meals (breakfast and lunch, scheduled at 8:00 AM and 1:00 PM). The variables being optimised are: the CHO content of the meals, $D_{g,1}$ and $D_{g,2}$ (bounds on breakfast: 5 – 40 g CHO; bounds on lunch: 30 – 70 g CHO); the glucose-dependent insulin infusion rate $u_s(t)$ (parameterised as a piecewise constant function, with n_z = 9 levels and n_{sw} = 8 switching times to optimise), the amount of insulin of the boluses $u_{bol,1}$ and $u_{bol,2}$ and the sampling times. The last sampling point can be taken not later than 600 min (10 h) from the beginning of the experiment (6:00 PM). The time interval between consecutive meals is not optimised, and the insulin bolus amount is not constrained to an insulin/CHO ratio. The end point constraints on the glucose concentration and on the derivative of the glucose concentration must be reached within 600 min.

Table 1 Comparison of different MPGT protocols. Superscript * indicates *t*-values failing the *t*-test (t_{ref} is the reference *t*-value and $\boldsymbol{\Theta} = [1.000\ 1.000\ 1.000\ 1.000]^T$)

Design	Parameter Estimate $\hat{\Theta}$	Conf. Interval (95%)	*t*-values	t_{ref}
MPGT-5	[1.045 0.729 1.222 1.043]^T	[±0.3665 ±1.2940 ±1.1300 ±0.2341]	[2.85 0.56* 1.08* 4.45]	1.795
MPGT-10	[0.997 0.915 1.059 1.007]^T	[±0.2955 ±0.5136, ±0.3217 ±0.0514]	[3.37 1.78 3.29 19.58]	1.745
MPGT-20	[1.008 0.808 1.114 1.015]^T	[±0.1878 ±0.4276, ±0.2603 ±0.0430]	[5.37 1.89 4.28 23.6]	1.705
MPGT-opt	[0.944 0.826 1.147 1.020]^T	[±0.2235 ±0.7305, ±0.4295 ±0.0713]	[4.23 1.12* 2.67 14.3]	1.705

As can be seen from Table 1 at least 10 samples are necessary to perform a statistically sound parameter estimation. Figure 1 (b) shows the contribution of each sample to the overall information (5) given by the trace of $\mathbf{H}_0$. The information threshold refers to a mean standard deviation of 10% on the final estimate. When the test duration is also optimised (MPGT-opt), the samples are concentrated at the very beginning of the test, where the test is scarcely informative. The test is safe and short (τ = 480 min = 8 h) but the estimation is not statistically satisfactory (although just marginally).



(a) (b)

Figure 1 (*a*) Predicted glucose response for the optimal test, measurements from this designed (simulated) experiment and subject glucose response after identification with test data for protocol MPGT-opt and (*b*) contribution to eq. (5) for different MPGT protocols.

3.2. Modified Oral Glucose Tolerance Test (MOGTT)

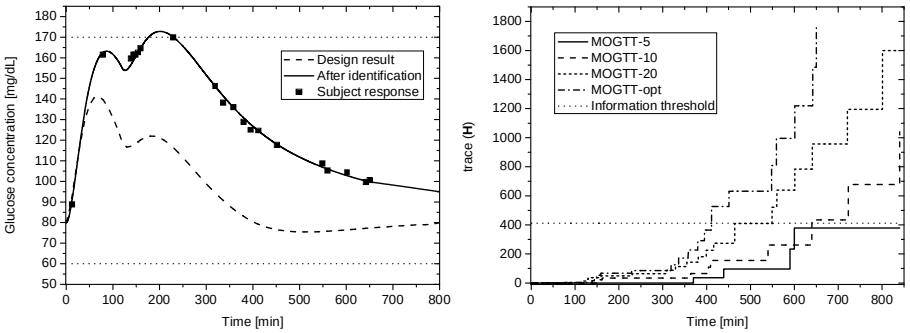
The identification test involves multiple glucose solution and insulin bolus intakes. The optimisation variables are: the glucose content of the four meals (glucose solution drink); the time interval between consecutive meals (acceptable range 15-840 min); the amount of each insulin bolus the sampling times, and duration not greater than 840 min (14 h). An additional constraint was imposed on the total amount of ingested glucose (acceptable range: 75-156 g CHO). The end point constraints on glucose concentration and its derivative must be fulfilled within 840 min from the last meal. The amount of bolus per meal was modelled as:

$$u(t) = u_{bas} + \alpha \sum_{i=1}^{N_{meals}} \delta_i(t) k_i D_{g,i} \tag{6}$$

where α = 52.63 mU/g CHO represents the optimal insulin/CHO ratio. The "relaxing factors" k_i are also optimised in the design. As underlined in [5], this test is more informative (compare the *y*-axis scale of Figure 1b and 2b), but a longer experiment duration has to be adopted since the four meals need to be spaced over a sufficiently long period of time. When τ = 840 h (14 h) the estimate is statistically satisfactory only with more than 5 samples (Table 2), but, interestingly, the MOGTT approach allows shortening the test duration from 14 to 10.8 h without much affecting the quality of the final estimate, while ensuring a safe and informative test (Figure 2).

Table 2 Comparison of different MOGTT protocols. Superscript * indicates *t*-values failing the *t*-test (t_{ref} is the reference *t*-value and $\boldsymbol{\Theta} = [1.000\ 1.000\ 1.000\ 1.000]^T$).

Design	Parameter Estimate $\hat{\Theta}$	Conf. Interval (95%)	*t*-values	t_{ref}
MOGTT-5	[1.096 0.800 1.231 1.024]^T	[±0.2437 ±0.4978 ±0.3028 ±0.1088]	[4.49 1.61* 4.06 9.41]	1.795
MOGTT-10	[0.902 1.007 1.089 1.022]^T	[±0.2139 ±0.1593, ±0.3054 ±0.1018]	[4.22 6.32 3.57 10.04]	1.745
MOGTT-20	[0.957 0.975 1.077 1.021]^T	[±0.1634 ±0.1176, ±0.2447 ±0.0817]	[5.86 8.29 4.40 12.49]	1.705
MOGTT-opt	[0.906 0.953 1.131 1.035]^T	[±0.1813 ±0.2277, ±0.3138 ±0.1014]	[5.00 4.19 3.61 10.22]	1.705



(a) (b)

Figure 2 (*a*) Glucose response modeled by design, after identification and subject response for protocol MOGTT-opt and (*b*) contribution to (5) for different MOGTT protocols.

4. Final remarks

A MBDOE has been applied to a complex model of glucose homeostasis with the purpose of evaluating the impact of sampling frequency and test duration on parameter estimation on modified Post Prandial and Oral Glucose Tolerance Tests. The results clearly show that an modified OGTT protocol is more effective with a higher number of samples and longer test duration. The results also demonstrate the flexibility of this protocol, which is simpler to be carried out and allows achieving a statistically sound parameter estimation in a short test without being invasive or harmful to the subject .

5. Acknowledgement

The financial support of the University of Padova under Progetto di Ateneo 2007 (cod. CPDA074133): "A process systems engineering approach to the development of an artificial pancreas for subjects with type-1 diabetes mellitus" is gratefully acknowledged.

References

[1] C. B. Landersdorfer, W. J. Jusko, 2008, Pharmacokinetic/Pharmacodynamic Modelling in Diabetes Mellitus, Clinical Pharmacokinetics, 47, 417-448.
[2] WHO department of non communicable disease surveillance, 2006, Definition and diagnosis of diabetes mellitus and intermediate hyperglicaemia.
[3] R. N. Bergman, L.S. Phillips, C. Cobelli, 1981, Physiologic evaluation of factors controlling glucose tolerance in man, J. Clin. Invest., 68, 1456-1467.
[4] G. Franceschini, S. Macchietto, 2008, Model-based Design of Experiments for Parameter Precision: State of the Art, Chem. Eng. Sci., 63, 4846-4872.
[5] F. Galvanin, M. Barolo, S. Macchietto, F. Bezzo, 2009, Optimal design of clinical tests for the identification of physiological models of type 1 diabetes mellitus, Ind. Eng. Chem. Res., 48, 1989-2002.
[6] M. E. Wilinska, L. J. Chassin, H. C. Schaller, L. Schaupp, T. R. Pieber, R. Hovorka, 2005, Insulin kinetics in type–1 diabetes: Continuous and bolus delivery of rapid acting insulin, IEEE Trans. Biomed. Eng., 52, 3–12.
[7] F. Pukelsheim, 1993, Optimal Design of Experiments, J. Wiley & Sons, New York, U.S.A.
[8] I. B. Hirsch, D. Armstrong, R. M. Bergenstal, B. Buckingham, B. P. Childs, W. L. Clarke, A. Peters, H. Wolpert, 2008, Clinical application of emerging sensor technologies in diabetes management: consensus guidelines for continuous glucose monitoring (CGM). Diabetes Technol. Ther., 10, 232-246.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

From a Generic Paradigm to a Generic Tool Set: Exploring Computer-Aided Multiscale Modeling

Aidong Yang, Yang Zhao

Division of Civil, Chemical and Environmental Engineering, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford GU2 7XH, UK

Abstract

Multiscale modeling is now widely regarded as a promising and powerful tool in various disciplines, including the broad area of process engineering. However, a multiscale model is usually much more difficult to develop than a single-scale model due to a range of conceptual, numerical, and software challenges. Currently, there is little support developed to facilitate multiscale modeling. This paper discusses the key challenges faced by computer-aided multiscale modeling (CAMM) and presents a methodology considered in an ongoing initiative which explores the way to a computer-based, generic and open supporting framework for multiscale modeling. Details are particularly provided on the development of a conceptual modeling tool, an important element of the envisaged tool set for CAMM.

Keywords: multiscale modeling, computer aided modeling, conceptual modeling

1. Introduction

Multiscale modeling is an emerging modeling paradigm which combines the models of different resolution scales of a complex system to obtain a high-quality characterisation of the system. This modeling paradigm is now widely regarded as a promising and powerful tool in various disciplines. As reflected by several reviews (e.g. Charpentier, 2002; Braatz et al., 2004; Vlachos, 2005), typical applications for chemical process systems include the simulation of growth of materials (such as thin films) and reaction systems involving a bulk phase and a catalytic surface, frequently characterized by combining kinetic Monte Carlo simulation with PDE/ODE-based continuum modeling. There are also process engineering applications which reply on the integration of different continuum scales for the purpose of improving computational efficiency (as opposed to handling physical heterogeneity across different scales), e.g. the coupling of a process simulator with a CFD package for modeling crystallisation and biochemical processes (e.g. Bezzo et al., 2004; Kulikov et al., 2005). Similar applications have also been developed in areas other than process engineering, including particularly material science, computational mechanics, and systems biology.

Developing a multiscale model is generally much more challenging than building a single scale model, because one has to determine what scales to be involved and how the involved scales should be connected. To tackle the difficulties, several efforts in formulating modeling methodologies have aimed to provide some kind of generic guidance to the development of new applications. The most notable activities are devoted to the classification of multiscale modeling approaches, schemes or frameworks which define the generic ways how submodels of different scales of a system can be introduced and coupled to form an integral model. Pantelides (2001) classifies scale integration strategies into serial, parallel, hierarchical, and simultaneous approaches.

This classification was expanded and adapted by Ingram et al (2004) which identifies a number of integration frameworks and discusses the compatibility issue. Li et al. (2005) articulate the difference between what is called descriptive, correlative, and variational methods. Vlachos (2006) makes the distinction between sequential or serial approaches and those referred to as hybrid, parallel, dynamic, and concurrent approaches.

While the conceptual guidance offered by the generalised modeling methodologies may help the "novice" modellers to some extent, what can provide much stronger support is computer-based tools (Marquardt et al, 2000; Ingram et al, 2004). Unfortunately, very little development exists in this direction. Consequently, the application of multiscale modeling still remains a special privilege of highly skilled modeling experts and its success is established very much on a case-specific basis (Fraga et al, 2006); the majority of scientists and engineers has yet to struggle with conceptual and practical difficulties, despite their high expectations on benefiting from this modeling paradigm.

To address the above problem, a research initiative has recently been started to explore the way to a computer-based, generic and open supporting framework for multiscale modeling. This paper discusses the starting point of this initiative, the key challenges to be addressed by computer-aided multiscale modeling (CAMM), as well as the concepts and methodology being developed. Particularly, it reports the development of a prototypical tool which supports the conceptual modeling step, as a key element in the envisaged tool set for CAMM.

2. Computer-Aided Multiscale Modeling: Starting Point and Key Challenges

In process engineering, the current lack of computer-based support to multiscale modeling is in sharp contrast with the advances in computer-aided process modeling (CAPM) prior to the era of multiscale modeling. This includes a number of commercially highly successful software packages for process modeling, simulation, and optimization. In academic research, an important methodological development in CAPM has been phenomena-based modeling, which is aimed at reducing the effort of human modellers by allowing them to work with process engineering concepts instead of mathematical details, as advocated by e.g. Stephanopoulos et al. (1990), Marquardt (1995), Perkins et al. (1996), Jensen & Gani (1996), Bieszczad (2000), and Bogusch et al. (2001). More recently, Yang et al. (2004) propose an ontology-based modeling approach to improve the generality and extensibility of phenomena-based modeling tools, which utilises ontologies (Gruber, 1993), i.e. explicit specifications of conceptualisation of domains of interest in terms of definitions of concepts and their relations, to establish conceptual models of the process systems being modeled.

Starting with the existing development of CAPM, two different sets of challenging issues must be addressed in the process of marching towards CAMM. Within the first set, there are conceptual, numerical, and software implementation related issues which will be faced by any modeller in a specific multiscale modeling application; the strategy and software framework to be developed in this work must provide corresponding solutions to these issues:

- *Conceptually*, and usually at the early stage of developing a multiscale model for a specific system/problem, the modeller has to decide what scales of the system or the subsystems should be considered, what characteristics of each of these scales should be modelled, and what connections should be established between these scales and between subsystems. Essentially "designing" the multiscale model, it often proves to be a nontrivial task especially for an inexperienced modeller.

- *Numerically*, one may expect that coupled simulation codes for different scales will run stably if each of these codes is numerically stable when being executed individually. This is however not necessarily the case (Pantelides, 2001, Rusli et al., 2004, Bezzo et al., 2005; Braatz et al., 2006). Numerical instability may arise in multiscale modeling due to e.g. temporal/spatial mismatch between the solutions at different scales or noises in the data passed between scales. Furthermore, modeling errors may occur as the consequence of handling numerical instability by means of filtering, or due to the aggregation and disaggregation of the information between different scales. To render successful coupling, one should be aware of the numerical consequences of implementing a particular coupling scheme, and apply appropriate methods to suppress numerical instability and to reduce modeling errors whenever possible.
- *With respect to software implementation*, a common issue in multiscale modeling is that different specialty modeling tools may be used for different scales of a system; these tools will have to be integrated to allow a multiscale model to be solved or executed as a whole. As these tools have usually been developed separately by different parties, significant heterogeneity often exists in their data structures, interfaces, and even supported operating systems, making the integration of these tools a very challenging task. A main research question in this aspect is how to leverage the enabling technologies provided by the CAPE-OPEN initiative (Braunschweig et al., 2000).

Additionally, there is a second set of challenging issues which are concerned with the generality and "powerfulness" of a *computer-aided* multiscale modeling solution. Briefly, the challenges there are about providing answers to the following two questions:

- *How to maximise the support to modellers at each step of model development*, given the very different nature of the conceptual, numerical, and software related problems to be tackled and given the varying levels of expertise possessed by different modellers?
- *How to make a computer-aided modeling system as generic as possible*, in light of the highly diverse nature of various multiscale modeling applications?

3. Towards a Methodology for CAMM

To effectively address the aforementioned challenges, a methodology with the following two aspects is proposed (cf. Fig. 1).

1. Establishing a theoretical framework of multiscale systems as the basis for computer-aided multiscale modeling

A unified theoretical framework may be established to cope with the diversity of applications that a generic solution for computer-aided multiscale modeling should support. This essentially follows the method proposed by Yang et al. (2004) for general systems modelling with extensions to handle multiple scales. More specifically, this framework may possess a hierarchical structure comprising a *fundamental level* and a *domain-specific level*. The former aims to provide *a unified theory for general multiscale systems*. As demonstrated in a recent publication (Yang & Marquardt, 2009), this may be achieved by a *deductive* approach on the basis of general systems theory and its formal formulations, allowing the formulation of explicit and rigorous definitions of fundamental concepts in multiscale modelling. These concepts may include (i) *system, subsystem, environment, inter-subsystem relation, system-environment relation*; (ii) *scale, inter-scale relation*, and (iii) *property and law (set) of*

all the above items. From these concepts and particularly the inter-subsystem and inter-scale relations, a variety of *generic multiscale modeling schemes* may be derived naturally and rigorously.

On top of this *fundamental level*, multiscale concepts for specific domains will be introduced. For the domain of process engineering, this level aims to provide generic definitions of scales typically involved in the multiscale modeling of process systems, such as *molecules*, *molecule clusters*, *particles*, *surfaces* and *phases*, *process units*, *plants*, *sites*, etc. Existing conceptualization of chemical process systems such as OntoCAPE (Morbach et al., 2007) may be adaptively reused for this purpose.

It is expected that, by building upon this unified, hierarchical theoretical framework, the generality of the CAMM solutions can be maximised: When a method for a particular problem (be it a conceptual, numerical, or software-related one) is developed, one can always try to base it, in the first place, only on the theory of general multiscale systems (i.e. the fundamental level). If it is successful, the resulting method will be applicable to all kinds of applications as long as the targets of modeling conform to this theory. If the method has to exploit domain-specific (process engineering for instance) concepts, the generic theory for multiscale process systems at the *domain-specific level* will come to play. This will ensure that this method be applicable to various process systems modeling applications.

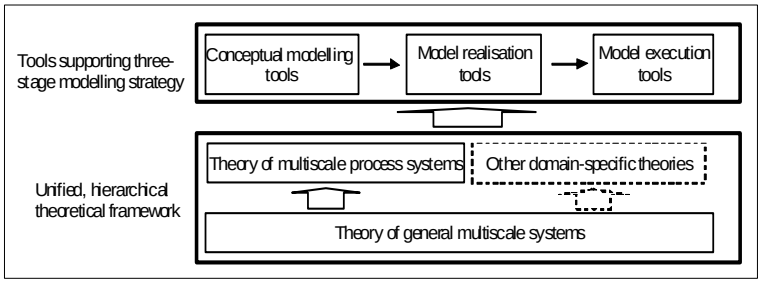


Figure 1. A methodology for computer-aided multiscale modeling.

2. Applying a three-stage strategy to maximise computer-based support

Motivated by the need of addressing different types of modeling issues while maximising computer-based support to human modellers, it would be instrumental to separate the most creative modeling tasks which inevitably require the input/intervention from the human modellers (referred to as Group 1), from other tasks which can be handled more "automatically" by software tools (referred to as Group 2). For this purpose and as an extension of the method previously proposed by Yang et al. (2004), activities in a multiscale modeling application may be organized into three successive stages: *conceptual modeling*, *model realisation*, and *model execution*. For a given modeling problem, the first stage results in a *conceptual model* which dictates ("by words") what scales to be included in the model and what kind of linkages should be established between these scales. This is a stage which mainly addresses the aforementioned *conceptual* challenges. Largely belonging to Group 1, the tasks in conceptual modeling can hardly be fully *automated*; they will rather be *supported* through the interactions between a computer-based conceptual modeling system and its users. Following this stage, the model realisation and execution stages deal with *(i)* the realisation of the conceptual model (to a form which is ready to execute) and *(ii)* the actual execution of model, respectively. The modeling tasks to be tackled in these two stages largely belong to Group 2, because they involve predominantly numerical and

software implementation related work as opposed to conceptual work. This type of work may be completed mainly by using designated software tools that implement numerical procedures for handling issues commonly encountered in multiscale modeling, hence requiring minimum effort from the human modeller.

4. A Prototypical Tool for Conceptual Modeling

Following the methodology proposed above, a prototypical tool to support conceptual modeling has been developed. It is based on the theory or conceptualization of general multiscale systems reported in Yang and Marquardt (2009) with minor extensions. This conceptualization is expressed using ontology modeling language OWL (cf. www.w3.org) and is referred to as the meta ontology. Any domain specific conceptualization or theory of multiscale systems should be derived from this meta ontology and expressed using the same modeling language. By doing so, a domain specific theory is manifested as a domain ontology which concretizes the general concepts defined in the meta ontology. For a specific modeling task, a domain ontology should offer the "top" concept, i.e. the one that represents the system to be modeled at the top level, as well as all the other concepts required for characterizing the decomposition of the system at lower levels. When applied to a specific task, this tool loads an appropriate domain ontology and establishes the possible scale structure according to the "top" concept of this task. When doing so, this tool relies purely on what is defined in the meta ontology, without needing to refer to the semantics of the domain specific terms. This feature is maintained in supporting the rest of the conceptual modeling process as schematically illustrated in Figure 2 for modeling a N-scale system.

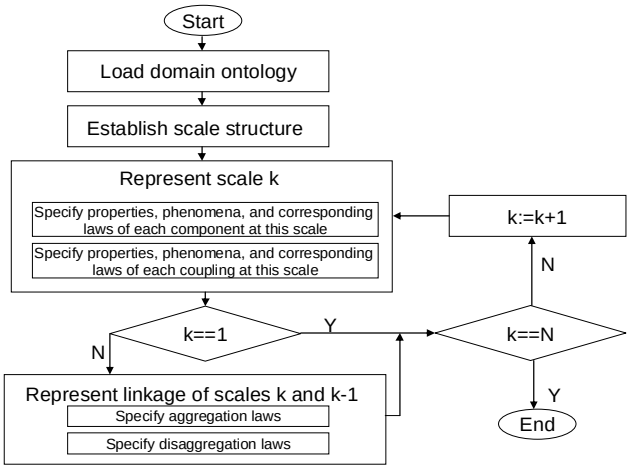


Figure 2. A simplified representation of the workflow for conceptual modeling.

The main function of the tool is to guide a human modeler by means of a series of modeling steps leading to a complete conceptual model of a specific (type of) system. The laws specified for the individual components at various scales may materialize in the form of either symbolic model building blocks or some kind of existing software code, a difference that will impact and be handled by the subsequent stages of model realization and model execution. The laws that govern the inter-scale relations will materialize as "glue" blocks to be offered as part of the tool set for the model realization

stage. These blocks will be joined by some others in this tool set which are to address common numerical issues as outlined earlier in Section 2. The development of the tools for modeling realization and execution is the subject of on-going work.

5. Concluding Remarks

As a generic and promising modeling paradigm, multiscale modeling is yet to gain computer-based support in order to help human modelers overcome difficulties in conceptual, numerical, and software implementation aspects. The development of such support may be guided by a methodology which is grounded on a unified, hierarchical framework of multiscale systems theories to achieve a high degree of generality, and which suggests building a set of tools around three successive modeling stages in order to maximum the facilitation the tools can offer to human modelers. This methodology has been partly implemented by the development of a prototypical tool for conceptual modeling, while work on the evaluation of the other parts of the methodology is ongoing. Given the fact that computer-aided multiscale modeling is still in its infant stage, the initiative reported in this paper is undoubtedly of an explorative nature; long term efforts with multidisciplinary collaborations are naturally expected in order to make substantial progresses in this area.

Acknowledgement

This work is supported by the UK Engineering and Physical Science Research Council (EPSRC) under Grant No EP/G008361/1.

References

Charpentier, J.-C. (2002). *Chemical Engineering Science*, 57, 4667-4690.
Bezzo, F., Macchietto, S., Pantelides, C. C. (2004). *Comput Chem Engng* 28, 501–511.
Bezzo, F., Macchietto, S., Pantelides, C. C. (2005). *AIChE Journal*, 51, 1169-1177.
Bieszczad, J. (2000). PhD Thesis, Department of Chemical Engineering, MIT.
Bogusch, R., Lohmann, B., Marquardt, W. (2001). *Comput Chem Engng, 25*, 963-995.
Braatz, R.D., Alkire, R.C., Rusli, E., Drews, T.O. (2004). *Chem Eng Sci* 59, 5623 – 5628.
Braatz, R.D., Alkire, R.C., Seebauer, et al. (2006). *Journal of Process Control*, 16, 193-204.
Braunschweig, B.L., Pantelides, C.C., and Britt. H.I. (2000). *Chem Eng Prog*, 96, 65–76.
Fraga, E.S., Wills, G., Fairweather, M., Perris, T. (2006). *Computer Aided Chemical Engineering*, 21,457-462.
Gruber (1993). *Knowledge Acquisitions*, 5, 199-220.
Ingram, G.D., Cameron, I.T., Hangos, K.M. (2004). *Chem Eng Sci* 59, 2171 – 2187.
Jensen, A. K., Gani, R. (1996). *Comput Chem Engng*, 20, S145-S150.
Kulikov, V., Briesen, H., Marquardt, W. (2005). *Chem Eng Res Des*, 83 (A6), 706-717.
Li, J., Ge, W., Zhang E. J., Kwauk M. (2005). *Chem Eng Res Des*, 83 (A6), 574–582.
Marquardt, W. (1995). *Comput Chem Engng*, 20 (6/7), 591-609.
Marquardt, L.V. Wedel and B. Bayer (2000). In: M.F. Malone, J.A. Trainham and B. Carnahan, Editors, *Foundations of computer-aided process design*, 96, 192–214.
Morbach, J., Yang, A., Marquardt, W. (2007). *Journal of Engineering Applications of Artificial Intelligence*, 20 (2), 147-161.
Pantelides, C. C. (2001). *Computer Aided Chemical Engineering*, 9, 15-26.
Rusli, E., Drews, T.O., Braatz, R.D. (2004). *Chem Eng Sci*, 59, 5607-5613.
Stephanopoulos, G., Henning, G., Leone, H. (1990). *Comput Chem Engng*, 14, 813-846.
Vlachos, D. G. (2005). *Adv. Chem. Eng.*, 30, 1.
Yang, A., Morbach, J., Marquardt, W. (2004). In: C. A. Floudas, R. Agrarwal (Eds.): *Proceedings of FOCAPD 2004*, 591-594.
Yang A., Marquardt, W. (2009). An ontological conceptualization of multiscale models. *Comput Chem Engng*, In Press.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Design of Experiments of Steam Deacidification of Edible Oils for Modeling Validation

Roberta Ceriani1,2, Klicia A. Sampaio1, Simone M. Silva1, Thiago Taham1 and Antonio A.J.A. Meirelles1

[1]*Faculty of Food Engineering, State University of Campinas*
[2]*Faculty of Applied Sciences, State University of Campinas*
P.O. Box 6121, 13083-862 Campinas – SP, Brazil
E-mail address: tomze@fea.unicamp.br

Abstract

The physical deacidification step must be designed in order to take the high difference in the volatility of undesirable compounds and the neutral oil under high temperatures and vacuum, without causing damage to nutraceutical compounds, excessive production of *trans* fatty acids and neutral oil loss. Improving the process conditions of this important stage of vegetable oil refining is a relevant topic nowadays. The present study aimed to investigate experimentally the batch steam deacidification of fatty systems with known compositions (two models systems and one vegetable oil), while keeping an accurate measurement and control of the main process variables (temperatures, pressure and mass percentage of injected steam). With this precise information, a validation of a simulation tool developed in the research group was achieved by comparing experimental results with computational ones.

Keywords: steam deacidification, computational simulation, buriti oil, palm stearin, palm olein.

1. Introduction

Steam deacidification and deodorization processes in the oil industry aim to vaporize odoriferous compounds and fat acids from the oil, and are based on the high differences in volatility between the oil and the majority of its unwanted substances. The separation is successfully accomplished by applying high temperatures and low pressures. However, these processing conditions also ease the occurrence of loss of neutral oil, the retention/recovery of nutraceutical compounds, and chemical reactions.

Steam deacidification of edible oils involves necessarily triacylglycerols plus fatty acids and, eventually, some minor compounds, as partial acylglycerols, tocopherols and carotenes. The batch process, although rarely used in industrial scale, gathers the main features of the continuous equipment and is suitable for research purposes. In fact, it has been extensively studied in the literature, but none of the previous works provide information in a way to allow a precise modeling of the entire system, and consecutively, the simulation of the process. In fact, this problem has been highlighted in recent articles published in the literature dealing with simulation of steam deacidification and deodorization of edible oils (Ceriani and Meirelles, 2007). In this context, our goal was to investigate the steam deacidification of model systems and of a real system (vegetable oil) with a reliable controlling and monitoring of processing

variables, such as temperature, pressure and amount of stripping steam injected, and with pertinent chemical analysis of the load (feed) and products (distillate and deacidified oil). Considering the goal was to obtain accurate data for validation of modeling/simulation, the independent variables (temperature and % of sparge steam) were monitored during each experiment in a five minute interval by means of two thermocouples installed at different points of the reactor and a microperistaltic pump.

2. Experimental Procedures

2.1. Materials

Pure tricaprylin (99%), commercial triolein (65-88%) and oleic acid (72%) (model system M_1). Stearin and olein from palm oil and oleic acid (model system M_2). Crude buriti (*Mauritia flexuosa*) oil (real system R_1).

2.2. Experiments

The experiments were conducted in a batch deodorizer developed by the research group, as an improved version of the equipment described by Decap et. al.(2004). In all experiments, steam was injected for 60 minutes after the desired temperature of the experiment has been reached, as usually occurs in such procedures (Petrauskaitè, et. al., 2000). The temperatures of the liquid in the reactor and of the volatile phase were measured in a five minute interval throughout the experiment. Pressure was also monitored using both a digital signal for vacuum and a mercury column for absolute pressure. Samples were collected after 60 minutes of stripping and immediately cooled, to avoid formation of acidity. These samples were analyzed in accordance with the procedures described in Section 2.3. Mass balances attested the reliability of the experiments. Each system (M_1, M_2 and R_1) was submitted to experimental trials as detailed below.

2.2.1. System M_1

The system M_1 was composed by 31:69 (m/m) of tricaprylin/triolein with an addition of 3.5 % of a free fatty acid (oleic acid). Experiments with 900g of this model system were done following a simple factorial design with 2 independent variables, temperature (ranging from 190°C up to 260°C) and steam flow rate (from 0.5 % up to 3.5 % of the load). Samples were collected and analyzed to obtain the responses of interest (acidity and neutral oil loss) and to evaluate the effects of process variables on these responses.

2.2.2. System M_2

The system M_2 was composed by different proportions of stearin and olein and also oleic acid, so the initial acidity was set as 3.0 %. Experiments with 1000g of this model system were done following a full factorial design with 3 independent variables, temperature (from 200°C up to 260°C), steam flow rate (from 0.5 % to 6.5 % of the load) and the proportion (m/m) between stearin and olein (from 0.1 up to 0.9). Samples were collected and analyzed to obtain the responses of interest (acidity and neutral oil loss).

2.2.3. Real System (R_1)

Crude buriti oil, extracted from a palm tree native of Brazil, has a high concentration of monounsaturated fatty acids (MUFA) and nutraceutical compounds, as tocopherols (antioxidant) and carotene (vitamin A activity). The selection of this oil as the real system, allowed studying also tocopherol loss and carotene degradation during deacidification process. Initial acidity of crude buriti oil was 3.11%, expressed as oleic acid. Experiments with 700 g of crude buriti oil were conducted following a full factorial design with 2 independent variables, temperature (ranging from 150°C to up to

260ºC) and steam flow rate (from 0.5 up to 8.5% of the load). In this case, samples were collected after 30 and 60 minutes of stripping.

2.3. Factorial design

For multivariable processes, such as food systems, a large number of factors (processing conditions) can influence the outcomes (responses) of the process. Correlating observed changes in individual factors (process variables) to changes in a response of interest is, usually, not an easy task. In this work, factorial design was applied to sketch the process conditions of experiments and simulation, according to Box et al. (1978).

2.4. Analytical Procedures

The free fatty acid content (acidity) was determined by titration with sodium hydroxide according to the IUPAC method 2201 (IUPAC, 1976). Tocopherol content was determined by HPLC combined with diode array detector following methodology proposed by SILVA et al. (2009). Total tocopherol content was obtained by fractions (α-, β-, γ- and δ-) sum. The moisture in the oily phases was determined by Karl Fischer titration according to the AOCS official method Ca 2e 55 (AOCS, 1998) and in the aqueous phase was determined according to the AOCS official method Ca 2c-5 (AOCS, 1998). These results were used in mass balances to assure the quality of the experiments. The neutral oil loss was calculated by difference.

2.5. Simulation

Oil composition and process parameters (temperatures and pressures measured during the experiments) are inputs of the simulation program. For each fatty system (M_1, M_2 and R_1), the fatty acid composition was determined by Gas Chromatography following official method Ce 1-62 (AOCS, 1998). The acylglycerol profile of crude buriti oil was determined following the method Cd 11c-93 (AOCS, 1998). The probable triacylglycerol (TAG) composition was determined using the method of Antoniossi Filho et al. (1995), cited by Ceriani & Meirelles (2007). Di- and monoacylglycerols of buriti oil were obtained from the stoichiometry of the hydrolysis reaction of TAG.

Ceriani and Meirelles (2007) modeled a laboratory-scale batch deodorization as a multicomponent differential distillation. In this process, a still (batch deodorizer) is fed and then heated until the deodorization temperature is reached. The injection of sparge steam begins and the process is performed until the required oil acidity is obtained. The volatile compounds flow overhead, are condensed and collected in a receiver. Thus, the whole deodorization process can be divided in two parts: heating (in absence of water) and stripping with sparge steam at constant temperature, which was allowed by the presence of water in the liquid-phase. The computational simulation program consisted of a differential distillation model with molar balances and equilibrium relationships. The energy balance was suppressed, and the molar vaporization rate was set as the experimental value found in each experiment. More detailed information about the computational tool/methods can be found referring to Ceriani and Meirelles (2007).

3. Results and Discussion

For sake of simplicity, this work presents only the experimental results and the comparison of them with the simulations. The results obtained by statistical models (for M2 and R1 systems) are not shown. The experimental results of the simple factorial design for model system M_1 and its comparison with the simulation results are shown in the Table 1.

The experimental results for model system M_1 indicated that: (1) while higher T (°C) and % steam improve the removal of free fatty acids, they also increase neutral oil loss

to unacceptable levels (trials 2 and 4), and (2) there is a synergetic effect between these two variables.

Table 1 – Experimental and simulation data from system M_1.

Trial	Variables		Acidity[a] (%)		NOL[b] (%)	
	T (°C)	% steam (mL/min)	Exp.	Simul.	Exp.	Simul.
1	190	0.5 (0.075)	3.23±0.02	3.40	0.06	0.06
2	190	3.5 (0.525)	0.14±0.01	0.15	7.53	7.17
3	260	0.5 (0.075)	2.60±0.02	3.22	0.17	0.14
4	260	3.5 (0.525)	0.12±0.01	0.01	13.59	13.06
5	225	2.0 (0.3)	2.19±0.02	2.22	0.58	0.64
6	225	2.0 (0.3)	2.11±0.01	2.49	0.67	0.68
7	225	2.0 (0.3)	2.27±0.03	2.69	0.51	0.52
RES[c]			0.25		0.14	

[a] values expressed as oleic acid; [b] Neutral Oil Loss, [c] Residue (Eq.1 - mean value).

Trials 5, 6 and 7 (central point of the factorial design) indicate a good replication of the experiments. In relation to the simulation of the process, a satisfactory correlation between its results and the experimental ones was generally achieved. Note that, in some cases, the simulation results coincided with the experimental values found. The residues shown in Table 1 were calculated using the absolute difference between the experimental and simulation results:

$$RES = ABS\ (experimental - simulation) \qquad (eq.1)$$

Table 2 shows the results for the complete factorial design of model system M_2.

Table 2 – Experimental and simulation results for model system M_2

Trial	Variables			Acidity[a] (%)		NOL[b](%)	
	T(°C)	%steam	stearin/olein	Exp.	Simul.	Exp.	Simul.
1	212.1	1.71(0.285)	0.262	2.46±0.02	2.70	0.04	0.02
2	247.9	1.71(0.285)	0.262	0.68±0.01	0.83	0.05	0.22
3	212.1	5.29(0.882)	0.262	1.21±0.02	1.28	0.02	0.13
4	247.9	5.29(0.882)	0.262	0.35±0.01	0.45	0.29	0.29
5	212.1	1.71(0.285)	0.738	2.16±0.02	2.44	0.06	0.04
6	247.9	1.71(0.285)	0.738	0.43±0.02	0.74	0.15	0.24
7	212.1	5.29(0.882)	0.738	1.13±0.07	1.45	0.17	0.12
8	247.9	5.29(0.882)	0.738	0.10±0.01	0.19	0.39	0.40
9	200.0	3.50(0.583)	0.500	2.45±0.03	2.42	0.06	0.04
10	260.0	3.50(0.583)	0.500	0.11±0.03	0.17	0.04	0.46
11	230.0	0.50(0.083)	0.500	2.20±0.02	2.87	0.16	0.01
12	230.0	6.50(1.083)	0.500	0.42±0.02	0.65	0.25	0.23
13	230.0	3.50(0.583)	0.100	1.37±0.01	1.51	0.07	0.11
14	230.0	3.50(0.583)	0.900	0.75±0.01	0.88	0.14	0.19
15	230.0	3.50(0.583)	0.500	0.74±0.02	0.90	0.20	0.19
16	230.0	3.50(0.583)	0.500	0.79±0.01	0.90	0.18	0.19
17	230.0	3.50(0.583)	0.500	0.75±0.02	0.91	0.17	0.19
RES[c]				0.19		0.07	

[a] values expressed as oleic acid; [b] Neutral Oil Loss, [c] Residue (Eq.1 - mean value).

Besides the effects of temperature and of % steam, already discussed, one can see the

effect of increasing the ratio stearin/olein (exemplified by the pair 13-14). Increasing the proportion of stearin in the system increases its volatility, which eases the removal of acidity and also the occurrence of NOL. Once more, the comparison between simulation results and the experimental ones gave low residues for acidity and neutral oil loss.

Experimental data for the deacidification of buriti oil (Table 3) shows that 30 minutes of stripping was not sufficient to get acidity below 0.3%, as recommended by Codex Alimentarius (1999). Sixty minutes of stripping, however, were enough to get more satisfactory values (0.31 and 0.29 for trials 4 and 6, respectively). It is interesting to note the evidence of hydrolysis reaction of triacylglycerols during trial 5: acidity values after 30 and 60 minutes of stripping (3.32% and 3.37%, respectively) were greater than the initial value.

Trial 6 presented the highest neutral oil loss, followed by trials 2 and 4, as an effect of higher temperatures. This was an expected value, because neutral oil losses are related with high temperatures. In general, the neutral oil loss values were very low in other trials.

Table 3 – Experimental data for buriti oil after stripping

	Variables		30 minutes		60 minutes		
Trial	T (ºC)	% steam (mL/min)	Acidity[a] (%)	Tocopherol (mg/100g)	Acidity[a] (%)	NOL[b]	Tocopherol (mg/100g)
1	166	1.66 (0.194)	3.04±0.01	135.9±5.1	3.01±0.01	0.00	135.6±2.6
2	244	1.66 (0.194)	2.61±0.04	124.7±0.6	1.43±0.02	1.84	122.3±1.6
3	166	7.33 (0.855)	3.10±0.01	119.7±6.5	2.93±0.02	0.68	128.4±1.6
4	244	7.33 (0.855)	1.02±0.01	136.5±2.0	0.31±0.02	1.33	132.8±2.9
5	150	4.5 (0.525)	3.32±0.01	113.8±1.4	3.37±0.01	0.31	106.4±2.3
6	260	4.5 (0.525)	0.85±0.01	129.3±3.7	0.29±0.01	5.28	105.7±2.5
7	205	0.5 (0.058)	3.01±0.03	135.7±0.3	2.91±0.01	0.00	133.5±4.0
8	205	8.5 (0.992)	2.74±0.02	135.8±2.4	1.96±0.01	1.06	130.8±1.8
9	205	4.5 (0.525)	---[c]	---[c]	2.19±0.03	0.30	132.9±2.7
10	205	4.5 (0.525)	2.73±0.02	148.3±2.3	2.17±0.02	1.34	131.6±1.5
11	205	4.5 (0.525)	2.82±0.01	135.9±5.1	2.41±0.01	1.02	119.7±2.0

[a] values expressed as oleic acid; [b] Neutral Oil Loss, [c] Sample not collected

Buriti oil used in this work had all fractions of tocopherols. The most important ones were α- and β- fractions (61.4 mg/100g and 68.7 mg/100g, respectively), followed by δ-tocopherol (13.6 mg/100g) and γ-tocopherol (5.0 mg/100g). Buriti oil also had, in lower concentrations, γ- and δ-tocotrienol (1.2 and 1.8 mg/100g, respectively).

Table 3 brings the results in terms of total tocopherols (α- and β- fractions) for each trial, after 30 and 60 minutes of stripping. As one can see, the main effect on the loss of tocopherols is the applied temperature. Besides this effect, which increases the volatility of tocopherols, there is also a consumption of tocopherols during the deacidification process caused by their antioxidant activity. This observation, in fact, explains the value found for trial 5 in which the lowest temperature was applied.

Table 4 shows the results of total tocopherols obtained by the simulation tool for the processing conditions of trials 1 to 11. As one can see, a very good representation of the experimental results was obtained for samples after 30 and 60 minutes of stripping. The average relative deviations between the experimental and simulation results were 5.52% and 7.96%, respectively.

Table 4 – Simulation results for total tocopherol (mg/100g) for buriti oil after stripping

Trials	T (ºC)	% steam (mL/min)	30 minutes	60 minutes
1	166	1.66 (0.194)	135.26	135.29
2	244	1.66 (0.194)	134.93	134.14
3	166	7.33 (0.855)	135.33	135.42
4	244	7.33 (0.855)	134.43	129.96
5	150	4.5 (0.525)	135.26	135.29
6	260	4.5 (0.525)	133.47	127.12
7	205	0.5 (0.058)	135.25	135.27
8	205	8.5 (0.992)	135.37	135.38
9	205	4.5 (0.525)	135.32	135.33
10	205	4.5 (0.525)	135.31	135.34
11	205	4.5 (0.525)	135.26	135.36
RES [a]			6.87	9.14

[a] Residue (Eq.1 - mean value).

Controlling and monitoring of the experiments allowed a better understanding of the process studied, enabling a better mathematical modeling of the process. The simulation tool was, then, efficient in predicting the losses of neutral oil, the acidity and tocopherols in the final oil, and gave a good perspective to help the quality of the final oil improving, by defining the process parameters according to raw material and final product specification.

References

A.O.C.S. (1998) Official Methods and Recommended Practices of the American Oil Chemist's Society, vol.1 – 2, 3 ed., Champaign

CODEX ALIMENTARUIS (1981) Codex Standard for Edible Fats and Oils – Codex Stan 19-1981 (Rev. 2-1999).

G.E.P. Box, W.G. Hunter, J.S. Hunter (1978) Statistic for experiments – an introduction to design, data analysis and model building. New York, John Wiley & Sons

IUPAC (1979) Standard methods for the analysis of oils, fats and derivatives, (6th ed., Part 1, sections I and II). In: PAQUOT, C., Pergamon Press

P. Decap, S. Braipson-Danthine,B. Vanbrabant, W. De Greyt, C. Deroanne (2004) Comparison of Steam and Nitrogen in the Physical Deacidification of Soybean Oil, J Am Oil Chem Soc, v.81, n.6, p.611 – 617

V. Petrauskaitè, W. De Greyt, M.J. Kellens (2000) Physical Refining of Coconut Oil: Effects of Crude Oil Quality and Deodorization Conditions on Neutral Oil Loss. J Am Oil Chem Soc, v.77, n.6, p.581-586.

R. Ceriani, A.J.A. Meirelles (2007) Formation of trans PUFA during Deodorization of Canola Oil: A Study through Computational Simulation, Chem. Eng. Proc., v.46, p. 375-385

S.M. Silva, K.A. Sampaio, T. Taham, S.A. Rocco, R. Ceriani, A.J.A. Meirelles (2009) Characterization of oil extracted from buriti fruit (*Mauritia flexuosa*) grown in the Brazilian Amazon Region, J Am Oil Chem Soc (*in press*)

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Interfacing IPOPT with Aspen Open Solvers and CAPE-OPEN

Weifeng Chen, Zhijiang Shao* and Jixin Qian

State Key Laboratory of Industrial Control Technology, Institute of Industrial control,Zhejiang University, Hang Zhou 310027, China

Abstract

CAPE-OPEN (CO) is brought forward as the standard for the next generation of process system simulation platform. In this work, the IPOPT algorithm and interface developed by (Lang and Biegler, 2005) is extended. IPOPT is encapsulated and is embedded into Aspen Plus through Aspen Open Solvers (AOS) Interface. A scaling module based on starting point is developed for scaling the models from Aspen Plus before optimization. The validity of AOS compliant IPOPT is tested by solving the economic optimization problem of depropanizer and debutanizer multi-column system and the reconciliation problem of large-scale air separation system. Compatibility of Aspen Plus for CO solvers is extended based on AOS interface through COM technology. In order to demonstrate validity of the extended compatibility, CO compliant IPOPT is integrated with Aspen Plus. It is tested by the same examples as AOS compliant IPOPT. Further analysis is made on effect of model scale on efficiency loss since the introduction of CO.

Keywords: Aspen Open Solvers, CAPE-OPEN, Scaling, Optimization

1. Introduction

Process system simulation and optimization have long been used as means to estimate the economic benefit and impact for environment of different design options. There have been many process simulation tools with excellent solvers, abundant unit models, or precise thermodynamic and physical properties. However, there are still critical demands for effective interactions among different modules.

CAPE-OPEN (CO) is brought forward as the standard for the next generation of process system simulation platform. It is conceived to achieve the goals of facilitating process modeling and design and increasing interoperability among various process simulation tools. The CO interoperability standard allows interoperation among heterogeneous Process Modeling Components (PMCs) and Process Modeling Environments (PMEs) from different software providers, independent of their algorithmic details, programming languages and operating system (Yang et al, 2007).

Many highly competitive software suppliers and famous research groups are involved in developing and utilizing CO compliant software, however, most of them focus on the development of CO compliant unit operations and CO compliant thermodynamic and physical properties (Ricardo et al, 2008; Testard et al, 2005). CO compliant numeric solvers, especially optimization solvers, are seldom concerned.

In this paper, the IPOPT algorithm and interface developed by (Lang and Biegler, 2005) is extended. Section 2 discusses IPOPT for Aspen Plus through Aspen Open Solvers (AOS) Interface. Section 3 presents compatibility of Aspen Plus for CO solvers based

*Corresponding author. E-mail address: zjshao@iipc.zju.edu.cn

on AOS Interface through COM technology. Section 4 then presents a number of case studies that test the validity of AOS compliant IPOPT and the compatibility of Aspen Plus for CO solvers. A further analysis is made on the effect of model scale on efficiency loss. Finally, some concluding remarks are given.

2. AOS Compliant IPOPT

This section deals with two issues for embedding IPOPT into Aspen Plus using Aspen Open Solvers (AOS) (Aspen Technology, 2003). First, a scaling module based on starting point is developed for scaling the models. And then, IPOPT is encapsulated into Dynamic Link Library (DLL) and an AOS wrapper for IPOPT is developed.

2.1. Scaling Module

Since the scaled models from Aspen Plus sometimes are not suitable for IPOPT to solve because of the ill-conditioned matrix. A scaling module based on starting point is developed for scaling the models again before optimization.

The scaling method is different from (Lid et al, 2008) which scales all the nonzero values of the Jacobian matrix of constraints. Original and scaled NLP formulations are

$$\begin{cases} \min\limits_{x} \quad f(x) \\ \text{s.t.} \quad c(x) = 0 \\ \qquad x \ \geq \ 0 \end{cases} \quad \Rightarrow \quad \begin{cases} \min\limits_{\tilde{x}} \quad \tilde{f}(\tilde{x}) \\ \text{s.t.} \quad \tilde{c}(\tilde{x}) = 0 \\ \qquad S_x \tilde{x} \ \geq \ 0 \end{cases} \tag{1}$$

where the scaled variable $\tilde{x} = S_x^{-1} x$, the scaled objective $\tilde{f}(\tilde{x}) = f(S_x \tilde{x})$, the scaled constraints $\tilde{c}(\tilde{x}) = S_c^{-1} c(S_x \tilde{x})$. S_x and S_c are fixed diagonal scaling matrices. The proposed scaling procedure is as follows,

1. Scale variables based on the starting point. If the absolute initial value of some variable is larger than one, it will be scaled. The variable scaling matrix is as follows, where x_0 is the starting point

$$S_x(i,i) = \begin{cases} |x_0(i)|, if \ |x_0(i)| > 1 \\ 1, else \end{cases} \tag{2}$$

2. Scale constraints based on the initial Jacobian matrix, $\partial c_j / \partial x_i$ is a column vector.

$$S_c(j,j) = \begin{cases} \left\| S_x(\partial c_j / \partial x_i) \right\|_\infty, if \left\| S_x(\partial c_j / \partial x_i) \right\|_\infty > 1 \\ 1, else \end{cases} \tag{3}$$

The scaling module exposes an interface (IScalingModelInfo) to supply IPOPT with all the required information.

2.2. Embedding IPOPT into Aspen Plus

We encapsulate IPOPT into DLL which exposes a class NLPStruct and a function AOS_IPOPT (NLPStruct* mynlp). NLPStruct is derived and the member methods of NLPStruct are overridden utilizing IScalingModelInfo interface. AOS_IPOPT is the entrance of the whole algorithm. PseudoSolver implementing AOSNumericNLPSystem and AOSNumericNLPSystemFactory interfaces exposes GetFactoryNLPSystem

method. When loading PseudoSolver, Aspen Plus looks for GetFactoryNLPSystem method and invokes it to create an instance of CNumericNLPSystemFactory class. For an AOSNumericNLPESO object in a being solved model, there is a corresponding CNumericNLPSystem instance allocated by the CNumericNLPSystemFactory instance in the PseudoSolver. Finally, AOSNumericNLPSystem is invoked and AOS_IPOPT function is called. IPOPT gets the required information of the model being solved by calling NLPStruct interface and CNLPStruct gets corresponding information by calling IScalingModelInfo interface. Interactions between IPOPT and Aspen Plus though AOS are shown as figure 1.

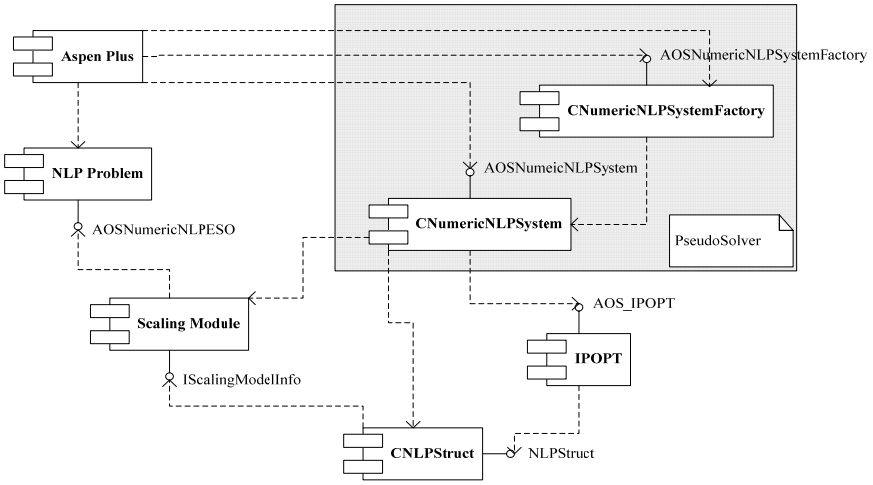


Fig 1 Component Diagram of Interaction between AOS Compliant IPOPT and Aspen Plus

3. Compatibility of Aspen Plus for CO Solver

Aspen Plus supports the CO Unit Operation and CO Thermodynamics standards. Unfortunately, so far there is no report that Aspen Plus is compatible with CO solvers. The compatibility of Aspen Plus for CO solvers is extended based on AOS Interface through COM technology.

The CO Optimization Interface issued by CO-LaN for MINLPs (CO-LaN, 2003a) is comprised of ICapeMINLP, ICapeMINLPSolverManager and ICapeMINLPSystem. One of the first steps for achieving the compatibility of Aspen Plus for CO compliant NLP solver is to implement the common CAPE-OPEN interfaces, such as Identification Common Interface (CO-LaN, 2003b), Parameter Common Interface (CO-LaN, 2003c) and Collection Common Interface (CO-LaN, 2003d). All the CAPE-OPEN components implement Identification Common Interface to provide their names and descriptions. Parameter Common Interface includes ICapeParameter, ICapeParameterSpec, ICapeOptionParameterSpec, ICapeRealParameterSpec, ICapeIntegerParameterSpec and ICapeBooleanParameterSpec. A CCapeParameterSpec class is created by implementing all the above referred parameter interfaces except ICapeParameter and it is contained in CCapeParameter class. The CCapeParameter class inherited from ICapeParameter represents all the parameters of the NLP solver. In order to facilitate handling parameters of the NLP solver, a CParameterCollection class is created by implementing ICapeCollection interface with map template class as a container.

NLP is a special case of MINLP, so it can be completely described by ICapeMINLP. CCapeNLP exposing ICapeMINLP is created by using IScalingModelInfo to represent the scaled model. CCapeMINLPSystem inherited from ICapeMINLPSystem connects a

selected NLP solver with a special CCapeNLP instance. CCapeMINLPSolverManager creates a CCapeMINLPSystem instance corresponding to a special CCapeNLP instance. It also handles the selection and configuration of the parameters of the selected solver.
In order to illustrate the progress of embedding CO solver (in this paper, we utilize CO IPOPT (Lang and Biegler, 2005) and it can be replaced by other CO solver immediately) into CO compliant Aspen Plus clearly, the component diagram of interaction between CO compliant IPOPT and Aspen Plus is given as figure 2. We can compare it with figure 1 to find out what makes Aspen Plus compatible with CO solver.

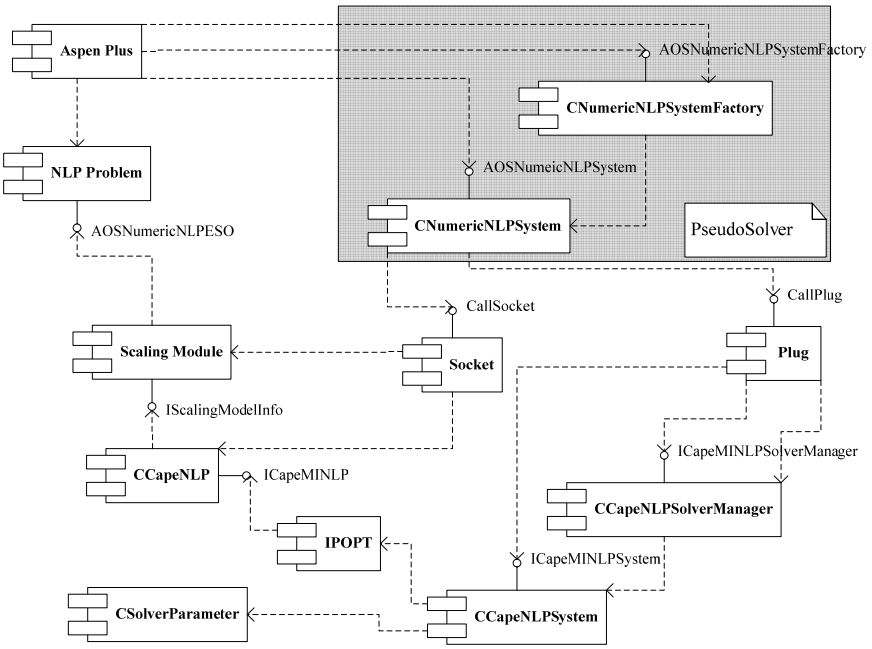


Fig 2 Component Diagram of Interaction between CO Compliant IPOPT and Aspen Plus

4. Numerical Experiments

This section deals with two issues. First, we test the validity of AOS compliant IPOPT by solving the economic optimization problem of depropanizer and debutanizer multi-column system and the reconciliation problem of large-scale air separation system. Simultaneously, we demonstrate the validity of the compatibility of Aspen Plus for CO solver by solving the same examples as AOS compliant IPOPT. Second, the effect of model scale on efficiency loss of CO compliant IPOPT is evaluated by solving two variable scale models.
All the numerical experiments are performed on the machine Dell Latitude D520 running Windows XP and with Intel(R) CPU T2300, 1.66GHz and 1.5G RAM memory.

4.1. Validity Test

Case 1 The depropanizer and debutanizer multi-column system is an important part in the ethylene production. We can refer to (Jiang et al, 2006) for the illustration of the flow sheet and the formulation of the optimization problem. For this NLP problem, number of variables is 3889, number of equality constraints is 3887, the degree of freedom is 2, and the number of nonzero elements of Jacobian in constraints is 76197. The results for solving this problem with AOS compliant IPOPT and CO compliant IPOPT respectively are presented in table 1. AOS represents the CPU time for solving

the problem with AOS compliant IPOPT. CO-InProcess represents the CPU time for solving the problem with CO compliant IPOPT run in "in process" (in the same process as Aspen Plus) mode while CO-Local is CO compliant IPOPT run in "out of process" (a separate process on the same computer as Aspen Plus) mode.

Tab.1 Numeric result of depropanizer and debutanizer multi-column system

objective	iteration	feasibility	AOS (s)	CO-InProcess (s)	CO-Local (s)
1.961e+1	11	1.95e-7	9.42	16.37	17.97

The problem is solved successfully and all the three runs have the same objective value, number of iterations and feasibility, we verify the validity of AOS IPOPT and the compatibility of Aspen Plus for CO IPOPT. Moreover, from table 1, we note that there is little difference in CPU time for solving this model in "in process" mode and in "out of process" mode.

Case 2 A large-scale air separation system was constructed which contains 36 units and 68 feed streams. Each stream has 3 components which are oxygen, nitrogen and argon. The measurement data of the process system reflects the run status of equipments and is the basis for process simulation and optimization. As the measurement error of data is inevitable, we need to revise the data before they are used as experiment data.

For this reconciliation model, the number of variable is 7184, the number of equality constraints is 7168, the degree of freedom is 16, and the number of nonzero elements of Jacobian of constraints is 40034. The results for solving this model using AOS compliant IPOPT and CO compliant IPOPT are shown in table 2.

Tab.2 Numeric result of large-scale air separation

objective	iteration	feasibility	AOS (s)	CO-InProcess (s)	CO-Local (s)
4.005e-2	173	9.44e-7	75.02	188.30	204.64

This model is also solved successfully and all the three runs have the same objective value, number of iterations and feasibility.

4.2. Efficiency Test

From section 4.1, we see that CPU time of CO wrapping increases much compared with AOS compliant IPOPT. Interoperability of Aspen Plus is gained at the expense of efficiency loss. A further test is made on the effect of model scale on efficiency loss. Two variable scale examples from (Biegler et al, 2000; Schmid, 1994) are tested.

Example 1

$$\min\ 0.5\sum_{i=1}^{n} x_i^2$$

$$x_1(x_{j+1}-1)-10x_{j+1}=0,\ j=1,\ldots,n-1$$

$$x_j^0=0.1$$

Example 2

$$\min\ \sum_{i=1}^{n-1}(x_i+x_{i+1})^2$$

$$x_j+2x_{j+1}+3x_{j+2}-1=0,\ j=1,\ldots,n-2$$

$$x_1^0=-4,\ x_i^0=1,\ i>1$$

The scale of the both examples varies from 5,000 to 200,000 with step 5,000. The relationship between the model scale and the CPU time cost is illustrated by figure 3. 'no CO' represents that the model is solved by AOS compliant IPOPT while 'CO' represents CO compliant IPOPT. 'Extra' represents the extra overhead because of the introduction of CO (namely 'CO' minus 'no CO'). Figure 3 indicates that a nearly linear relationship holds between extra overhead and model scale. The reason for this phenomenon is the implementation of the CO interface involving extra memory allocation and data copy.

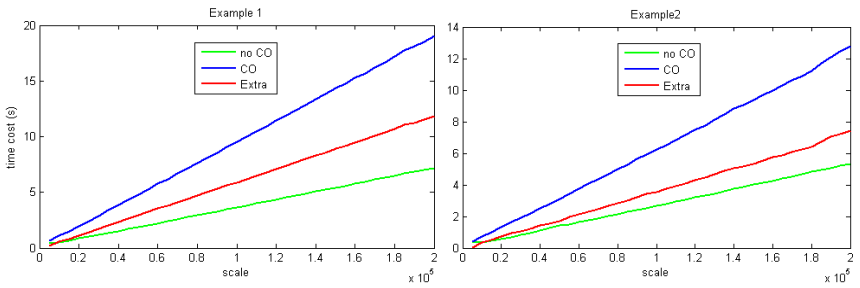


Fig 3 Relationship between the model scale and the CPU time cost

5. Conclusion

In this paper, we descirbe the development of AOS compliant IPOPT. The compatibility of Aspen Plus for CO solvers is extended based on AOS Interface through COM technology. The validity of AOS compliant IPOPT is tested by solving the economic optimization problem of depropanizer and debutanizer multi-column system and the reconciliation problem of large-scale air separation system. Moreover, validity of the extended CO compatibility of Aspen Plus is tested with the same examples as AOS compliant IPOPT.

IPOPT is embedded into Aspen Plus successfully as a new tool for solving process optimization problems. Compatibility of Aspen Plus is extended for CO compliant solver. A nearly linear relationship holds between the efficiency loss and model scale.

References

Aspen Technology, Inc. ,2003, Aspen Open Solvers 12.1 User Guide, available at http://support.aspentech.com/webteamasp/My/FrameDef.asp?/webteamasp/AllDocsDB.asp

L. T. Biegler, J. Nocedal et al.,2000, Numerical experience with a reduced Hessian method for large scale constrained optimization, Computational Optimization and Applications, 15, 45-67.

CO-LaN Consortium, 2003a, Open Interface Specification: Optimization Interface, version 2. http://www.colan.org/Spec%2010/Optimisation%20Interface%20Specification.pdf

CO-LaN Consortium, 2003b, Open Interface Specification: Identification Common Interface, version 3. http://www.colan.org/Spec%2010/Identification%20Common%20Interface.pdf

CO-LaN Consortium, 2003c, Open Interface Specification: Parameter Common Interface, version 5. http://www.colan.org/Spec%2010/Parameter%20Common%20Interface.pdf

CO-LaN Consortium, 2003d, Open Interface Specification: Collection Common Interface, version 2. http://www.colan.org/Spec%2010/Collection%20Common%20Interface.pdf

A. P. Jiang, Z. J. Shao, et al., 2006, Simulation and Optimization of Distillation Column Sequence in Large-Scale Ethylene Production. Chem. Ind. Eng. (China), 57, 9, 2128-2134.

Y. D. Lang, L. T. Biegler, 2005, The IPOPT Interface to CAPE-OPEN. CAPE-OPEN Update, Vol. 10, available at http://www.colan.org/CO%20Update/IPOPTmain_Vol10.htm

T. Lid, S. Skogestad, 2008, Scaled steady state models for effective on-line applications. Computers & Chemical Engineering, 32, 990-999.

M. R. Ricardo, G. Rafiqul, et al., 2008, Use of CAPE-OPEN standards in the interoperability between modeling tools (MoT) and process simulators (Simulis Thermodynamics and ProSimPlus). Chemical Engineering Research and Design, 86, 7, 823-833.

C. Schmid, 1994, Reduced Hessian Successive Quadratic Programming for Large-Scale Process Optimization, Ph. D. thesis, Carnegie Mellon University.

L. Testard, J. P. Belaud, 2005, A CAPE-OPEN based framework for process simulation solution integration. Computer Aided Chemical Engineering, 20, 1, 607-612.

A. Yang, B. Braunschweig, et al., 2007, A multi-agent system to facilitate component-based process modeling and design. Computers & Chemical Engineering, 32, 2290-2305.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Software Factory for the Generation of CAPE-OPEN compliant Process Modelling Components

Amine Lajmi,[a] Sylvie Cauvin,[a] Mikal Ziane[b]

[a]*IFP - Technology Computer Science and Applied Mathematics Division, 1 & 4 avenue de Bois-Préau, 92852 Rueil-Malmaison Cedex, France*
[b]*Laboratoire d'Informatique de Paris 6 (LIP6), 104 avenue du Président Kennedy, 75016 Paris, France*

Abstract

The maintenance of CAPE-OPEN compliant process modelling components is a complex task. It requires accurate knowledge about three interconnected domains: the process itself, the CAPE-OPEN interfaces specification, and the middleware (COM, CORBA, and .NET). Consequently, maintenance tasks require the collaboration of several experts throughout the entire component lifecycle. Tools that assist experts in performing these tasks are thus required. This paper presents a tool that embeds enough knowledge about these three aspects, to automatically generate compliant code. Our approach is a software factory that takes as inputs three separated models describing: the process modelling component, the standard specification, and the middleware. These models are combined and refined using successive model transformations, until code is generated. In order to anticipate the evolution of the three domains, transformations are expressed on stable abstractions with respect to expected changes.

Keywords: CAPE-OPEN, Model Driven Engineering (MDE), Software Evolution

1. Introduction

The increasingly growing complexity of Computer-Aided Process Engineering (CAPE) tools, and the diversity of suppliers and technologies, has resulted in the emergence of interoperability needs, and led to the CAPE-OPEN standard [2]. This standard defines the interfaces that off-the-shelf Process Modelling Components (PMCs) [1] must implement to communicate with each other, in any compliant process modelling environment (e.g. PRO-II, Hysys). Since the adoption of the standard by the CAPE community, several tools that assist the developers of process modelling components have been made available by the CO-LaN [6]. These tools are used to guide the developer with producing CAPE-OPEN compliant components.

Most of these tools follow the same principle: the source code which does not support the calculation logic is automatically generated, while the implementation of the calculation code is left to the developers. However, in almost all of these tools, the generated code becomes hard to maintain when the component configuration changes, the standard evolves, or the middleware changes.

Indeed, the generated code includes concepts from three domains: the component model (e.g. ports, parameters), the standard specification (e.g. packages, interfaces, types) and the middleware specification (e.g. object structures, persistence, registration). Since in most existing wizards, these specifications are hard-coded, the developers need to collaborate with the wizard suppliers in order to understand how a given change affects the calculation routines. Moreover, they have to do the task manually because most existing wizards override the calculation code.

In this paper, we propose an approach that allows expressing changes through three separated views, and automatically propagating the effects of these changes to the already generated code, without losing any manually-added code. This way, the separation of views and the automation of code updates allow reducing considerably the maintenance effort of the generated code. This paper is organized as follows. Section 2 gives examples of changes impacting the three considered views. Section 3 presents our approach. Section 4 describes the validation process, and demonstrates the tolerance of the approach regarding changes impacting the three view. Section 5 summarises some related work, while a short conclusion is made in section 6.

2. Taking changes into account

Changes are considered along three axes: the process modelling component, the CAPE-OPEN specification, and the middleware. For example, a change in CAPE-OPEN can be moving operations from one interface to another (part of the transition from version 0.93 to 1.0). A change in the middleware can be the migration from COM to .NET or vice versa. A change in the process modelling component can be a modification in a unit operation's configuration, for example changing its ports and/or parameters.

As an illustrating example, suppose given a unit operation used to mix adiabatically two inlet streams to obtain a single outlet stream. A first step towards the development of such a unit is to use an existing wizard, generally specific to a standard version (e.g. CAPE-OPEN 1.0), and dedicated to COM middleware. One can start by configuring the unit ports, for example by specifying each port's name, type, and direction. Once the source code is generated, the calculation code is to add manually in the body of the Calculate method, which is localized in the ICapeUnit interface [2].

A common change impacting the unit operation is to add a supplementary inlet port, such that the unit mixes now three inlet streams instead of two inlet streams. In practice, there are two ways to take this change into account. The first way is to modify both the generated code and the calculation code manually. This requires from the developer additional effort such as declaring and initializing the new inlet port outside the calculation algorithm (namely in the constructor of the unit). The second way is to re-configure the unit using the same wizard and regenerate the code, and then modify the calculation code manually. This solution is less effort consuming for the developer, however, it requires from the wizard additional mechanisms that allow to preserve the calculation code during change propagation.

3. The Approach

Changes are expressed as editing operations on graphical models. In Model-Driven Engineering (MDE) terminology, a model is a graph of objects typed over -and constrained by- a meta model, which is a graph of classes (for example, a class diagram). A meta model is, in turn, typed over -and constrained by- a standard language, for example the Eclipse Modeling Framework (EMF)'s Ecore [11]. Model transformations are automatic operations performed on models in order to produce other models, or textual data. We call a view an association between a model and its typing meta model. The approach is illustrated in Fig. 1. Three separated views of the unit operation (1), the standard specification (2) and the middleware (3) are first introduced. The code generation process is a sequence of model transformations performed in two steps: first, an intermediate model (4) is produced, next code is generated.

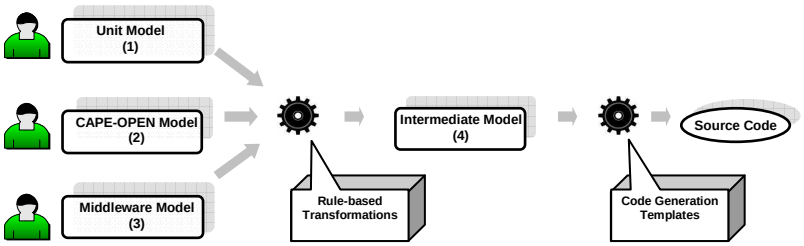


Fig.1. Overview on the approach

3.1. The unit operation view

A simplified unit meta model is presented in Fig.2. The developer configures the unit operation ports and parameters, typically using a graphical editor, which is automatically generated, by EMF, from the meta model.

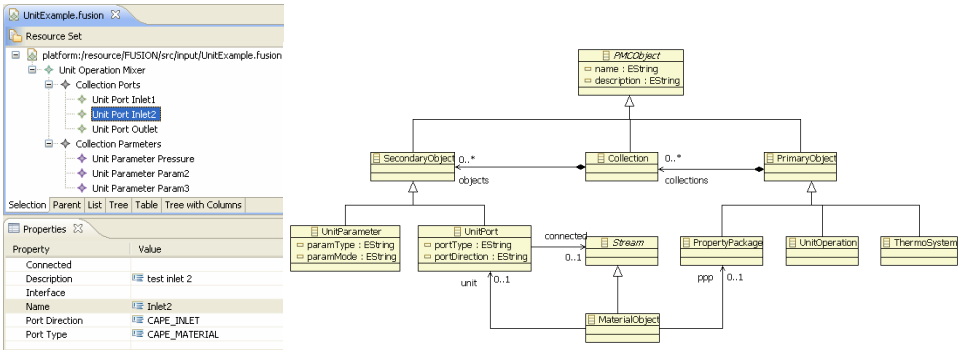


Fig.2. A unit model (on the left), the unit meta model (on the right)

3.2. The CAPE-OPEN view

The CAPE-OPEN meta model has been elaborated in order to give the user a dedicated language to express the standard types definitions. This allows specifying the standard using abstractions that are middleware-independent.

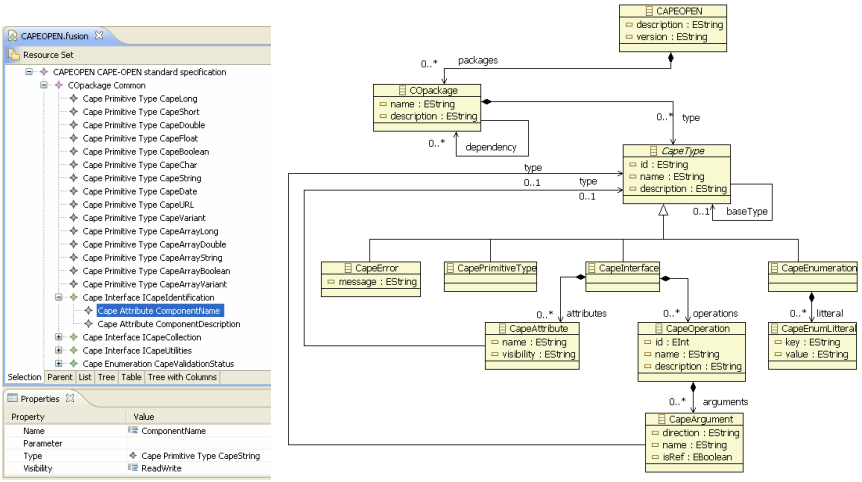


Fig.3. A CAPE-OPEN model (on the left), the CAPE-OPEN meta model (on the right)

3.3. The middleware view

The .NET meta model (Fig. 4) is inspired from [10]. Concepts relevant to COM Interop technology have been added to make the generated unit available to COM simulators. Externally, the generated unit is seen by any COM simulator as a COM based unit, assuming the .NET framework is installed in the hosting machine.

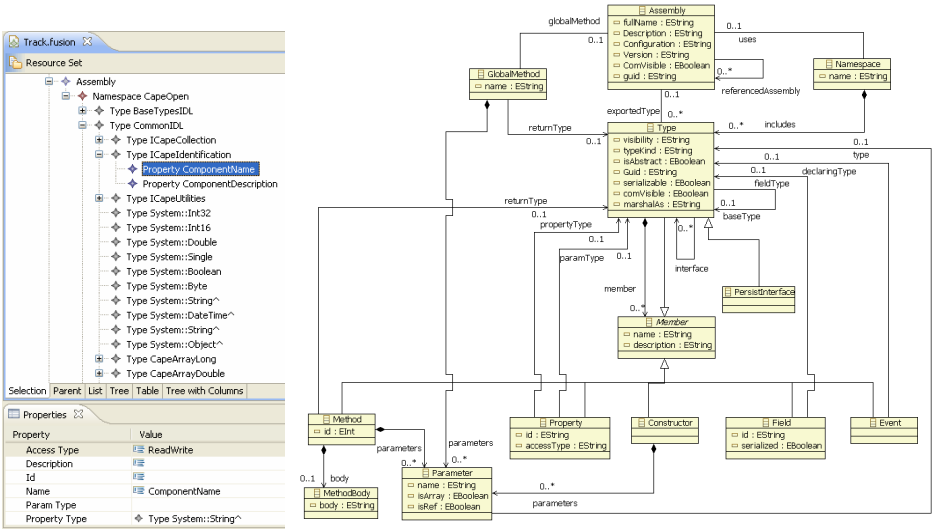


Fig.4. A .NET model (on the left), the .NET meta model (on the right)

3.4. The code generation process

The first transformation (see Fig. 1), designed using Tiger EMF [8], is based on the algebraic graph transformation theory [5]. It takes as inputs a unit operation model (1), a CAPE-OPEN model (2), and generates an intermediate model (4). The intermediate meta model is the union of the three meta models, plus traceability links used to guide the transformation strategy. The second transformation, written in Xpand language [9], is a combination of templates that take as input the model (4), and generate code.

4. Validation

In this section we give some elements towards the validation of the approach. Then, we discuss the tolerance of the approach regarding some elementary changes impacting the three views. Finally, we validate the current implementation on the example of section 2, and show how the generated code is updated while the calculation code is preserved

4.1. Validation of the approach

A first level of validation consists in validating each artefact involved in the proposed code generation process, which means meta models and transformations.

- Meta models

To validate the meta models, we adopted an iterative domain modelling approach based on code abstraction. First, an existing unit operation code has been refactored such that loosely coupled code blocks are clearly identified. Next, at each step, an abstraction is validated, and thus added to the meta model, if a change in its instances produces a code that can be compiled and tested in a compliant simulation environment. Such continuous meta model updates impact the development of the code generation process. However, the impact is quite limited in the first transformation because rules are

designed using the graphical editor of Tiger EMF, and their associated API is automatically generated by the tool.

- Rule-based transformations

Formal validation of the syntactical correctness of the first transformation has been made on general EMF models in [4]. Hence, the produced model fulfils both requirements: (i) it is typed over the output meta model, and (ii) it fulfils all the associated constraints.

- Code generation templates

The templates are validated with respect to two criteria: (i) the syntactical correctness of the generated code, and (ii) the ability of the templates to avoid writing over existing manually-added code.

4.2. Validation on case studies

Assuming meta models and transformations are validated, a second level of validation consists in quantifying the effort of taking into account some preliminary changes that impact the three considered views.

From the CAPE-OPEN view, elementary changes (e.g. changing operations signatures, moving operations) are performed manually using the associated graphical editor, and are automatically propagated by the code generation process. However, these changes have to be propagated also to the simulation environment, so that it can implement the same (changed) version of the standard.

From the middleware view, the use of .NET is due to the fact that the evolution from COM to .NET has been already anticipated by Microsoft. Indeed, mechanisms that allow making .NET components available to COM (e.g. COM Interop) have been integrated to the .NET framework as additional APIs. Hence, we expect to use the actual middleware view to generate both COM and .NET unit operations.

As to the unit operation view, some elementary changes have been tested (e.g. adding/removing ports, adding/removing parameters). In the example of section 2, we wanted to check if changes to the unit operation view were easily taken into account (e.g. adding a unit port). The following template code is responsible of instantiating all the ports configured in the unit operation model.

```
CapeOpen::Mixer::Mixer(void)
{
          «EXPAND TUnitPort::AddPorts»
}
«DEFINE AddPorts FOR TrakingModel»
«LET unitoperation.collections.get(0) AS ports»
«FOREACH ports.objects.typeSelect(UnitPort) AS aPort»
          this->PortCollection->Add(gcnew CUnitPort("«aPort.name»", "«aPort.description»",
          CapePortDirection::«aPort.portDirection», CapePortType::«aPort.portType»));
«ENDFOREACH»
«ENDLET»
«ENDDEINE»
```

The generated C++ code (below) is localized in the unit constructor, and called by the simulator when the unit is added to the flowsheet. This code is out of the scope of the calculation routines. A run is required to update it, unless the developer does the task.

```
CapeOpen::Mixer::Mixer(void)
{
          this->PortCollection->Add(gcnew CUnitPort("Inlet1", "test inlet 1",
          CapePortDirection::CAPE_INLET, CapePortType::CAPE_MATERIAL));
          this->PortCollection->Add(gcnew CUnitPort("Inlet2", "test inlet 2",
          CapePortDirection::CAPE_INLET, CapePortType::CAPE_MATERIAL));
          this->PortCollection->Add(gcnew CUnitPort("Outlet", "test outlet",
          CapePortDirection::CAPE_OUTLET, CapePortType::CAPE_MATERIAL));
}
```

In order to avoid losing the calculation code, the body of the Calculate method has been put into a protected region, so that its corresponding part in the generated code is not overridden by the next run. This preliminary solution allows updating the constructor of the unit operation, without writing over the calculation code.

```
«PROTECT CSTART "/*" CEND "*/" ID this.name+"::calculate"»
/**********************************/
/*    ADD CALCULATION CODE HERE    */
/**********************************/
/**********************************/
«ENDPROTECT»
```

5. Related Work

Several software tools are used to guide the developer on the steps to be taken towards producing CAPE-OPEN compliant components. Currently, two COM specific wizards are available on the CO-LaN website: the first one is compliant with version 0.93, and the second is compliant with version 1.0. Other unit operation wizards are also available: one written in C++ (donated by IFP) and one written in FORTRAN 90 (donated by TOTAL). Each of these tools is specific to a version of the standard, and a middleware. More recently, the development and integration of CAPE-OPEN compliant PMCs using the .NET framework has been discussed in [3]. A detailed COM/.NET interoperability guideline is also available on the website.

6. Conclusions and Perspectives

In this paper, we proposed an iterative model-driven approach for the generation of CAPE-OPEN compliant process modelling components. The approach takes three separated models as inputs, representing three expert views on the development of process modelling components. The combination of these models and code generation are delegated to the tool, using a sequence of model transformations. Future work will include a decomposition of changes into sequences of atomic editing operations that leads to tracking the history of changes within each expert view, and the elaboration of a more expressive unit meta model that allows generating parts of the calculation logic.

References

1. B. Braunschweig, and R. Gani, "Software Architectures and Tools for Computer Aided Process Engineering". 2002: Elsevier. 659.
2. J.-P. Belaud, B. Braunschweig, M. White, "The CAPE-OPEN Standard: Motivations, Development Process, Technical Architecture and Examples". In: Software Architectures and Tools for Computer Aided Process Engineering, vol. 11, (2002), pp. 303-332
3. W. M. Barrett Jr., J. Yang, “Development of a chemical process modelling environment based on CAPE-OPEN interface standards and the Microsoft .NET framework”. In: Computers and Chemical Engineering 30 (2005), pp. 191-201
4. Enrico Biermann, Claudia Ermel, Gabriele Taentzer: Precise Semantics of EMF Model Transformations by Graph Transformation. In: MoDELS 2008, pp. 53-67
5. H. Ehrig, K. Ehrig, "Overview of Formal Concepts for Model Transformations Based on Typed Attributed Graph Transformation, Electronic Notes in Theoretical Computer Science 152 (2006), pp. 3-22
6. The CAPE-OPEN Laboratories Network, http://www.colan.org
7. COCO: The CAPE-OPEN to CAPE-OPEN Simulator, http://www.cocosimulator.org
8. TIGER EMF Transformation Project, http://tfs.cs.tu-berlin.de/emftrans/
9. openArchitectureWare Project, http://www.openarchitectureware.org
10. Abd-Ali, J. and Guemhioui, K. E. 2005. An MDA-Oriented .NET Metamodel. In: Proceedings of the Ninth IEEE international EDOC Enterprise Computing Conference (September 19 - 23, 2005). EDOC. IEEE Computer Society, Washington DC, 142-156.
11. Eclipse Modeling Framework (EMF), http://www.eclipse.org/emf

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Multiscale Modelling for Computer Aided Polymer Design

Kavitha Chelakara Satyanarayana[a], Jens Abildskov[a], Rafiqul Gani[a], Georgia Tsolou[b], Vlasis G. Mavrantzas[b]

[a]*CAPEC, Department of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark.*
[b]*Department of Chemical Engineering, University of Patras & Institute of Chemical Engineering and High-Temperature Chemical Processes (ICE-HT/FORTH), Patras GR 26504 Greece*

Abstract

A multiscale modelling approach for designing polymers is presented in this work. Here, the desired properties of a polymer is given as input in the polymer design problem, computer aided molecular design (CAMD) algorithm using group contribution plus models (macro-meso scale) gives out the polymer repeat unit structures satisfying the desired properties as output. The arrangement of polymer repeat unit structures to form a polymer chain and the properties corresponding to the generated polymer are studied in micro-scale approach. A case study using this multiscale approach is presented in this paper.
Keywords: CAMD, Polymer design, Multiscale modelling, Polyisobutylene.

1. Introduction

The search for new polymeric compounds with desired end-user properties is an important problem in chemical, pharmaceutical and food industries. There are a number of methods that can be used for designing new polymers with desired properties. But each method has its own set of advantages and limitations. For example some are too time consuming (like quantum mechanics based methods) while some are capital intensive (like laboratory scale preparation of too many polymers). The ability to predict the physical and chemical properties of macromolecular materials from their molecular structure prior to synthesis is of great value to the design of products with higher knowledge content, new functionalities and improved performance. Computer-aided approach can expedite the design process by establishing input-output relations between the type and number of molecular groups in a polymer repeat unit and the desired macroscopic properties. Integration of such computer-aided approaches on the basis of the molecular descriptor used to describe the polymer structures (*i.e.* based on the time and length scale shown in figure 1) is considered in this work which is called a *multi scale approach* for computer aided polymer design.

Identification of a basic polymer repeat unit at macro scale level requires an appropriate computer aided molecular design (CAMD) algorithm (where fragments are used to design the homopolymer repeat unit structures) and a set of property prediction models. Group contribution based property prediction models[1-3] are sufficiently simple and reliable for use in CAMD, but due to the limited availability of experimental data for polymer properties, one cannot expect to fill out the entire group parameter table. At times there can also be a polymer repeat unit structure that cannot be totally represented

by groups for a specific group contribution method. This is overcome by a meso-scale approach, where an atom-connectivity index method is applied to determine the missing groups and their contributions automatically without the need for additional experimental data. This integration of group and atom-connectivity index based models is called group contribution plus (GC^+) model. Integration of Marrero/Gani (MG) group contribution methods and atom-connectivity index methods that has been proposed as MG-GC^+ models for polymers in Satyanarayana et al[3] is used in this work. The next step is to decide on the type of CAMD algorithm. Different types of CAMD algorithms have been proposed for homopolymer repeat unit structure identification for a specified set of property criteria[5-8]. A multi-step, multi-level hybrid CAMD algorithm by Harper and Gani[8] suppresses the combinatorial explosion problem; it is also capable of generating a large number of structures. This method has been modified for using the GC^+ models of Satyanarayana et al[7] and it is the one employed in this work.

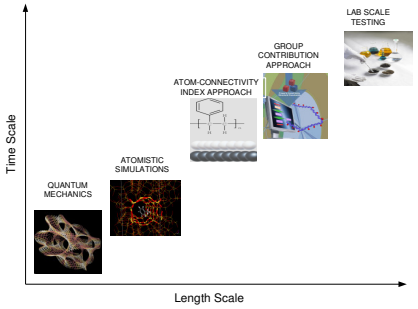


Fig 1. Multi scales for developing new polymer product

The identified basic repeat unit(s) are then used as input in molecular simulation algorithms (micro-scale) to generate a polymer chain with the desired number of repeat units and then to either predict or validate the required set of properties (some properties can vary with the varying chain length or varying number of repeat units). The accuracy of the results (e.g., thermodynamic, structural, conformational, etc.) obtained from molecular simulations depend on the efficiency of the simulation method used and the reliability of the force-field employed for describing the intra- and inter-atomic interactions. At low temperatures or for highly complicated polymer structures, these simulations are computationally intensive and time consuming. To overcome this, hierarchical simulation methodologies[9] are followed or multi-scale modeling algorithms are developed based on the definition of super-atoms. The latter approach is a method followed to reduce chemical complexity in a polymer chain without losing the chemical identity of the molecule. It is based on grouping a certain number of atoms or repeat units along the chain together into "super atoms" (or United Atoms) connected by effective bonds and governed by softer or smoother effective non-bonded interactions (shown in Stage 1 of figure 2). The resulting chain sequence is simpler and amenable to fast thermal equilibration and simulation through a brute-force application of the molecular dynamics (MD) method or by a Monte Carlo algorithm properly modified to account for the presence of the few different species along the chain.

By integrating the different scales, from macro- to micro-scale, the uniqueness of each method is retained while their limitations are avoided to some extent. Thus, this offers a promising avenue to the materials design problem, where one should propose structures satisfying a set of performance criteria. This methodology is illustrated for a case where materials performance is judged based on permeability (solubility and diffusivity) properties to small gaseous molecules.

2. Case Study

A case study using the detailed methodology (shown in figure 2) is performed. We consider here a problem referring to the design of a hermetic stopper for medical containers, which intend to secure against the entry of air or microorganisms and

maintains the safety and quality of the contents stored in their containers. Natural rubber (can cause latex allergies) and cork (has varying pore size) can also be used as sealants. It would be interesting to design a polymeric material for use as a hermetic stopper which can retain the advantages of existing stopper materials and overcome their disadvantages.

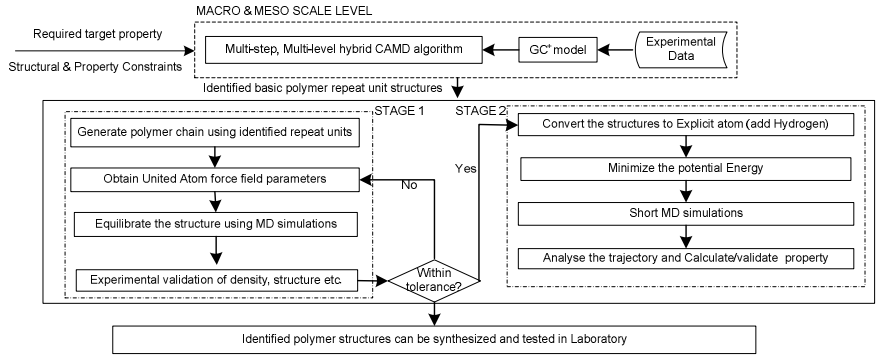


Fig 2. Schematic work-flow of the multiscale modeling.

2.1. Application of the macro-meso scale approach

*T*he basic homo-polymer repeat unit structures are identified in this approach. This requires specifying initially, a set of structural and property constraints.

2.1.1 Property Constraints (a) Glass transition temperature (T_g) – the polymeric stopper is required to be elastic over a wide range of temperature, and it is preferred to be made of a rubbery amorphous polymer so that its T_g is less than or equal to the room temperature (300 K)[1]. The upper limit for T_g can be relaxed to 350 K. By doing this, the effect of missing some candidates that have the required property (but are not identified due to inconsistency of the experimental data reported and also considered while developing the property prediction models) is minimized to some extent. (b) Dielectric constant (ε) – some contents stored in the medical containers need to be constantly agitated with electrical agitators before use. It is therefore required that the polymeric stopper should not conduct electricity and electrolyze the stored contents in the medical containers; this implies a low ε. Here, we impose a maximum constraint of 2.6 on the value of the ε. (c) Low permeation to moisture and air – the stopper should protect the stored contents from moisture and air, and to exhibit very low water (or moisture) absorption[1]. As the GC^+ model for predicting water absorptivity has not yet been developed, Van Krevelen's[1] group contribution model is used to predict this property (to verify if the identified repeat units have low water absorptivity) for the identified repeat units that satisfy both T_g and ε constraints. Moreover, low permeation to air implies low permeation to nitrogen (N_2) and oxygen (O_2). Since there is no direct group contribution method to study the solubility and diffusivity in the polymer, the micro-scale approach is also used for determining these properties (along with the arrangement of repeat units to form polymer chains).

2.1.2 Structural Constraints The focus here is to design a polymer with low permeability to moisture, therefore the polar and hydrogen bond forming functional groups should not be included in the basis set. Reason being, they often enhance chain-chain attractions, which tends to enhance crystallinity and tensile strength. Therefore, only the groups forming olefinic polymers are allowed. It is also known that the double bonds in polymers render them poor for weather resistance, thermal resistance, oxygen

resistance and ozone resistance. This applies also to olefinic groups with triple bonds. Taking this into account, olefinic groups with double and triple bonds are excluded from the basis set, which therefore included CH_3, CH_2, CH and C groups only. As the side chain forming groups (CH and C) are included in the basis set, in order to avoid too many side chains in the basic polymer repeat unit, the number of times each group can occur in the repeat unit is limited to 2 implying that the maximum number of groups that can occur with the combination of all the groups included in the basis set is 6. This is shown in eq. 1, where, n_j is the number of groups of type j.

$$2 \le \sum_j n_j \le 6 \qquad \text{(Eq. 1)}$$

As this case study aims at generating a basic homo-polymer repeat unit that has two free ends, the criterion of having two free ends in each generated structure is also specified (based on valency and octet rule) as shown in eq. 2, where n_j, v_j are the number and valency, respectively, of groups of type j and m =1, 0 or -1 for acyclic, monocyclic and bicyclic groups, respectively. As only olefinic groups are considered, the value of m=1.

$$\sum_j (2 - v_j) n_j = 2(m-1) \qquad \text{(Eq. 2)}$$

Table 1: Identified polymer repeat unit structures using CAMD algorithm

Property & Structural Constraints	Min	Max	
Glass Transition temperature (K)		350 K	
Dielectric Constant		2.6	
Building blocks : CH3, CH2, CH, C			
Number of groups	6	2	
Basic polymer repeat units	T_g (K)	ε	W_{abs} (g H_2O/g polymer)
-[-CH_2-$C(CH_3)_2$-]n-	*342.1*	*2.23*	*0.00003*
-[-$C(CH_3)((CH_2)_2CH_3)$-]n-	307.43	2.12	0.00003
-[-$CH(CH_3)$-]n-	302.51	2.48	0.00004
-[-CH_2-CH_2-]n-	168.72	2.47	0.00004
-[-CH_2-$CH(CH_3)$-]n-	257.91	2.37	0.00004
-[-$CH(CH(CH_3)_2)$-]n-	303.13	2.28	0.00004
-[-$CH(CH_3)$-$CH(CH_3)$-]n-	302.51	2.26	0.00004
-[-$CH((CH_2)_2CH_3)$-]n-	235.61	2.25	0.00004
-[-$CH(CH(CH_3)CH_2CH_3)$-]n-	275.75	2.14	0.00004
-[-CH_2-$CH(CH(CH_3)_2)$-]n-	276.24	2.17	0.00004
-[-$CH(CH(CH_3)((CH_2)_2CH_3)$-]n-	257.91	2.03	0.00004
-[-$CH((CH_2)_2$-$CH(CH_3)_2)$-]n-	258.32	2.05	0.00004

Structural and property constraints were ser as input to the CAMD algorithm. 112 candidates of basic repeat unit structures with isomers were designed and 106 candidates among them were selected based on the given constraints. Removing the isomers, 12 basic polymer repeat units that satisfy all constraints are listed in table 1. The repeat unit structure of polyisobutylene (PIB) is one of the identified repeat units from the CAMD algorithm. PIB is important elastomers with low T_g and high friction coefficient. From different literature sources[10-12] it is also reported that this polymer exhibits low permeability properties to small gas molecules in comparison to other elastomers due to the presence of the two bulky pendant methyl groups on each basic repeat unit structure. Permeability is described as the product of solubility and diffusivity. Lower solubility and diffusivity values always correspond to low permeability. As it is highly time consuming to study the entire identified basic polymer repeat unit structures at the micro-scale level (at this point of time), only PIB was addressed in the present case study

2.2. Application of the micro scale approach

Recently, Tsolou et al[9] undertook a systematic study of PIB system models ranging from 8 up to 80 repeat (monomer) units per chain at different temperatures (from 600 K to 300 K). The development and implementation in the molecular simulations of a new UA force field for PIB, which reproduces faithfully the conformational, structural and thermodynamic properties of the polymer is the key feature of this work. Guided by Tsolou et al[9], the first stage of the micro-scale approach of the present case study involved the execution of atomistic MD simulations with a C_{320} PIB system at several temperatures. The outcome of these simulations was an ensemble of well-relaxed configurations, which were used next as input to all-atom (AA) simulations (simulations

in which the atomistic detail is restored with the insertion of the missing hydrogen atoms along each PIB chain) for the accurate prediction of the solubility and diffusivity properties of PIB to small molecules such as O_2 and N_2 of interest here. Through such an approach, Tsolou et al[9], for example, obtained excellent predictions of O_2, N_2, He and Ar solubility in PIB as a function of temperature. In the present study, the approach was extended in order to calculate the diffusivity of O_2 and N_2 in PIB. Overall, the steps undertaken in the micro-sale approach are: (a) 3-chain PIB system with 80 monomers per chain was subjected to long atomistic MD simulations in the NPT ensemble using the parallel LAMMPS code and the new UA model introduced by Tsolou et al[9]. (b) Well-relaxed configurations from these UA simulations at four different temperatures (350, 450, 500 and 550 K) were selected and converted to AA representations by re-inserting the hydrogen atoms that had been neglected in the UA simulations. (c) After a short energy minimization of the chosen PIB system, five O_2 molecules were randomly inserted at different positions inside the matrix at the temperature of interest, and the energy of the system was minimized again. (d) Long MD runs (up to 3 ns depending on the temperature) in the NPT statistical ensemble were carried out, which allowed us to monitor the time evolution of the 5 oxygen molecules. In all simulations with the all-atom model, the very detailed and accurate COMPASS force field was used.

The diffusivity of oxygen molecules in the PIB matrix was studied in terms of the time evolution of their mean square displacement (MSD) in the course of the MD simulation. The diffusion coefficient, D, can then be readily computed from the linear part of the curve through the Einstein relation[13], where $\mathbf{r}_i$ denotes the position vector of the center of mass of the diffusant and angular brackets denote ensemble average over five (*i.e.*, N_a=5) O_2 molecules and over all possible time origins in the course of the trajectory.

$$D = \frac{1}{6N_a}\lim_{t\to\infty}\frac{d}{dt}\sum_{i=1}^{N_a}\left\langle\left[\mathbf{r}_i(t)-\mathbf{r}_i(0)\right]^2\right\rangle \qquad \text{(Eq. 3)}$$

Clearly, as the temperature decreases, the diffusivity of O_2 in PIB slows down considerably. By plotting log(D) versus 1/T (Fig. 3), a linear plot is obtained whose slope provides an estimate for the activation energy of the diffusive process defined as:

$$E_a^{app} = -\frac{R}{0.434}\frac{\partial(\log D)}{\partial\left(1/T\right)} \qquad \text{(Eq. 4)}$$

The value obtained from the data of Fig. 3 is 5.7kcal/mol. (the experimental value of 11.9 kcal/mol reported by Amerongen[14] from diffusivity data measured over a range of lower temperature values). In fact, Amerongen reports that the activation energy is not constant but increases with decreasing temperature. This indicates that the relatively higher value of E_a^{app} predicted from our micro-scale approach is quite reasonable.

2.3. Case study results and discussion

This case study has justified the success of the potential application of a multilevel approach to the design of a polymer with a set of pre-specified properties. Using the macro-meso-scale approach, a set of basic polymer repeat units promising to have the required target properties is obtained. Water absorptivity values for the identified basic polymer repeat units were comparatively lower as shown in table 1. On the other hand, the macro-meso-scale approach alone cannot decide about the optimal arrangement of the repeat units or the chain length; in the context of the proposed multi-level approach, this is the task of the micro-scale component. In the present case, for example, we chose to simulate PIB with C_{320} and compute its permeability properties to small molecules like O_2 and N_2. The simulation results from micro-scale work seem to verify the low permeability of O_2 in the designed polymer. Given that the molecular size of O_2 is smaller than N_2, we can safely conclude that PIB will also be characterized by low

permeability to N_2. Therefore, this polymer will have a low permeability to air. Other PIB systems with varying carbon number (C_{80}, C_{120} and C_{240}) will also be studied in the future to verify which system has lowest permeability to air. But the next step is to recommend the polyisobutylene with a chain length of 80 monomer repeat units for laboratory scale synthesis, to experimentally validate the properties claimed here.

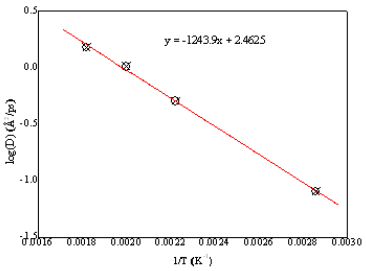


Fig 3. Log(D) versus 1/T for O_2 in PIB-C_{320} matrix

3. Conclusions

A multi-level approach for the design of polymers with a set of required properties, which can substantially help decrease the time for the development of new products have been presented in this paper. This approach relies first on the application of GC^+ property models in the CAMD algorithm for the identification of the polymer repeat unit structure properties; this greatly enhances the application range of problems that can be handled. Since the CAMD algorithm does give complete information on how the basic polymer repeat units should be arranged for the resulting polymeric structure to have the desired properties, a micro-level approach is also proposed to be employed. By providing the link to microstructure, it is capable of deriving reliable estimates of the desired properties based on the principles of quantum and statistical mechanics, and statistical thermodynamics. We have applied such a multilevel approach to a case study aimed at designing a polymer that can be used as a hermetic stopper. One of the basic polymers identified from the macro-meso-scale approach for such an application was polyisobutylene. In comparison to other elastomers and based on literature data, this polymer is characterized by a low glass transition temperature and markedly low permeability to O_2, N_2 and air. We concluded that PIB would be a candidate for constructing hermetic stoppers for use in medical applications.

References

1. Van Krevelen, D.W. (1990). Properties of Polymers, Amsterdam: Elsevier.
2. Bicerano, J. (2002). Prediction of Polymer Properties. New York: Marcel Dekker Inc.
3. Satyanarayana, K.C., Gani, R., Abildskov, J. (2007). Fluid Phase Equilibria, 261, 58.
4. Gani, R., Harper, P.M., Hostrup, M. (2005). Ind. Eng. Chem. Res., 44, 7262.
5. Manaras, C.D. (1996). Ind. Eng. Chem. Res.,35 (10), 3403. ‘
6. Venkatasubramanian, V., Chan, K., Caruthers, J. M.(1994). Computers &.Chemical Engineering, 18, 9, 833.
7. Satyanarayana, K.C, Abildskov, J. , Gani, R.(2008). Computers & Chemical Engineering. (In press).
8. Harper, P.M., Gani, R. (2000). Computers & Chemical Engineering, 24, 2-7, 677
9. Tsolou, G., Mavrantzas, V.G., Makrodimitri, Z.A., Economou, I.G., Gani, R. (2008). Macromolecules, 41 (16), 6228.
10. Kamio, K., Moorthi, K., Theodorou, D.N. (2007). Macromolecules 40, 710-722.
11. Abrams, C.F., Kremer, K. (2003). Macromolecules 36(1), 260-267.
12. Harper, P.M., Gani, R., Kolar, P., Ishikawa, T. (1999). Fluid Phase Equilibria 158-160, 337.
13. Karayiannis, N.C., Mavrantzas, V.G., Theodorou, D.N. (2004). Macromolecules 37, 2978.
14. Van Amerongen, G.J. (1950). Journal of polymer science, 5 (3), 307.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Predicting refinery energy losses due to fouling in heat exchangers

Francesco Coletti and Sandro Macchietto

Department of Chemical Engineering - Imperial College London
South Kensington Campus
London SW2 2AZ (UK)

Abstract

Energy efficiency is paramount in an oil refinery and heat integration is critical, especially in the energy-intensive atmospheric distillation unit. Thermal and hydraulic performance of each exchanger in the network used to pre-heat the crude is greatly reduced by the progressive deposition of unwanted material on the thermal exchange surfaces. Here, a detailed mathematical model for a shell and tube heat exchanger undergoing crude oil fouling is used to predict local and average fouling rates and identify critical performance areas in the exchanger. The ease of configuring the model for industrial applications is detailed. With reference to an industrial unit, it is shown that unit averages mask very different local behavior, and that it is possible to achieve precise calculation of heat duty loss caused by fouling over time.

Keywords: Fouling, crude oil, modeling, refinery, energy, heat exchanger.

1. Introduction

Fouling is a major cause of energy losses in oil refineries. In particular, fouling in the crude pre-heat train of the atmospheric distillation unit is source of extra costs and greenhouse gas emissions, disruptions of operation and hazards. Losses due to fouling in the US alone were estimated to be in the order of $1.2 billion per annum in 1992 (ESDU 2000), or $1.8 billion in 2008 simply adjusting for inflation. The pre-heat train (PHT) is a network of shell and tube heat exchangers (of the type in Figure 1) that is used to recover heat from the hot products of the distillation column. This allows the refinery to decrease energy requirements at the furnace which brings the oil to the final temperature required for the fractionation. Although all the units in the PHT experience fouling, it is particularly severe in the hot-end, where the temperature is high enough to activate chemical reactions between crude components. Over months, a deposit layer builds up on exchange surfaces with the two-pronged effect of increasing the pressure drops across the unit and decreasing the overall heat transfer coefficient. As a result, more pumping power and more energy at the furnace are required, adding extra costs to operations. Moreover, large production losses arise when the furnace firing limit is reached as it is necessary to reduce throughput to maintain a constant inlet temperature at the column. Further large losses arise from periodic cleaning shut-downs.

The oil industry has faced this problem for decades and the efficiency of the pre-heat train is a primary concern for refinery operators. Operations are scheduled so as to minimize the energy lost due to fouling. The heat duty is constantly calculated (and fouling monitored) from plant data. However, tools to help refiners to analyze plant data

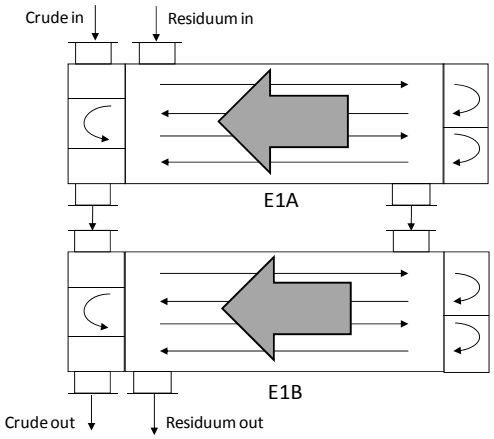


Figure 1 Heat exchanger geometry considered in the case study. The unit comprises two shells with four pass each. Crude is flowing on the tube side, residuum on the shell side.

and predict fouling behavior with accuracy are still lacking. Models based on the threshold concept (Ebert and Panchal 1995) have been used to include crude oil fouling when designing or retrofitting single units (Polley *et al.* 2002), whole PHTs (Nasr and Givi 2006; Yeap *et al.* 2004) and to assist in cleaning scheduling to improve the network operability and mitigate maintenance-related costs (Ishiyama *et al.* 2007; Smaili *et al.* 2001; Wilson and Vassiliadis 1997). All above use lumped models based on a macroscopic thermal balance and quasi-steady state approximations. They do not take into account local variation in conditions along the units (known to be significant) and dynamics. In this paper a distributed, dynamic first principle model (Coletti and Macchietto 2009) is used to predict fouling rates and identify critical performance areas in the heat exchanger. Moreover, it is shown that use of the model can help solve the difficult problem of calculating the effects of fouling on the loss of heat duty for exchangers with highly varying inlet conditions.

2. Approach

The conceptual framework developed to simulate refinery data is as follows:

1. Develop a high fidelity, first principle model
2. Analyze refinery data to identify unreliable data
3. Estimate model parameters from reliable data and validate model
4. Use validated model to predict behavior as function of process conditions

The model used for a tubular heat exchanger undergoing crude oil fouling in step 1 has been presented in (Coletti and Macchietto 2008; 2009). Table 1 summarizes the main model equations in each domain considered: tube side ($\Omega_{T,n}$); deposit layer ($\Omega_{L,n}$); tube wall ($\Omega_{w,n}$) and shell side (Ω_S). The model is dynamic and distributed along the length of the heat exchanger and captures the time- and space- varying effects of process variables (geometry, temperature, velocity and crude oil properties) on deposition mechanisms. For this purpose, the classic Ebert-Panchal model for chemical reaction fouling was used locally. The model accounts for important interactions between the fouling layer and fluid flow. Step 2 (refinery data analysis) and step 3 (parameter estimation) above are given in a previous paper (Coletti and Macchietto 2009) for the specific industrial application described in the following, leading to excellent match to

Table 1 Summary of main model equations. Ω is the domain considered (see text).

Ω	Equations
Ω_T	$\frac{\partial}{\partial t}\left(A_{flow}\rho e\right) = -\frac{\partial}{\partial z}\left(A_{flow}\rho u e\right) + \frac{\partial}{\partial z}\left[A_{flow}k\frac{\partial T}{\partial z}\right] + Ph_t\left(T_w - T\right) \quad \frac{dR_{f,n}}{dt} = \alpha\,\mathrm{Re}^{-0.66}\,\mathrm{Pr}^{-0.33}\exp\left(\frac{-E}{RT_{f,n}}\right) - \gamma\tau$
	$T_{f,n}(z) = T_n(z) + 0.55\ \left(T_{L,n}(z, R_{flow}) - T_n(z)\right);\qquad \tau = C_f\left(\frac{\rho u^2}{2}\right);\qquad \frac{d\delta_n(z)}{dt} = k_L\frac{dR_{f,n}(z)}{dt}$
Ω_L	$\frac{\rho_L c_{p,L}}{k_L}\frac{\partial T_{L,n}}{\partial t} = \frac{1}{r}\frac{\partial T_{L,n}}{\partial r} + r\frac{\partial^2 T_{L,n}}{\partial r^2};\qquad q_{L,n} = -k_L\frac{\partial T_{L,n}}{\partial r}$
Ω_w	$\frac{\rho_w c_{pw}}{k_w}\frac{\partial T_{w,n}}{\partial t} = \frac{1}{r}\frac{\partial T_{w,n}}{\partial r} + r\frac{\partial^2 T_{w,n}}{\partial r^2};\qquad q_{w,n} = -k_w\frac{\partial T_{w,n}}{\partial r}$
Ω_S	$\frac{\partial}{\partial t}\left(\rho_S c_{p,S} T_S\right) = \frac{\partial}{\partial z}\left(\rho_S c_{p,S} T_S u_S\right) + \frac{\partial}{\partial z}\left(k_S\frac{\partial T_S}{\partial z}\right) + \frac{1}{A_S}\sum_{i=1}^{N_p} P_{S,i}h_S\left(T_S - T_{w,i}\right)$

actual plant data. The model could predict the outlet temperatures of a shell and tube heat exchanger undergoing crude oil fouling, as a function of process conditions and time, within ±1°C.

2.1. Model implementation

The model partial, differential and algebraic equations are solved using the gPROMS™ modelling environment exploiting its hierarchical structure. A single tube model is defined once and re-used as many times as needed within desired multi-pass configurations to describe a full single shell, tubular heat exchanger unit (Figure 2). Such exchanger model is then itself defined as an object that can be easily replicated and flexibly interconnected with other units in a flowsheet (e.g. double shell unit in Figure 1). This feature is particularly handy considering that pre-heat trains can comprise up to 60 heat exchangers interconnected in a single network (Panchal and Huangfu 2000). Configuring a model requires geometric exchanger data and thermo-physical oil properties. The exchanger geometries were inputted directly from standard company data sheets, in an Excel spreadsheet, via gPROMS' built-in goEXCEL interface. This feature makes it extremely simple to set-up the model for a unit under investigation. Various thermodynamic packages accessed via a CAPE-OPEN interface allow defining oils in terms of either their gas chromatographic analysis or true boiling point curve. The thermodynamic package can then be used to calculate, at each point in time and axial length, physical properties such as density, viscosity, heat capacity and thermal conductivity of the fluids flowing in both shell and tube sides.

2.2. Case Study

The model was applied to simulate an industrial unit in an ExxonMobil pre-heat train. The unit comprises two identical shells with 4 tube pass per shell, with crude in the tube-side and residuum from the distillation column in the shell side (Figure 1). For this case study, the physical proprieties of tube- and shell-side oils where not calculated through the thermodynamic package but were assumed constant and values estimated from a short set of plant data (inlet and outlet temperatures, and flowrates). Operating conditions over 150 days after a mechanical cleaning were then simulated, with inlet temperatures and flow rates from measured plant data. The main model parameters used in the simulations are summarized in Table 2.

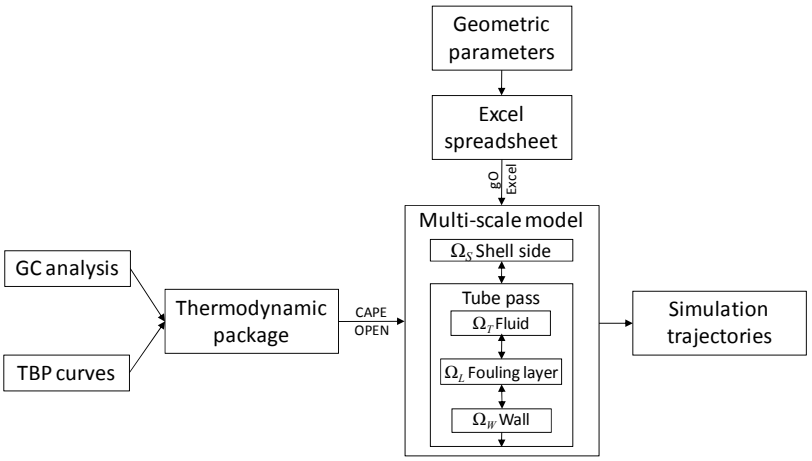


Figure 2 Implementation of the model in process simulator. Physical proprieties are calculated through a thermodynamic package which is interfaced via CAPE-OPEN. Necessary geometries are set interfacing the process simulator with an Excel spreadsheet.

3. Results

Model simulations allow tracking the trajectory of key variables over time and space. Figure 3.a shows the averaged (over length) fouling resistance, R_f, in each pass in each shell in the unit, for 150 days. The fouling resistance calculated in shell E1B is larger than that in E1A. This is expected as temperatures in the latter are lower than those in the former. An overall fouling resistance for the unit (the arithmetic average between the two shells) is plotted in the same figure as a solid line. A large difference is evident between the shells, and the unit-average resistance seriously underestimates local fouling effects (therefore occlusion and pressure loss) at the hot end of tubes. This can be observed by tracking the local thickness of the fouling layer. In Figure 3.b the foulant thickness, δ, is reported at the inlet and outlet of the unit, over time. Again, the difference in thickness after 150 days is evident. The qualitative difference in fouling behavior is also noted: at the tube inlet, foulant thickness reaches an asymptotic value whereas at the tube outlet the behavior can be assimilated to a linear fouling rate.

Table 2 Summary of main model parameters.

Parameter		Tube	Shell	
			E1A	E1B
Density, ρ	kg m^{-3}	817	880	
Viscosity, μ	Pa s	2.002×10^{-3}	-	
Heat capacity, C_p	J K^{-1} kg^{-1}	2184	2462	
Thermal conductivity, k	W m^{-1} K^{-1}	0.125	0.180	
Tube length, L	m	6.1	-	
Tube diameters, d_o/ d_i	mm	25.40/19.86	-	
Shell diameter, D_S	m		1.397	
Activation energy, E	J mol^{-1}		28931	28626
Deposition constant, α	m^{2}K kW^{-1} h^{-1}		1.689×10^{-3}	1.785×10^{-3}
Suppression constant, γ	m^{2}KW^{-1}s^{-1}Pa^{-1}		6.8×10^{-11}	2.1×10^{-11}
Pass number			4	4
Tube count, N_T			880	880

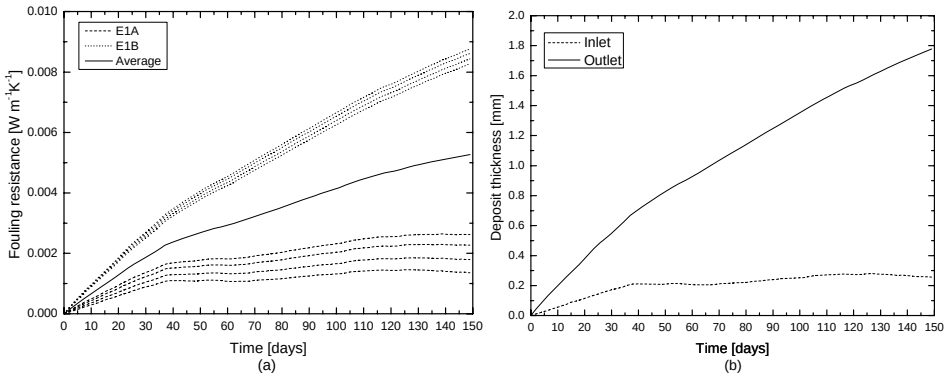


Figure 3 Fouling resistance for each tube side pass in the two shells and average value for the unit (a) and deposit thickness at inlet and outlet (b).

By repeating a simulation without fouling in the model (R_f=0) it is possible to calculate, for the given process conditions, the difference between the heat exchanged in clean and fouled conditions (Figure 4.a). At time zero (clean exchanger), there is no difference between the two scenarios. However, after few days of operations the difference becomes large. The same figure also reports the heat duty loss due to fouling calculated as the difference between the heat duties in the two scenarios considered. The maximum difference occurs after 135 days when the calculated loss in heat duty is over 2.5 MW. Losses at any one time depend on (variable) inlet temperatures and flowrates, which makes it difficult to isolate the fouling effects. The shaded area below the heat loss curve represents the energy actually lost because of the thermal inefficiencies caused by fouling and is given by the integral over time of the difference in heat duty clean and fouled. The underlying energy loss trend can be noted from a straight line linear fitting, as indicated at the bottom of Figure 4.a, or a rolling average. Figure 4.b shows the integral of the heat duty loss over time representing the cumulative energy loss. After 150 days the cumulative energy loss due to fouling is nearly 140 MJ, equivalent to a cost of extra fuel burnt in the furnace of over US$ 30,000 (considering a cost of the fuel at 27 $/KWh and 90% furnace efficiency).

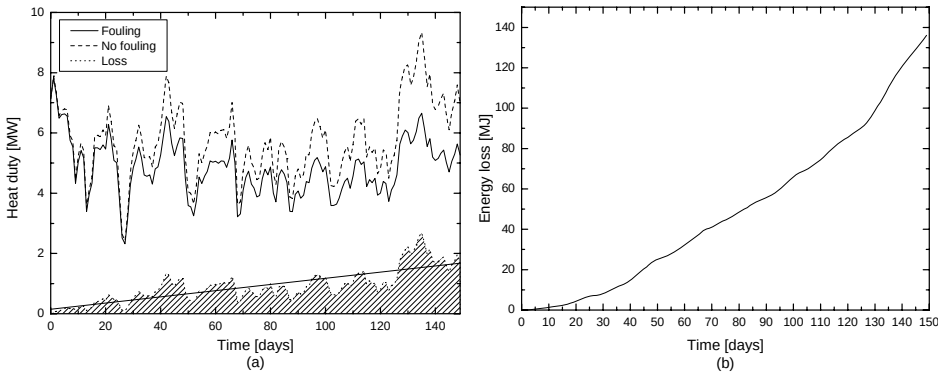


Figure 4 Heat duty from model simulations in case of fouling and no fouling over time (a). Dashed line in the bottom indicates the difference between the two. The shaded area is the energy lost due to fouling, and the straight line is obtained by linear fitting. Cumulative energy loss is given over time in (b).

4. Conclusions

A novel dynamic model for a multi-pass shell and tube heat exchanger was developed for use in refinery preheat trains, where fouling deposition results in severe energy inefficiency. It considers several aspects of the fouling phenomenon such as the local interactions between the onset and growth of the fouling layer and the crude oil flowing inside the tubes. The model allowed carrying out a detailed analysis of the related spatial and temporal evolution of thermal and hydraulic performance inside an industrial heat exchanger unit. The model was set up with the industrial unit configuration, tuned to local conditions for a clean unit, and then driven according to actual measurement of the input variables from the plant. Results showed the difference between the heat duty in clean and fouled conditions as a function of operating conditions, allowed a precise estimation of energy losses (hence extra energy and emissions duties at the furnace) and highlighted the very different performance in different parts of the exchangers.

A variety of applications are envisaged, including for improved operations (fouling monitoring of individual exchangers, energy loss calculation at PHT network scale, control mitigation applications, prediction of time to clean) and for exchanger design and retrofit. Further model improvements will include other complex phenomena such as ageing of the deposits and the effect of oil composition on physical properties.

5. Acknowledgements

The authors gratefully acknowledge ExxonMobil for the data provided.

References

F. Coletti and S. Macchietto (2008). Minimising efficiency losses in oil refineries: a heat exchanger fouling model. Sustainable Energy UK: Meeting the science and engineering challenge", St Anne's College, Oxford.

F. Coletti and S. Macchietto (2009). "A dynamic, distributed model of shell and tube heat exchangers undergoing crude oil fouling." Submitted for publication.

W. A. Ebert and C. B. Panchal (1995). Analysis of Exxon crude-oil-slip stream coking data. Fouling Mitigation of Industrial Heat-Exchange Equipment, San Luis Obispo, California (USA), Begell House inc.

ESDU (2000). Heat exchanger fouling in the pre-heat train of a crude oil distillation unit, Data Item 00016. London, ESDU International plc.

E. M. Ishiyama, W. R. Paterson and D. I. Wilson (2007). The effect of fouling on heat transfer, pressure drop and throughput in refinery preheat trains: optimisation of cleaning schedules. Fouling and Cleaning of Heat Exchangers VII, Tomar, Portugal.

M. R. J. Nasr and M. M. Givi (2006). "Modeling of crude oil fouling in preheat exchangers of refinery distillation units." Applied thermal engineering 26(14-15): 1572-1577.

C. B. Panchal and E.-P. Huangfu (2000). "Effects of Mitigating Fouling on the Energy Efficiency of Crude-Oil Distillation." Heat transfer engineering 21(3): 3-9.

G. T. Polley, D. I. Wilson, B. L. Yeap and S. J. Pugh (2002). "Use of crude oil fouling threshold data in heat exchanger design." Applied Thermal Engineering 22(7): 763-776.

F. Smaili, V. S. Vassiliadis and D. I. Wilson (2001). "Mitigation of fouling in refinery heat exchanger networks by optimal management of cleaning." Energy & fuels 15(5): 1038-1056.

D. I. Wilson and V. S. Vassiliadis (1997). Mitigation of refinery fouling by managment of cleaning. Understanding Heat Exchanger Fouling and its Mitigation, Lucca, Italy, Begell House.

B. L. Yeap, D. I. Wilson, G. T. Polley and S. J. Pugh (2004). "Mitigation of crude oil refinery heat exchanger fouling through retrofits based on thermo-hydraulic fouling models." Chemical engineering research & design 82(1): 53-71.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Numerical Simulation of Coal Boiler at Electric Thermal Plants Using Computational Fluid Dynamics

Jairo Z. Souza[a], Leonardo P. Rangel[a], Henrique C. Monteiro[a], Marcelo Bzuneck[b], Luiz Felippe[b], Artur R. F. Ellwanger[b]

[a]*ESSS – Engineering Simulation and Scientific Software LTDA.*
Rod. SC 401, km 1, Parq. Tec. Alfa, Ed. CELTA, Sl. 4.01, CEP 88030-000, Florianópolis - SC, Brazil
[b]*Tractebel Energia S.A – Usina Termelétrica Jorge Lacerda C (UTLC)*
Av. Paulo Santos Mello, s/n, CEP 88745-000, Capivari de Baixo – SC, Brazil

Abstract

One of great challenges found in electric thermal plants boiler operation is to avoid the erosion problem on water wall ducts. This problem is generally caused by three different sources: flame misalignment, thermal attack and erosion due to the contact with chemicals. This work focus on investigate the effect of flame misalignment. This problem can be caused by fluid dynamics factors due the burner geometry, which can be very complex. The key part of the burner the swirler, whist working generates a spiral flow starting in the burner and following towards the rear of the boiler. The numerical simulation of this system has been developed using the commercial software of Computational Fluid Dynamics (ANSYS CFX). It has been chosen based on its robustness in mesh generation and to solver conservation equations by finite volume method. The mathematical model has been developed by two approaches: homogeneous (only air has been simulated) and heterogeneous (the coal particle trajectories inside the boiler has been calculated in the Euler-Lagrange approach). The results presented velocity profiles, pressure profiles, streamlines and other data that is helpful to understand the fluid flow phenomena inside the equipment. The coal particles trajectories simulated allow identifying where particles collide with the boiler wall. All results are in very good concordance with field observations according to engineers at UTLC from Tractebel Energia S.A.

Keywords: CFD, Boiler Simulation, Multiphase Flow

1. Introduction

Keep high efficiency of boilers in steam generation process is a great challenge for thermal power plants. The efficiency of boilers can be analyzed by reducing maintenance stops due to the expensive coasts in this equipment.
Tractebel asked ESSS to guide engineers to understand the fluid flow within the boiler using a numerical simulation. The main objective of this study is to analyze the possible causes of excessive erosion in a specific boiler wall region: the region situated near of the right burner in the third row of burners, as showed in the Figure 1.

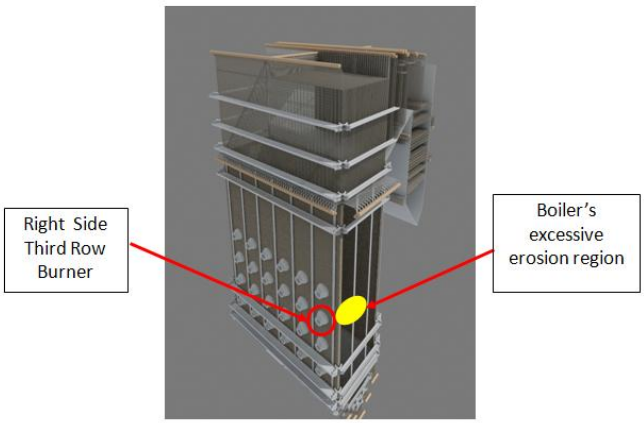


Figure 1 – Boiler region affected by excessive erosion

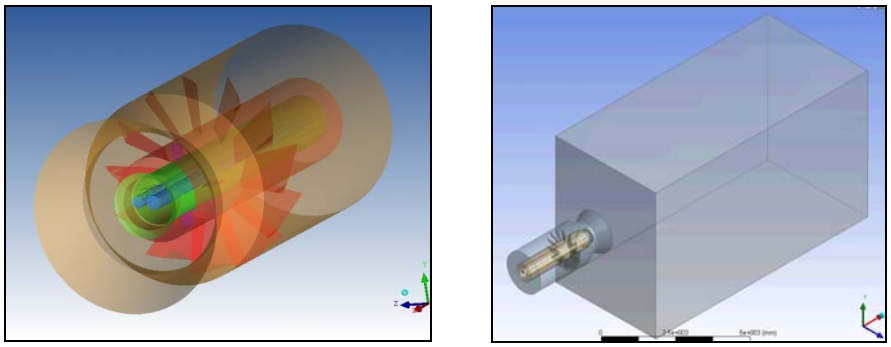
It has been previously identified that a possible origin for the excessive erosion could be due to the non-burned particles flow as a result of a misaligned flame.
The swirler promotes a spiral flow inside the equipment and is developed to stabilize the flame and the air-coal mixture flow within the burner.
The strategy to visualize the flow has been to develop a detailed air and coal mixture within the region allowing the detection of recirculation zones. This can be observed with velocity profiles and particle coal pathlines. It has been assumed that the erosion problem is associated to the particles path lines driven to the collision against the boiler wall.
This work has been developed with an adiabatic flow model. Thermal effects have not been considered since the hot air entered and no combustion model has been applied. For an additional study, these factors must be considerer to quantify the wall erosion rates, but the computational effort increases significantly. According to what has been presented, these results are qualitative, but enough to understand the boiler internal flow and wall erosion causes.

2. Methodology

The proposed boiler-burner numerical simulation has been developed using Computational Fluid Dynamics (CFD) with the commercial CFD software ANSYS CFX®. A CFD work is made of five stages: geometry, mesh, pre-processing, solver and pos-processing. These stages are detailed below.

2.1. Geometry and Computational Mesh

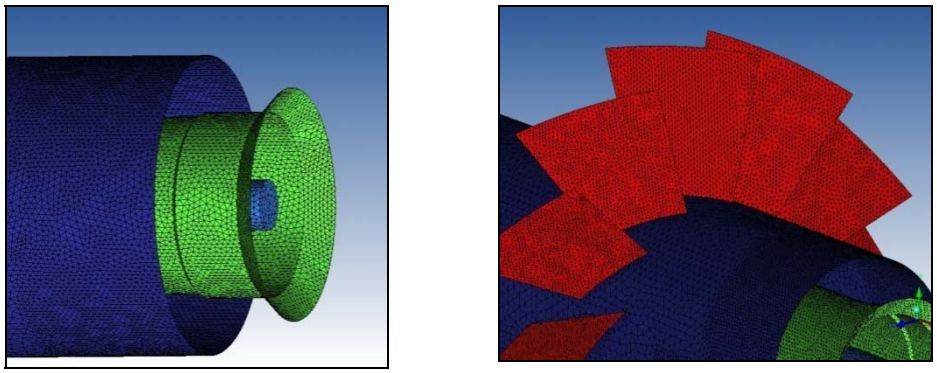
The geometry is developed using a CAD model that must specify all important details that are necessary to the correct equipment description. For this work, the initial data available has been only 2D plants and photographs.
The geometry has been build using ANSYS® Design Modeler. The final burner CAD model is showed in Figure 2(a), while the Figure 2(b) shows the computational domain.
The computational mesh is the representation of finite volumetric elements which are applied the conservation equations. In this stage, has been used the software ANSYS® ICEM CFD. To keep the correct geometry in the discretized domain, the elements sizes are established in accordance to the local length of each equipment part.



(a) (b)

Figure 2 – Burner (a) and boiler (b) geometries

Equipment computational mesh examples are shown in Figures 3, where there are two burner parts: deflector (a) and swirler (b). The final computational mesh has 2,2 millions of nodes. This size characterized a high detailed mesh which requires high computational effort.



(a) (b)

Figure 3 – Computational mesh in deflector (a) and swirler (b) regions

2.2. Mathematical Model

The main assumptions in the boiler numerical simulation are (i) steady state, (ii) isothermal, (iii) incompressible, (iv) turbulent, (v) lower solid volumetric fraction within the equipment (Euler-Lagrange).

With assumptions above, the mathematical model consists by mass continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla . \left(\rho . \vec{V} \right) = 0 \tag{1}$$

and the Navier-Stokes momentum equations

$$\rho \frac{D\vec{V}}{Dt} = \rho g - \nabla p + \mu \nabla^2 \vec{V} \tag{2}$$

For the turbulence model, it has been assumed the k-□ model, the most utilized in industrial equipments with rotation. For the multiphase model, the Eulerian-Lagrangian approach calculates the particle individual track, through the integration of Ordinary Differential Equation (ODE) for particle velocity and position. The particle position can

be represented as in the Figure 4, which integrated, results the particle position within the equipment.

$$\frac{d\vec{r}_P}{dt} = \vec{U}_P \tag{3}$$

Looking Figure 4, a particle with mass m_P moves with velocity $\vec{U}_P$ and it is influenced by $\vec{F}$ force generated by continuous phase, which has velocity $\vec{U}_F$.

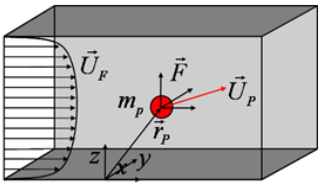


Figure 4 – Particle schematic representation

The particle velocity is obtained through the particle momentum balance, as shown equation 4.

$$m_P \frac{d\vec{U}_P}{dt} = \vec{F} \tag{4}$$

The $\vec{F}$ force acting over the particle has its origin in different phenomena, which depends of particle physical properties, continuous phase, relative velocity, gravity, turbulence, etc. In this model, it has been considered the drag force, the weight and the turbulence influence on particles tracks, through the turbulent dispersion force. To enclosure the mathematical model, the boundary conditions have been obtained through operation data at thermal unity, where has been measured the burner air flow values. The wall has considered as “no-slip” condition.

2.3. Solver

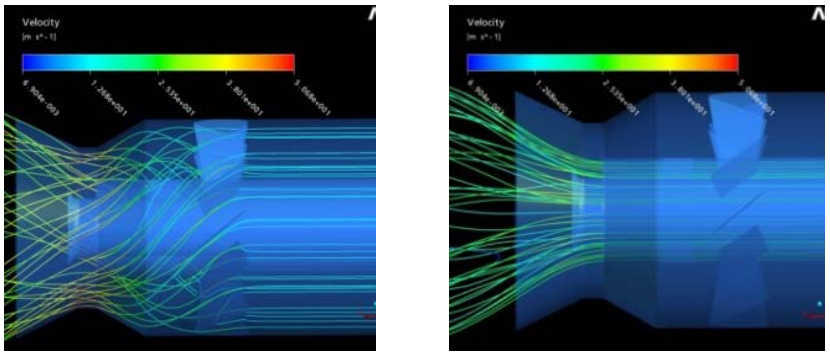
ANSYS® CFX-Solver has been applied to solve all conservation equations. It implies the method of finite volume based on finite elements. The convergence criteria has been the RMS (Root Mean Square), which has been monitored up to the residual reaches the acceptable criteria. Several monitor points throughout the equipment have been indicating velocity and pressure values aiming to guarantee the convergence stability. The third convergence criterion is the closure of mass imbalance.

3. Discussion of Results

The CFD results allow detailed flow visualization, as recirculation zones, high velocities, etc. These results are generally presented in profiles planes, streamlines or vectors, showing the flow directions and preferential paths. In this work, it has focused the velocity field and particle track. The results allowed identify the equipment erosion tendencies. The numerical solution pos-processing has been realized with ANSYS® CFX-POST.

3.1. Burner and Boiler Flow

The burner is the main boiler component and focused in this work. In it that is generated the air flow in a spiral format. Figure 5 shows the streamlines in this part of equipment. It was possible capture this phenomenon through the turbulence model.



(a) (b)

Figure 5 – Burner Geometry effects throughout the primary and secondary inlets

In the Figure 5 (a), the spiral flow starts on the blades of the swirler, promoting a spin. As mentioned before, this swirl stabilizes the flame and homogenizes the air-coal flow. The mass flow in this secondary air pipe is higher than the primary air-flow; therefore this spiral flow is predominant. Figure 5 (b) shows the streamlines of primary air pipe in contact with deflector. The radial opening homogenizes this flow with secondary air. In this duct that coal is transported. Deflection and swirlers angles are geometric parameters which can be tested with CFD, supplying tendencies of optimal configuration. The flame destabilization inside the boiler may be caused by two internal pipes in the middle of burner. These are used for flame ignition. Figures 6 and 7 show that these pipes promote a non-symmetric region.

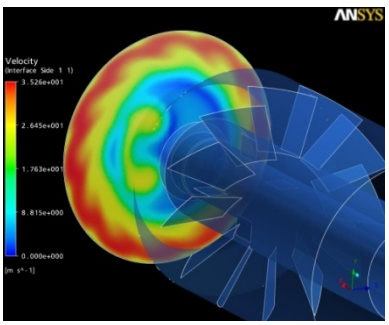


Figure 6 – Velocity Profile in the Interface between boiler and burner

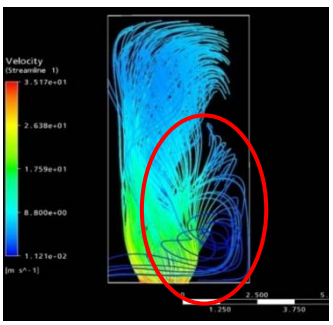


Figure 7 – Streamlines Inside the Boiler

The external annular region present an axial symmetry, however, the central region is non-symmetric. This shows that the internal pipes have significant influence in the flow. Looking towards the air flow within the boiler in a superior view in the Figure 7, it can be observed the spiral flow tendency, follow throughout the boiler. However, there are some streamlines that disengages of principal flow, going to the wall. The streamlines overtopped by the red circle can carry coal particles not burned into the wall, resulting in the excessive erosion. The causes of this disengages streamlines are unknown and it is considered that internal ignition pipes have influence on it.

3.2. Multiphase Analysis

The multiphase approach used in this work has been the Eulerian-Lagrangian, with one way coupling between the phases. With this approach only the air flow influences in the

solid flow. This simplification is possible due the low volumetric solid concentration, as explained in the mathematical model section. Figure 8 shows coal particle trajectories and collision against the boiler wall. These are few trajectories (the most part follows towards the back of the boiler), but these particles can attack the wall with erosion.

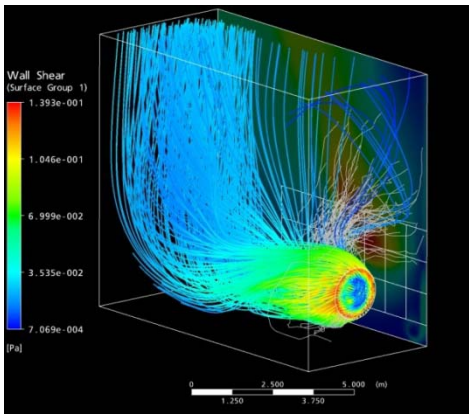


Figure 8 – Coal Particles Tracks

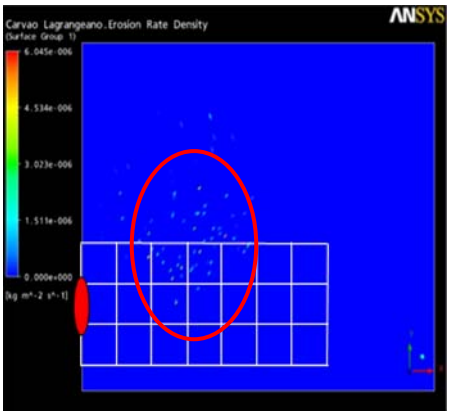


Figure 9 – Boiler Wall Erosion

Figure 9 shows wall contact of particles. The red ellipse in left side is the burner position. The white lines are only for reference. The erosion showed by colored points within the red circle are calculated through the material properties (coal–steal), velocity and collision angle between the coal particles and wall. The reds points are the more intense collision points. The region where the coal particles collide against the wall is in very good concordance with the field observations from engineers of Tractebel. The showed region is about 3.5 m of burner offset.

4. Conclusions

Erosion problem on water wall ducts of coal boilers can be analyzed by numerical simulation using CFD. In this work, it has been shown how burner swirl and deflection affect the particles and air path lines by calculating the turbulent multiphase flow with a commercial CFD tool. Velocity profile and streamlines showed the flow tendencies within the boiler with particles collision at 3.5 m of burner offset, the exact position of the highest detection of erosion. This is caused because there is no symmetry of two internal ignition pipes within the burner. This fact generates a misalignment in the spiral flow, what may promote the flame destabilization as well. A quantitative erosion analysis has not been performed at this stage, but it can be developed with addition of thermal and chemical effects like combustion. The qualitative analysis provided good results is real agreement with the problem at Tractebel Energia Power Plant validating the proposed model of this work.

References

Bird, R. B. ; Stewart, W. E. ; Linghtfoot, E. N. (1960) Transport Phenomena. NY: John Wiley.

Foust, A. S. (1982) Princípios das operações unitárias. 2. ed. Rio de Janeiro: Guanabara Dois.

Maliska, C. R. (2004) Transferência de Calor e Mecânica dos Fluidos Computacional. 2° Edição. Editora LTC.

Miller, A. Gidaspow D. (1992) Dense, vertical gas-solid flow in a pipe. AIChE Journal, v.38, n.11, p.1801-1815.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimization of Transformations for Convex Relaxations of MINLP Problems Containing Signomial Functions

Andreas Lundell and Tapio Westerlund

Process Design & Systems Engineering
Åbo Akademi University, Biskopsgatan 8, 20500 Turku, Finland

Abstract

In this paper, a method for determining an optimized set of transformations for signomial functions in a nonconvex mixed integer nonlinear programming (MINLP) problem is described. Through the proposed mixed integer linear programming (MILP) problem formulation, a set of single-variable transformations is obtained. By varying the parameters in the MILP problem, different sets of transformations are obtained. Using these transformations and some approximation techniques, a nonconvex MINLP problem can be transformed into a convex overestimated form. What transformations are used have a direct effect on the combinatorial complexity and approximation quality of these problems, so it is of great importance to find the best possible transformations. Variants of the method have previously been presented in Lundell et al. (2007) and Lundell and Westerlund (2008). Here, the scope of the procedure is extended to also allow for minimization of the number of required transformation variables, as well as, favor transformations with better numerical properties. These improvements can have a significant impact on the computational effort needed when solving the transformed MINLP problems.

Keywords: deterministic global optimization, signomial functions, MINLP problems, convex relaxations.

1. Introduction

MINLP problems containing signomial terms which can be transformed to a relaxed convex form using the transformation techniques in this paper are of the form

$$\begin{aligned} &\text{minimize} && f(\mathbf{x}), \qquad \mathbf{x} = (x_1, x_2, \ldots, x_I), \\ &\text{subject to} && \mathbf{Ax} = \mathbf{a}, \quad \mathbf{Bx} \le \mathbf{b}, \quad \mathbf{g(x)} \le \mathbf{0}, \\ & && \mathbf{q(x)} + \boldsymbol{\sigma}\mathbf{(x)} \le \mathbf{0}, \end{aligned} \tag{1}$$

where the functions f, g, and q are convex, σ is signomial and the variables x_i are positive reals or integers. A signomial function is defined as the sum of terms consisting of products of power functions multiplied with real constants, i.e.,

$$\sigma(\mathbf{x}) = \sum_{j=1}^{J} c_j \prod_{i=1}^{I} x_i^{p_{ji}}.$$

Signomials appear in many different types of optimization problems. In the special case when the constants c_j are positive, the function is called a posynomial. Posynomials can be found for example in geometric programming, where often exponential transformations are used to solve the optimization problem. In the more general case, when a MINLP problem is of the form in Eq. (1), i.e., contains linear, convex and signomial constraints, the exponential transformation is no longer suitable. Such problems are difficult to solve, since simple signomial terms such as bilinear and trilinear terms can make the problem nonconvex, possibly with several locally optimal solutions.

2. The transformation procedure

The transformation procedure consists of two steps. In the first step, the signomial terms are convexified using single variable transformations on the individual variables. This does, however, not transform the MINLP problem to a convex form, since nonlinear equality constraints defining the inverse transformations must also be added. Therefore, in the second step, the inverse transformations are approximated by piecewise linear functions (PLFs) in a way so that the feasible region of the original MINLP problem is overestimated by a convex relaxed region. The whole procedure has been described previously in several papers, e.g., in Lundell et al. (2007) and Westerlund (2005). The scope of this paper is the first step of the transformation procedure only, i.e., the transformations of the signomial terms themselves.

Depending on the sign of the signomial terms, different transformations can be used; for positive terms the exponential and the power transformations (ET and PTs) are valid. The PTs can further be classified into two different types – the positive and negative power transformations (PPT and NPT respectively). Negative signomial terms are transformed using PTs only. In the ET the transformations $x_i = \exp(X_i)$, and in the PTs the transformations $x_i = X_i^{p_i Q_i}$ are respectively applied to the individual variables. Different values for the transformation power Q_i are used in the different PTs. Note also that whenever the transformation power Q_i is equal to one, no transformation occurs.

2.1. Convexifying positive signomial terms

When convexifying positive signomial terms, all variables x_i with positive powers must be transformed by either an ET or a PT. A general signomial term is then transformed according to

$$c\prod_i x_i^{p_i} \rightarrow c\prod_{i:p_i>0} e^{p_i X_i} \cdot \prod_{i:p_i<0} x_i^{p_i} \quad \text{(ET)} \quad \text{and} \quad c\prod_i x_i^{p_i} \rightarrow c\prod_{i:p_i>0} X_i^{p_i Q_i} \cdot \prod_{i:p_i<0} x_i^{p_i} \quad \text{(PTs).}$$

For the ET no additional conditions are required, but for the NPT all transformation powers Q_i must be negative. In the PPT the requirements are that one transformation power is positive (and larger than or equal to one), the rest negative, and that the sum of the powers is larger than or equal to one. For details see Lundell et al. (2007).

2.2. Convexifying negative signomial terms

For negative signomial terms, PTs with $Q_i < 0$ are used for variables with negative powers and $Q_i > 0$ for variables with positive powers. Thus, all powers $p_i Q_i$ in the transformed term are positive. Additionally, the sum of the powers must be less than or equal to one. Therefore, a negative signomial term is transformed according to

$$c\prod_i x_i^{p_i} \rightarrow c\prod_{i:p_i\neq 0} X_i^{p_i Q_i}.$$

3. The optimization method for the transformations

In this section, the method for optimizing the set of transformations is presented. This method is based on formulating a MILP problem, where different strategy parameters determine what transformations are favored for use on the individual variables in the signomial terms. Since the same variables can occur in several signomial terms in the problem, it may be possible to use the same transformations and, hence, also reuse the same transformation variables. Because each transformation variable is approximated with different PLFs, this can have a large impact on the solution time of the transformed MINLP problem. Other aspects to be considered when choosing the transformations are, e.g., their numerical stability and what types of transformations are preferred.

The objective function of the MILP problem for optimizing the set of transformations for convexifying the signomials in the MINLP problem described in Eq. (1) is

$$\begin{aligned}\text{minimize}\quad & \delta_{\mathrm{R}} \sum_{i:x_i \in R} r_i B_i + \delta_Z \sum_{i:x_i \in Z} r_i B_i + \delta_I \sum_{i=1}^{I} \sum_{j_1=1}^{J_T} \sum_{\substack{j_2=1 \\ j_1 \neq j_2}}^{J_T} \gamma_{j_1 j_2 i} \\ & + \sum_{i=1}^{I} \sum_{j=1}^{J_T} \left(\delta_{NT} b_{ji} + \delta_{NS} \Delta_{ji} + \delta_{ET} b_{ji}^{ET} + \delta_{PT} b_{ji}^{PT} + \delta_P \beta_{ji} \right).\end{aligned}$$

For simplicity, all signomial terms in the problem have been renumbered according to $j=1,\ldots,J_T$ (where J_T is the total number of signomial terms in the problem), since it does not matter in the transformation step in what constraints the terms are. The impacts of the different parameters δ in the objective function are described in Table 1 and the variables in Tables 2 and 3. The parameter r_i penalizes transformations of variables with large domains and can for example be calculated as follows for $x_i \in [x_{L,i}, x_{U,i}]$:

$$r_i = \begin{cases} 1+\varepsilon(x_{U,i} - x_{L,i}), & \text{if } x_i \in R, \\ 1+\varepsilon(x_{U,i} - x_{L,i} - 1), & \text{if } x_i \in Z. \end{cases}$$

In addition to the objective function, the additional linear constraints described in the rest of this section, are also included in the problem formulation. The following linear conditions – used to express the relation between the binaries b and B – are included for all variables x_i, $i=1,\ldots,I$:

$$\forall i: \quad \sum_{j=1}^{J_T} b_{ji} \leq J_T B_i \,.$$

In the MILP formulation two additional parameters are used: a large positive value M and a small positive value ε. The parameter ε can, e.g., be calculated as $\varepsilon = 1/M$. Because the convexification requirements are different for positive and negative terms, the constraints guaranteeing transformations resulting in convex terms are different depending on the sign of the term. The additional constraints needed for negative and positive signomial terms are described in Sections 3.1 and 3.2 respectively.

3.1. Additional constraints for negative terms

For a negative signomial term, the condition for convexity is that all powers in the terms are positive, and that their sum is between zero and one. Therefore, all variables with negative powers must be transformed but not necessarily all with positive powers

(whenever $Q_i=1$). When a variable x_i is transformed in the j^{th} term the binary b_{ji} should be one, and otherwise zero. Furthermore, the sum of the powers in the transformed terms should be less than or equal to one. These conditions can be written as

$$\forall i: p_{ji} > 0: \quad 1 - b_{ji}(1-\varepsilon) \le Q_{ji} \le 1 - \varepsilon b_{ji}, \quad \forall i: p_{ji} < 0: \quad \begin{cases} -M \le Q_{ji} \le -\varepsilon, \\ b_{ji} = 1, \end{cases} \quad \text{and}$$

$$\sum_{i=1}^{I} p_{ji} Q_{ji} \le 1.$$

To obtain transformations with good numerical properties, i.e., not to small powers in the transformations, the following expressions are included for each variable in a negative signomial term:

$$\forall i: p_{ji} \ne 0: \quad \begin{cases} \Delta'_{ji} \ge \left| p_{ji} Q_{ji} - \frac{1}{I_j} \sum_{i=1}^{I} p_{ji} Q_{ji} \right|, \\ \Delta_{ji} \ge 1 - \sum_{i=1}^{I} p_{ji} Q_{ji} + \varepsilon \Delta'_{ji}. \end{cases}$$

The parameter I_j, specifies the number of variables in the j^{th} signomial term. Note that the expression with the absolute values can be rewritten as two different expressions with different signs on the right-hand side.

3.2. Additional constraints for positive terms

The convexity requirements for positive signomial terms allow for more freedom in choosing the transformations than those for negative terms since we have three different types of applicable transformations, the ET, NPT and PPT. When using the first two, no restrictions on the sum of the powers in the transformed term is required for convexity, however, when using the PPT the sum of the powers in the transformed term should be greater than or equal to one. Furthermore, in the PPT only one of the powers should be left positive after transformation. These requirements can be formulated as

$$\sum_{i=1}^{I} p_{ji} Q_{ji} - M \sum_{i=1}^{I} \alpha_{ji} + M \sum_{i=1}^{I} b_{ji}^{ET} \ge 1 - M \quad \text{and} \quad \sum_{i=1}^{I} \alpha_{ji} - M \sum_{i=1}^{I} b_{ji}^{ET} \le 1.$$

For variables x_i with positive powers, the following conditions are added to guarantee correct transformations:

$$\forall i: p_{ji} > 0: \quad \begin{cases} -M + (M+1)\alpha_{ji} \le Q_{ji} \le M\alpha_{ji} - \varepsilon(1-\alpha_{ji}), \\ b_{ji}^{PT} \ge 1 - \alpha_{ji}, \\ \varepsilon(Q_{ji} - 1) \le b_{ji}^{PT} \le (1-\varepsilon) Q_{ji} + M(1-\alpha_{ji}), \\ b_{ji}^{ET} \ge \frac{1}{I_j} \sum_{i=1}^{I} b_{ji}^{ET}, \\ b_{ji}^{ET} + b_{ji}^{PT} \le 1, \\ b_{ji} = \max\{b_{ji}^{ET}, b_{ji}^{PT}\} \Leftrightarrow b_{ji} \ge b_{ji}^{ET}, b_{ji} \ge b_{ji}^{PT}\}. \end{cases}$$

No transformations are needed for variables with negative powers in positive terms, so the variables are fixed to the following values:

$$\forall i : p_{ji} < 0 : \quad \begin{cases} Q_{ji} = 1, \\ b_{ji}^{ET} = b_{ji}^{PT} = b_{ji} = \alpha_{ji} = \beta_{ji} = \Delta_{ji} = 0. \end{cases}$$

Similarly to the case with negative signomial terms, conditions favoring numerically more stable powers in the PTs should be added. Here, the preferred values for the powers in the PPT and NPT are set using the parameters $P_{POS} \geq 1$ and $P_{NEG} < 0$ in the following conditions:

$$\forall i : p_{ji} > 0 : \quad \begin{cases} M(\beta_{ji} - 1) \leq p_{ji} Q_{ji} \leq M\beta_{ji}, \\ \left| p_{ji} Q_{ji} - P_{POS} \right| \leq \Delta_{ji} + M(1 - \beta_{ji}), \\ \left| p_{ji} Q_{ji} - P_{NEG} \right| \leq \Delta_{ji} + M\beta_{ji}. \end{cases}$$

3.3. Favoring identical transformations for the same variables

The variables γ indicates whether the same transformations are used for the same variables in different terms (see Table 3). The same transformation variable and PLF can be used for the same variable in different terms if the transformations are equal. Since this results in transformed problems of less combinatorial complexities, the number of different transformations has a direct impact on the time required to solve the problem. Since we want $\gamma_{j_1 j_2 i}$ to be equal to one if different transformations are used for the variable x_i in the j_1^{st} and j_2^{nd} term, the following conditions are included:

$$\forall i : \forall j_1, j_2 \in \{1, \ldots J_T\} : (j_1 \neq j_2) \wedge (p_{j_1 i}, p_{j_2 i} \neq 0) \wedge (\operatorname{sgn} c_{j_1} = \operatorname{sgn} c_{j_2}) :$$
$$Q_{j_1 i} - Q_{j_2 i} + b_{j_1 i}^{ET} - b_{j_2 i}^{ET} - M(1 - b_{j_1 i}) - M(1 - b_{j_2 i}) \leq M\gamma_{j_1 j_2 i} \quad \wedge \quad \gamma_{j_1 j_2 i} = \gamma_{j_2 j_1 i}.$$

4. Conclusions

In this paper, the MILP method for optimizing the set of transformations was extended to provide some important additional features: favoring identical transformations of the same variable in different terms and numerically more stable transformations. The main use for the optimization method presented here, is to include it as a preprocessing step in a global optimization solver for MINLP problems containing signomial functions. Thus, creating a solver being able to solve problems of the mentioned form to global optimality without requiring the user to do the transformations manually.

References

A. Lundell, J. Westerlund and T. Westerlund, 2009, Some transformation techniques with applications in global optimization, Journal of Global Optimization, 43, 391-405.

A. Lundell and T. Westerlund, 2008, Exponential and power transformations for convexifying signomial terms in MINLP problems, Proceedings for the 27th IASTED MIC Conference.

T. Westerlund, 2005, Global optimization: From theory to implementation: Some transformation techniques in global optimization, Eds. Liberti L. and Maculan N., Springer, New York.

Parameter	Description	Domain
δ_R	Penalizes the number of transformed real variables	$\boldsymbol{R}_+$
δ_Z	Penalizes the number of transformed integer/binary variables	$\boldsymbol{R}_+$
δ_I	Penalizes different transformations for the same variable in different terms	$\boldsymbol{R}_+$
δ_{NT}	Penalizes the number of transformations	$\boldsymbol{R}_+$
δ_{NS}	Penalizes numerically unstable transformations	$\boldsymbol{R}_+$
δ_{ET}	Penalizes the ET in positive terms	$\boldsymbol{R}_+$
δ_{PT}	Penalizes the PTs in positive terms	$\boldsymbol{R}_+$
δ_P	Favors the PPT if negative, otherwise the NPT	$\boldsymbol{R}$

Table 1 The strategy parameters in the objective function of the MILP problem formulation

Variable	Description	Value
B_i	The variable x_i is transformed in any term	$B_i = 1$
	The variable x_i is not transformed in any term	$B_i = 0$
b_{ji}	The variable x_i is transformed in the j^{th} term	$b_{ji} = 1$
	The variable x_i is not transformed in the j^{th} term	$b_{ji} = 0$
b_{ji}^{ET} *	The variable x_i is transformed using the ET in the j^{th} term	$b_{ji}^{ET} = 1$
	The variable x_i is not transformed using the ET in the j^{th} term	$b_{ji}^{ET} = 0$
b_{ji}^{PT} *	The variable x_i is transformed using a PT in the j^{th} term	$b_{ji}^{PT} = 1$
	The variable x_i is not transformed using a PT in the j^{th} term	$b_{ji}^{PT} = 0$
α_{ji} *	$1 \le Q_{ji} \le M$	$\alpha_{ji} = 1$
	$-M \le Q_{ji} < -\varepsilon$	$\alpha_{ji} = 0$
β_{ji} *	$0 \le p_{ji} Q_{ji} \le M$	$\beta_{ji} = 1$
	$-M \le p_{ji} Q_{ji} \le 0$	$\beta_{ji} = 0$
$\gamma_{j_1 j_2 i}$	Different transformations are used on x_i in the j_1^{st} and j_2^{nd} terms	$\gamma_{j_1 j_2 i} = 1$
	Identical transformations are used on x_i in the j_1^{st} and j_2^{nd} terms	$\gamma_{j_1 j_2 i} = 0$

Table 2 Binary variables (variables marked with * are defined for positive terms only)

Variable	Description	Domain
Q_{ji}	The transformation power used on the variable x_i in the j^{th} term	$\boldsymbol{R}$
Δ_{ji}	For negative terms: The deviation from one for the sum of the powers in the term plus a penalty from Δ'_{ji}. For positive terms: The deviation from P_{POS} or P_{NEG} if the PPT or the NPT respectively is used on x_i in the j^{th} term.	$\boldsymbol{R}$
Δ'_{ji}	For negative terms: The deviation for the product $p_{ji}Q_{ji}$ from the mean of the powers in the term.	$\boldsymbol{R}_+$

Table 3 Real variables

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Model-Based Methodology for Simultaneous Design and Control of a Bioethanol Production Process

Merlin Alvarado-Morales[a], Mohd Kamaruddin Abd Hamid[a], Gürkan Sin[a], Krist V. Gernaey[b], John M. Woodley[b] and Rafiqul Gani[a]

[a]*CAPEC,* [b]*BioEng, Departmenf of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark*

Abstract

In this work a model-based methodology to solve an integrated process design and control (*IPDC*) problem for a bioethanol production process is presented. The *IPDC* problem is formulated and solved such that the economic performance is optimized in terms of a cost effective design and controllable process. The concepts of attainable region (*AR*) and driving force (*DF*) are used within this methodology, to determine the optimal design-control of the process as well as to generate feasible alternatives. Based on this methodology, the optimal solution to the design-control problem is found by locating the maximum value of *AR* and *DF* for reactor and separator, respectively. The use of *DF* and *AR* concepts are shown to provide an optimal design with respect to energy consumption for the downstream separation units and with respect to controllability for the simultaneous saccharification and fermentation (*SSF*) bioreactor unit, respectively, used in the bioethanol production process.

Keywords: bioethanol, integrated process design and control, model-based methodology, attainable region, driving force.

1. Introduction

The production of bioethanol as a fuel from renewable resources has recently attracted much interest. However, the production process is complicated, and bioethanol has therefore only become a viable and realistic product in the energy market due to rising environmental concerns and the increase in oil price. The economic competitiveness of bioethanol as a liquid fuel depends strongly on the resources used during its production and also the plant operability. Furthermore, the production of fuel ethanol from biomass generates large volumes of wastewater, which must be processed in order to minimize the environmental impact. Therefore, the development of a cost-effective, sustainable, and controllable process for bioethanol production is a priority for many research centers, universities, and private companies.

The starting point for our work is based on a bioethanol production process which has been documented by the National Renewable Energy Laboratory (*NREL*) [1]. The main unit operations are highlighted in Figure 1. The feedstock, in this case hardwood chips, is delivered to the feed handling unit. It is then sent to the pretreatment process where the biomass is decomposed into cellulose, and hemicellulose sugars are solubilized. Subsequently, the hydrolysate slurry is separated into liquid and solid fractions. The liquid fraction undergoes detoxification to eliminate inhibitory compounds and the pH is adjusted with lime. The liquid and solid fractions are recombined to be sent to the

simultaneous saccharification and co-fermentation (*SSCF*) process. In the *SSCF* process two different operations are performed, saccharification of the remaining cellulose to glucose using enzymes and fermentation of the resulting glucose and other sugars into ethanol using microorganisms. Finally, effluent from the fermentor is sent to the downstream separation process.

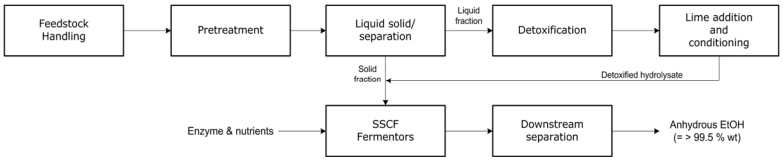


Figure 1. Bioethanol production process flowsheet from lignocellulosic biomass based on the NREL process [1].

The process flowsheet (base case design) was simulated using the PROII® simulator, and simulation results were analyzed in terms of process economic feasibility and sustainability. As a result, an improved bioethanol production process in terms of sustainability was obtained [2]. To further improve the economics of the process, the design-analysis of the downstream separation is considered by using the driving force (*DF*) based-method [3]. However, as the bioreactor, where enzymatic hydrolysis and fermentation occurs, has the inconvenience that the optimal conditions for hydrolysis and fermentation are different, implying thereby a difficult control and optimization of process parameters [4], it is worth investigating the question of integrated process design and control for this bioreactor unit. This paper therefore highlights the use of a model-based methodology for the integrated process design and control (*IPDC*) of the bioreactor and the downstream separation units for the bioethanol production process.

2. Methodology

2.1. Integrated process design and control methodology

The model-based methodology for *IPDC* aims at identifying the optimal operating point where both design and control objectives are best satisfied [5]. The *IPDC* problem is decomposed into four sequential hierarchical stages based on the formulation of a mixed-integer dynamic optimization problem to find the optimal point which gives the best performance (economic, operability, etc.): (1) pre-analysis, (2) steady-state analysis, (3) dynamic analysis, and (4) evaluation stage. Stage 1 in this model-based approach defines the design-control targets by locating the maximum values of the attainable region (*AR*) and driving force (*DF*) for design feasibility (reaction or separation are not possible if the *AR* or *DF* is zero). In stage 2, the established target is validated by finding the feasible values of design-control variables (candidates) that match the target. In Stage 3, the selected optimal operating point identified in the previous stage together with the corresponding design-control variables are analyzed in terms of performance indices such as controllability. The objective here is to verify that at the selected point the performance of the system is best satisfied at the minimum value of the derivative of *AR* or *DF* (minimum value of derivative determines the sensitivity and controllability). The last stage analyzes the economic performance of the process in terms of energy or cost. Therefore, design-control of a reactor and/or separator at the largest total *AR* and/or *DF* leads to lower sensitivity, better controllability, and the minimum energy consumption/cost process.

2.2. Design of a bioreactor through the attainable region approach

The starting point for the design of a single reactor are the reaction rates, visualized in the attainable region diagram, where the concentration of the desired product is plotted as a function of the concentration of the reactant as shown in Figure 2a.

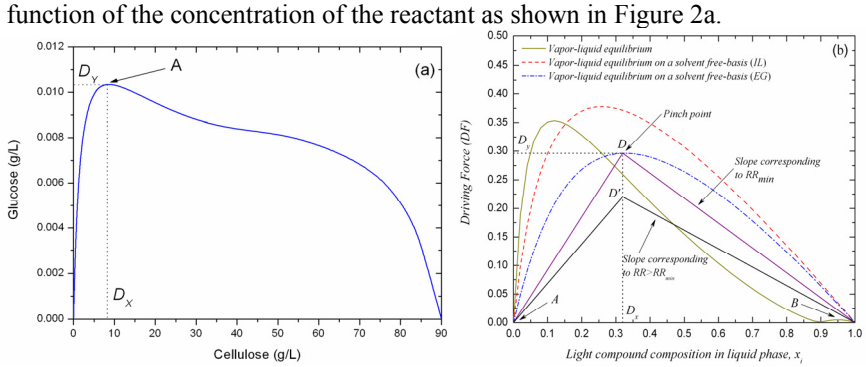


Figure 2. a: Attainable region diagram; b: Solvent-free basis driving force diagram for ionic liquid and ethylene-glycol with illustration of the design parameters.

The design steps involve the following: 1) Develop the attainable region diagram; 2) Identify the maximum value (point *A*) of the attainable region diagram and its corresponding point D_X and D_Y; 3) Compute the reaction temperature, *T* (if applicable) and residence time, τ_R from their respective models; 4) Compute the reactor volume ($V=F/\tau_R$); 5) Compute the real reactor volume ($V_R=1.1V$).

2.3. Design of a separator through the driving force approach

The starting point for *DF*-based methods is the availability of phase composition data, and secondly the graphical representation of the phase composition data as shown in Figure 2b. The *DF*-based method [3] has been used for the synthesis and design of the downstream separation process.

3. Results

3.1. Reactor design results

In order to have a graphical representation of the *AR* analysis for a bioreactor, kinetic models describing the *SSF* process were taken from South *et al*. [6]. The objective of this design is to obtain the highest concentration of ethanol within the range tolerated by the microorganisms (*Eth*(*w*/*v*)≤10%). Note indeed that the *SSF* bioreactor is considered instead of the *SSCF* bioreactor, due to the large uncertainties of the experimental data as well as the lack of reliable kinetic models for the *SSCF* process. Nevertheless, the design-analysis with the *AR* based method for the *SSF* bioreactor model serves the purpose of illustrating the main concepts (see also section 3.3). The *AR* analysis was performed at different values of enzyme loading at two different initial cellulose concentrations (Case 1 = 1 *g/l* and Case 2 = 90 *g/l*). The results are shown in Figure 3 and 4. For Case 1, a low enzyme loading is required in order to obtain the attainable ethanol concentration at the highest concentration of glucose. The optimum ethanol concentration of 0.08 *g/l* can be obtained at 0.048 *g/l* of glucose with 45 *FPU/g* of enzyme loading as shown in Figure 3a. This is corresponding to 0.458 *g/l* of cellulose (Figure 3b). Conversely, for Case 2, a high enzyme loading is required to obtain the attainable ethanol concentration at the highest concentration of glucose. In this case (see Figure 4a), the ethanol concentration of 43.98 *g/l* is obtained at the enzyme loading of 200 *FPU/g* (see Figure 4b). Clearly, this is a very high enzyme loading, compared to the

SSCF bioreactor process in which the enzyme loading is usually between 3.5 and 15 *FPU/g* [1]. Interestingly, it can be noted that all the specifications of the *SSCF* cannot be matched with the *SSF* bioreactor. Matching the low enzyme loading leads to a low ethanol concentration at low initial concentration of cellulose, while matching the ethanol concentration leads to a high enzyme loading but for the correspondingly higher initial concentration of cellulose. The reason for this difference is that the *SSF* process model does not account for conversion of C5-sugars into ethanol. Conversion of C5-sugars is important for the process economy, since *SSCF* gives a higher reported ethanol yield from cellulose than *SSF* and requires lower amounts of enzyme [1].

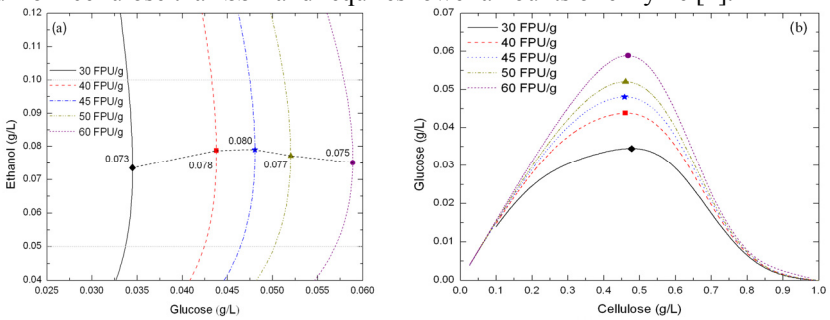


Figure 3. *AR* space-concentration diagram for a: Ethanol-Glucose and b: Cellulose-Glucose at initial concentration of cellulose of 1 *g/l* (Case 1).

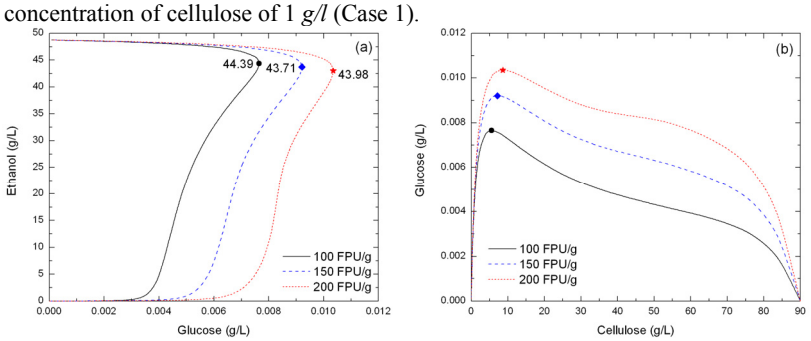


Figure 4. *AR* space-concentration diagram for a: Ethanol-Glucose and b: Cellulose-Glucose at initial concentration of cellulose of 90 *g/l* (Case 2).

3.2. Separator design results

The downstream separation process objective is to obtain anhydrous ethanol from an ethanol-water mixture. A solvent-based separation process has been considered by using ethylene-glycol (*EG*) in order to highlight the *DF*-based method for synthesis and design of separation processes. Table 1 summarizes the results with respect to process design variables.

Table 1. Downstream separation design results

	Distillation column	Extractive column	Recovery column
Number of stages	32	30	15
Feed stage	17	22	5
Reflux ratio	3.2	0.52	0.54
DF_{max}	0.35	0.48	0.59
Energy consumption	1633 $kcal/kg_{Eth}$		

The process separation scheme has been validated by rigorous simulation to evaluate performance in terms of energy consumption. For the base case design, an energy

consumption of 1743 $kcal/kg_{Eth}$ is obtained, whereas the obtained energy consumption for the solvent-based separation process is 1633 $kcal/kg_{Eth}$, a reduction of 6.3 % with respect to the base case design. The reason for this reduction is that the solvent-based separation process is now designed at the maximum value of *DF*. Designing separation processes at the largest *DF* leads to more energy efficient separation processes.

3.3. Integration of design and control

In this section, results of the *IPDC* analysis for the *SSF* reactor unit (Case 2) are presented. As explained in the previous section, a reactor should be designed at the maximum point of the *AR*. For control, at the selected point the controllability performance of the system is best satisfied. In order to show that, the absolute value of derivatives of glucose with respect to cellulose is plotted for the sensitivity analysis as shown in Figure 5. At this point the derivative of glucose with respect to cellulose has a minimum value. In Figure 5 also, there are two other points (B and C) representing two alternative operating points which are below the maximum *AR* point. From a design point of view, they are not feasible since point B generates a lower ethanol concentration and point C requires larger residence time. Thus, from an integrated design-control point of view, only the control issues will be highlighted. Since the derivative at point A is smaller than at points B and C, the process sensitivity of point A is low. From a control point of view, any changes in cellulose will give smaller changes in glucose at point A compared to points B or C. Therefore, by maintaining the cellulose at point A, the desired glucose can more easily be obtained (controlled) than for the other points. Consequently, the corresponding ethanol concentration can easily be controlled. In order to maintain cellulose at its desired value, one option is to manipulate the enzyme loading (manipulated variable).

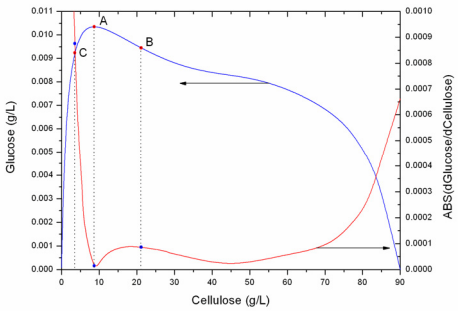


Figure 5. *AR* space-concentration diagram for Glucose-Cellulose and its corresponding derivative (in absolute value) with respect to Cellulose for Case 2.

Let us suppose a disturbance (-5% change in input cellulose) moves cellulose away from its set points (points A, B, C). According to Figure 5, any changes in the cellulose concentration at points B and C will easily move the glucose concentration away from its steady state value and as a result, it will be more difficult to maintain the ethanol concentration at these points than at point A. Figure 6 shows the output response of cellulose, glucose, ethanol and its corresponding control actions to reject a step disturbance (-5% input cellulose) using a P-controller. Although, there is some off-set in the output cellulose concentration at point A (see Figure 6a), the off-set in output for glucose is negligible while the off-set in output for ethanol is quite small (see Figure 6b and 6c). Notably, point A shows better performance in terms of off-set than other points. From a design point of view, design at point A is feasible since it satisfies the design criteria of ethanol concentration, whereas points B and C are not. From a control point of view, candidate A has better controllability performance in terms of disturbance

rejection (see Figure 6d). Consequently, candidate A is the optimal bioreactor design that satisfies design and control objectives.

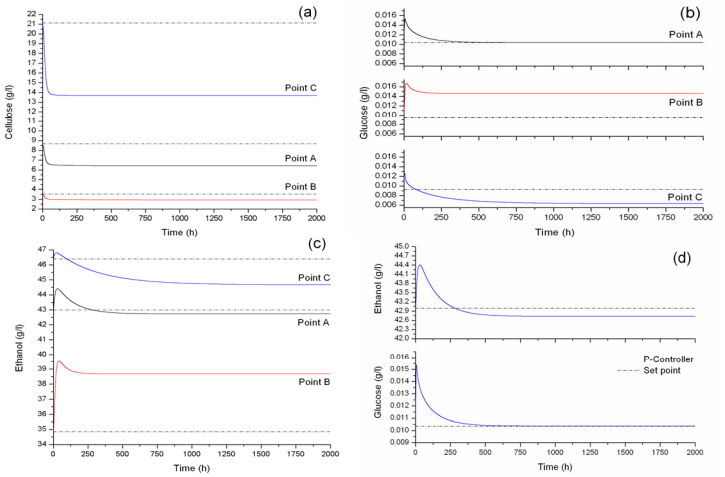


Figure 6. Rejection of a step disturbance (-5% input cellulose) for Case 2 using a P-controller: Output responses of all points for (a) cellulose; (b) glucose; (c) ethanol; and (d) ethanol and glucose at point A.

The same analysis is also valid for the *DF*-based separation design and control analysis since the graphical concept of *AR* and *DF* regions are similar. Design of the separation system should be performed to maximize the total *DF*. If the separation becomes more difficult, the operation and control become more expensive.

4. Conclusion

In this paper, a model-based methodology provides useful information related to the solution of the integrated process design and control of a bioethanol production process. By taking the maximum value of the *AR* or *DF*, the optimal solution with respect to the design, control and economic performance can be estimated before the rigorous simulation is attempted. For bioreactor design, we successfully demonstrated the implementation of the *AR* concept to provide an optimal design with better dynamic performance. Numerical simulation results indicated that the resulting reactor design is able to maintain its steady state conditions with better dynamic performances. For the downstream separation unit, the use of the *DF*-based concept is shown to provide an optimal design with respect to energy consumption. Current and future work is focused on applying this methodology considering the *SSCF* process, as well as adding design-control analysis for all unit operations of the bioethanol production and other processes involving reaction and separation.

References

1. R. Wooley, M Ruth, J. Sheehan, K. Ibsen, H. Majdeski, A. Galvez, 1999, Technical Report. NREL/TP-580-26157, Golden Colorado USA.
2. M. Alvarado-Morales, J. Terra, K. V, Gernaey, J. M. Woodley, R. Gani, 2008, *ChERD*. Submitted.
3. R. Gani, E. Bek-Pedersen, 2000, *AIChE J.*, 46(6), 1271-1274.
4. P. A. M. Claassen, J. B. van Lier, A. M. Lopez Contreras, E. W. J. van Niel, L. Sijtsma, A. J. M. Stams, S. S. de Vries, R. A. Weusthuis. 1999, *Appl. Microb. and Biotech.,* 52, 741–755.
5. M. K. A. Hamid, R. Gani 2008, *In Proceedings of the FOCAPO 2008*, MA., USA, 205-208.
6. C. R. South, D. A. L. Hogsett, L. R. Lynd, 1995, *Enzyme Microb. Technol.*, 17, 797-803.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

New generic approach for the analysis of energy conversion system models

Raffaele Bolliger, Helen Becker, François Maréchal

Industrial Energy Systems Laboratory (LENI), Ecole Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland

Abstract

This paper proposes a new approach in the field of energy conversion systems analysis and synthesis. The method is based on a generic and multi-platform description syntax, that clearly separates the information concerning the physical behavior of the modeled technology (e.g. mass and energy balances and chemical reactions) from the information necessary to apply one or more system analysis methods (e.g. process integration, Life Cycle Impact Assessment, thermo-economic evaluation,...).
The description syntax also contains other informations about the model, namely about its history, quality, scope or documentation.
The approach encourages the development of reusable models, which can be easily assembled to create large superstructures from which optimal system configurations can be extracted. By dissociating technology models from the analysis and synthesis method, the approach allows the independent development of analysis methods and the consistent data transfer between models of different scales.
The study of a fine chemical industry waste incineration system is presented to demonstrate the flexibility of the approach.
Keywords: Energy conversion system analysis computer aided process engineering process design process integration

1. Introduction

The design of energy conversion systems is based on models which describe the mass and energy balances for the different process units and their integration into the whole system. These models generate the data needed to analyze the overall system efficiency and to establish performance indicators, using for example exergy analysis, process integration with pinch analysis, life-cycle assessment or thermo-economic evaluation. The increasing complexity of the system, the highest degree of integration and the increasing number of energy conversion options together with the demand of applying different performance indicators require more systematic approaches that go beyond the use of simulation models. This paper proposes a methodology to systematically tackle the integrated system design by dissociating technology modeling from the methods for the analysis and the synthesis of integrated systems.

2. Literature review

In recent years, research activity in energy conversion system analysis and design evolved towards more complex and more integrated systems, often composed by combining smaller sub-systems. The domain covers multiple system scales from equipment design [11] to process design (e.g. biomass conversion processes [5]) to industrial processes (e.g. fine chemicals batch plants [2]) and even urban systems [10].

In order to address the problem of handling complex models, research recently focused on developing tools for exchanging informations and allow the interoperability of modeling softwares. For example, The DOME platform (distributed object-based modeling environment) [3] implements a model based co-current system design and engineering platform. It has been successfully applied to urban systems [7] allowing to interconnect versatile sub-systems models accessed from the internet using web services. The CAPE-OPEN [4] initiative on the other hand has been developed by the process engineering community to allow the interoperability of flow-sheeting tools, unit models and thermodynamic packages. Although these methods give the opportunity to construct very complex models, most of the time they are only concerned with the process flowsheet calculation problem.
The development of the system model however only solves the first part of the problem. The design of the energy conversion systems requires the application of one or more analysis and synthesis methods allowing one to deduce the performance indicators and the information about the interactions between the various technologies included in the system with a system holistic vision. Examples of such methods are energy and process integration, life-cycle analysis, thermo-economic evaluation or exergy analysis.
The major drawback of the existing approaches is that they do not separate energy and mass balances modeling from information relative to the application of system analysis and synthesis methods. Models are built in one single block containing all the information. The reuse of the same models in different study contexts becomes therefore very difficult and often requires a partial recoding of the model.

3. Proposed method

We propose a new method to handle and connect models by separating the chemical and physical models of the process units from the methods used for system-scale analysis and synthesis. The method is based on a generic syntax which describes the different sub-system models and the interfaces available to model their integration in the system. A schematic representation of the concept is shown in Figure 1.

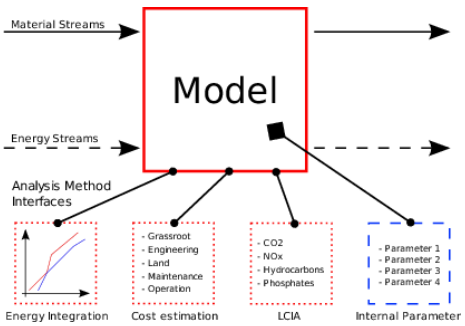


Figure 1 Schematic view of the separation of physical model from method analysis related data

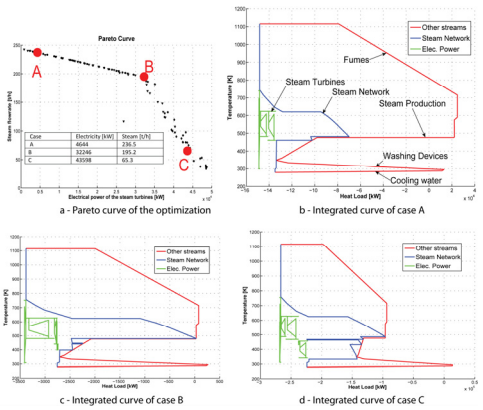


Figure 2 Case study results

The syntax provides a generic abstraction layer that describes the models independently from their modeling environment and considers their reuse in different projects. This also allows one to develop a data base of technology models that may be used to systematically develop process superstructures with interconnected technologies.
Furthermore, the use of a generic and homogeneous syntax promotes the development of analysis and synthesis methods that are independent of the technology models, hence expanding the scope of application and allowing multiscale approaches.
A model representing a technology is the combination of a set of modeling parameters used to represent the characteristics of the technology and its environment and of a set of equations that describes the thermo-physical and chemical conversion operations of the technology. Models can be connected via their input and output streams and possibly by sharing some model parameters.
Each system analysis method applied to the model has a dedicated interface, where all the required informations are stored in a structured variable. The system analysis interfaces include the extraction and the generation of the required data for the analysis methods from the model results. For process integration application, the interface includes the definition of the hot and cold streams of the technology as well as the related energy flows (electricity, water, resources,...). For the cost estimation, the interface provides the values required to estimate the cost (e.g. the size, pressure and temperature conditions, the materials, ...). For the environmental impact assessment, the model provides information about the materials used and the size of the technology that are required to compute the direct and the indirect emissions of the technology. For the exergy analysis, not only the input, output, enthalpy and entropy are given but more detailed data are also provided such as exergy losses in the chemical reactions, pressure losses, etc... The definition of the analysis method interfaces implies therefore a detailed analysis not only of the technology itself but also of how it will interact with the other technologies in the analyzed system. This analysis is however done independently of the other technologies considered and is therefore independent of the system analyzed.
The system analysis and synthesis methods will use the system analysis interfaces to generate the performance indicators at the system level considering the possible integration of each technology in the system.
The interest of the approach is the separation of tasks: the development of a new analysis method does not require the modification of the physical model and does not affect the information handled by the other methods. However, it may require specific development when particular data are needed.
The syntax also defines other data, in relation with the inclusion of the model into a shared database. These informations concern the model classification, its documentation, the history of modifications and more generally data allowing to establish the model accuracy and quality. Table 1 defines the domains covered by the syntax.

Table 1: Domains covered by EnergyTechnologies

Domain	Description
Physical model	Definition of input and output flows (material and energy) crossing the model boundaries. Description of the model parameters used to control its behavior.
System Analysis interfaces	Definition of data required to apply system analysis and synthesis methods.
Classification	Information used to place the model inside searchable database.
Files, version	Information about model content, required software, versions.
Changesets, references	Information about the authors, model changes and improvements.
Documentation	References to model documentation and bibliographic references.
Validation	Information about model quality and validation status.

The syntax defined is generic: it can be implemented in any programming language accepting structured variables; the content of the structures is suitable to describe any energy conversion technology and more generally any model which can be built with a black-box technique.

4. Case study - waste incineration system

The method has been applied to study the integration of a liquid waste incineration system composed by four units.

The first sub-system model represents the liquid waste incineration plant. The unit model includes the burner, the fumes washing devices and a deNOx. The thermodynamic model is built with a commercial flow-sheeting software.

The second model represents the steam network. It includes the heat recovery steam generator, the steam turbines and the pumps. The model is formulated as a mixed integer linear programming optimization problem, where the optimal layout of the steam network in terms of electricity production is computed according to the available heat profile in the steam generator [9, 1].

The third sub-system model represents the cooling water system, which is used to cool down the fumes and the steam in the condenser. The fourth model represents the industrial plant where the waste incineration plant is situated. The model defines the heat requirement as the amount of steam to be produced to satisfy the heat requirement of the plant. The third and fourth models are simple representations of a cold and a hot stream, defined in the energy integration interface described in Figure 1, without any model applied. When the model is available, then the unit will be substituted without having to redevelop whole model.

If a new alternative method exists for treating the liquid waste (e.g. a wet oxidation process), the model will be added to the system and the system analysis will allow one to compare the two options.

An energy and process integration analysis is performed on the system in order to study the interactions between the four models. The heat recovery opportunities will be studied in order to maximize the combined steam and electricity production.

A multi-objective optimization using an inhouse advanced evolutionary algorithm [8] is performed to find the optimal layout of the steam network in order to maximize two objectives in competition: steam and electricity production. The optimization is run in an inhouse computational platform developed to apply the proposed approach[5].
This example let us show the flexibility of the approach in terms of system integration. The process integration of the utilities is visualized by the integrated composite curves (Fig. 2 b,c and d), while the multi-objective optimization systematically generates competing technical solutions and is an extremely valuable computer aided decision tool for the engineer.

4.1. System analysis method: process integration

The objective of pinch analysis and process integration is to identify the heat recovery potential between the hot and cold streams in a system [6]. All the cold and hot streams of the process have to be identified in the various units of the process and the utilities. Based on the definition of a minimum temperature difference (ΔT_{min}) the minimum energy requirement (MER) is computed. The maximum energy recovery between the process streams is also calculated.
In the case of the proposed example, the waste incineration system and the steam requirement are defined as process units, while the steam network and the cooling water are defined as utilities whose sizes have to be optimized. The process integration model [9] computes the optimal flows in the utility system in order to satisfy the process requirement. At the same time, the mechanical power produced by the steam turbines is maximized.
The composite curves are used to analyze the solutions and to visualize the process integration. The composite curve of the hot and cold streams shows the minimum energy requirements of the process.
The grand composite curve shows if the utilities can satisfy the process requirements and displays the potential heat recovery improvements.
Integrated curves will be used to analyze the integration of a particular unit with the rest of the whole system.

4.2. Optimization results

The multi-objective optimization of the problem leads to a Pareto optimum set (Fig. 2a) which represents the best compromise between the two objective functions. Solutions appear to span over two different zones. The table in Figure 2a displays electricity production and steam flow rates corresponding to cases A to C. Figure 2a, and c display the integrated curve of the steam network for cases A to C.
The analysis of the results shows that case B corresponds to the point where the maximum steam flow is reached in the high pressure header. The increase in electricity production going from case B to case C is performed at the expense of the steam production that quickly decreases. On the other hand, steam production is only slightly affected with the increase of electricity production at high temperature, from case A to B. Solution B is therefore the best compromise in terms of steam/electricity production ratio.

5. Conclusions

The proposed approach addresses the problem of the size increase in the domain of the analysis and synthesis of energy conversion systems. In order to overcome the complexity of the models, which are often developed on a per-study bases, this paper proposes a new method, which separates the modeling of a technology from the

information necessary to apply system analysis methods. This is done by defining a generic and platform-independent syntax, which can be used to describe models built with commercial or inhouse packages.
The approach encourages the development of more generic and reusable models, which can be stored in a shared database and later on be reused in future studies. The proposed approach allows to uncouple the technology model development from the development of system analysis and synthesis methods whose models and tools become independent from the technology itself.

References

[1] Bolliger, Raffaele and Favrat, Daniel and Maréchal, François. Advanced Power Plant Design Methodology using Process Integration and Multi-Objective Thermo-Economic Optimisation. *ECOS 2005, 18th International Conference on Efficiency, Cost, Optimization, Simulation and Environmental Impact of Energy Systems*, pages 777--784, Trondheim, Norway, 2005.

[2] Bolliger, Raffaele and Palazzi, Francesca and Maréchal, François. Heat Exchanger Network (HEN) costs and performances estimation for multi-period operation. *18TH European Symposium on Computer Aided Process Engineering*. Elsevier.

[3] Borland,N. and Wallace,D. Environmentally conscious product design: A collaborative Internet-based modeling approach. *Journal of Industrial Ecology*, 3(2-3):33-46, 1999.

[4] Ray Dickinson. Expanding Process Modelling Capability through software interoperability standards: Application, extension and maintenance of CAPE OPEN standards. In Bertrand Braunschweig and Xavier Joulia, editors, *18th European Symposium on Computer Aided Process Engineering* in Computer Aided Chemical Engineering, pages 537 - 537. Elsevier, 2008.

[5] Gassner, Martin and Maréchal, François. Methodology for the optimal thermo-economic, multi-objective design of thermochemical fuel production from biomass. *Computers and chemical engineering*, 33(3):769--781, 2009.

[6] Kemp,I.C. Pinch Analysis and Process Integration: a user guide on process integration for the efficient use of energy. *Second Edition, Elsevier, Butterworth-Heinemann, UK*, 2007.

[7] Kraines,S. B. and Wallace,D. R. Urban sustainability technology evaluation in a distributed object-based modeling environment. *Computers, Environment and Urban Systems*, 27(2):143-161, 2003.

[8] G. Leyland. *Multi-objective optimisation applied to industrial energy problems*. PhD thesis, Ecole Polytechnique Federale de Lausanne, 2002 annote = This PhD thesis presents the development of MOO.

[9] Francois Marechal and Boris Kalitventzeff. Targeting the optimal integration of steam networks: Mathematical tools and methodology. *Computers And Chemical Engineering*, 23:S133-S136, 1999.

[10] Marechal, Francois and Weber, Celine and Favrat, Daniel. *Energy Systems Engineering*, chapter Multi-Objective Design and Optimisation of Urban Energy Systems, pages 39-81. Wiley, 2008.

[11] Palazzi, Francesca and Autissier, Nordahl and Maréchal, François and Van herle, Jan. A Methodology for Thermo-Economic Modeling and Optimization of SOFC Systems. *Chemical Engineering Transactions*, 7:13--18, 2005.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

An Improved Quasi-Sequential Approach to Large-Scale Dynamic Process Optimization

Weirong Hong [a], Pengcheng Tan [a], Quoc Dong Vu [b], and Pu Li [b]

[a]*Institute of Process Equipment, Zhejiang University, Hangzhou 310027,China*
[b]*Institute of Automation and Systems Engineering, Ilmenau University of Technology, PO Box 100565, Ilmenau 98684, Germany*

Abstract

We present an improved quasi-sequential approach to large-scale dynamic process optimization. This new approach incorporates the quasi-sequential approach with the interior point method to handle inequality constraints. In this way the eventual optimization problem to be solved becomes a NLP without constraints. Mathematical derivations and computation schemes are developed. We first take a two-dimensional constrained optimization problem as an example; the result is compared with the simultaneous and quasi-sequential approach in terms of path solution with a graphical interpretation. A highly nonlinear reactor optimal control problem is also taken to demonstrate the effectiveness of this approach.

Keywords: Dynamic Optimization, Quasi-Sequential Approach, Interior-Point, NLP

1. Introduction

Dynamic optimization can be applied in many decision-making situations in process systems engineering. It remains as a challenging research task for high complex as well as high dimensional applications. These optimization problems are based on a rigorous model and often lead to many highly nonlinear equality and inequality constraints. Even though a solution may be found in some cases, the computation efficiency, convergence rate and initial-value sensitivity may be unsuitable for modern process industries, especially for real-time applications. The available solution approaches can be generally classified into two groups, sequential and simultaneous strategies (Biegler et al., 2002, Biegler, 2007). A new approach, the quasi-sequential approach that takes the advantages of both the simultaneous and the sequential strategies, was recently proposed (Hong et al., 2006). In this study, we extend this quasi-sequential approach by using the interior-point method to handle the nonlinear programming problem inside the solution framework. This is due to the fact that the interior-point method can efficiently handle a large number of inequality constraints. We call it interior-point quasi-sequential approach. In this new approach, state and control variables are first discretized completely as in the simultaneous approach with collocation on finite elements, so as to overcome the difficulty in the sequential approach with satisfying state path constraints. Second, model equations and state variables are eliminated in the same manner as in the sequential approach by a simulation step, so that the optimization problem is reduced to a small NLP only with inequality constraints and control variables. Finally, the primal-dual interior-point method is employed to handle the inequality constraints. In this way the final optimization problem to be solved becomes a NLP with equality constraints by adding slack variables to the inequalities. Mathematical derivations and computation schemes for the interior-point quasi-sequential approach have been developed.

Taking a two-dimensional equality constrained optimization problem as an example the resulting approach is compared with the simultaneous and quasi-sequential approaches in terms of path solution with a graphical interpretation. A highly nonlinear reactor optimal control problem is also considered to demonstrate the effectiveness of this approach. It can be concluded that this interior-point quasi-sequential approach is better than the quasi-sequential approach for solving highly nonlinear large-scale dynamic optimization problems with multiple inequality path constraints. Therefore it provides a promising solution to nonlinear model predictive control and real-time dynamic optimization of engineering processes.

2. NLP formulation

The constrained dynamic optimization problem of interest in this paper can be stated as follows:

$$\min_{x,u} j(z,u) \tag{1}$$

$$s.t. \quad F(\dot{z},z,u) = 0 \tag{2}$$

$$z_{\min} \le z \le z_{\max} \tag{3}$$

$$u_{\min} \le u \le u_{\max} \tag{4}$$

$$z(0) = z_0 . \tag{5}$$

In this formulation, $z(t)$ and $u(t)$ are state (dependent) and control (independent) variables, respectively. Eq. (2) defines a general system of DAEs, Eq. (3) and (4) describe path constraints on the state and control variables, and Eq. (5) the initial condition of the state variables, respectively. In the framework of interior-point methods, inequality constrained functions will be reformulated through the addition of slack variables with lower bounds of zero. These slack variables are added to the state variable vector. Correspondingly, the reformulated inequality constraints which are now equalities are added to the system of DAEs.

The quasi-sequential approach (Hong et al., 2006) possesses advantages of both the simultaneous and the sequential approach. It is based on a completely discretization of the state and control variables with collocation on finite elements like the simultaneous approach, and also it eliminates the state variables and the DAEs in the manner as the sequential approach. Only control variables and inequalities will be handled in the NLP solver. By using the collocation method, the differential equations of the DAE system are converted to a set of algebraic equations. The DAE model is solved successively in each element at each NLP iterate, and the sensitivities of the state variables with respect to control variables are computed in parallel with the DAE solution. We describe the interior-point quasi-sequential approach in more detail as follows. After discretization with collocation, Eqs. (1)-(5) can be represented as the following NLP problem:

$$\min_{z\in\Re^m, u\in\Re^{n-m}} f(z,u) \tag{6}$$

$$\text{s.t.} \quad c(z,u) = 0 \tag{7}$$

$$z_U \ge z \ge z_L \tag{8}$$

$$u_U \ge u \ge u_L \tag{9}$$

After the solution of model equations in the quasi-sequential framework, the problem is reduced to the form:

$$\min_{u\in\Re^{n-m}} f(z(u),u) \tag{10}$$

$$\text{s.t. } z_U \geq z(u) \geq z_L \tag{11}$$

$$u_U \geq u \geq u_L \tag{12}$$

This problem consists only of control variables as well as inequality constraints which are enforced at each collocation point. Without loss of generality, the problem formulated in (10)-(12) can be rewritten as following:

$$\min_{u\in\Re^{n-m}} f(z(u),u) \tag{13}$$

$$\text{s.t. } z(u) \geq 0 \tag{14}$$

$$u \geq 0 \tag{15}$$

We now replace the inequalities by a logarithmic barrier term that is added to the objective function and yield:

$$\min_{u\in\Re^{n-m}} j_m(u) = f(z(u),u) - m\left(\sum_{i=1}^{m}\ln(z(u)^{(i)}) - \sum_{i=1}^{n-m}\ln(u^{(i)})\right) \tag{16}$$

with the barrier parameter $m>0$. Now we have an unconstrained NLP problem to solve. For a given m, multiplier estimates, n_1 and n_2, are introduced and the optimality conditions for (16) can be written in the primal-dual form:

$$\begin{aligned} \nabla f - (\frac{dz}{du})^T n_1 - n_2 &= 0 \\ V_1 Ze - me &= 0 \\ V_2 Ue - me &= 0 \end{aligned} \tag{17}$$

where $e = [1,\ldots,1]^T$, $v_1 = mZe$ and $v_2 = mUe$, Z and U are diagonal matrices with z and u on their diagonals, respectively. The Newton method can be used to solve the system of nonlinear equations (17). Then the search direction ($d_k^u, d_k^{n_1}, d_k^{n_2}$) at iterate k ($u_k, v_{1,k}, v_{2,k}$) can be obtained as a solution of the linearization of (17), that is:

$$\begin{bmatrix} W_k & (\frac{dz}{du})_k^T & -I \\ V_{1k}(\frac{dz}{du})_k & Z_k e & 0 \\ V_{2k}e & 0 & U_k e \end{bmatrix} \begin{pmatrix} d_k^u \\ d_k^{v_1} \\ d_k^{v_2} \end{pmatrix} = -\begin{pmatrix} \nabla f_k - (\frac{dz}{du})_k^T v_{1k} - v_{2k} \\ V_{1k}Z_k e - m_k e \\ V_{2k}U_k e - m_k e \end{pmatrix} \tag{18}$$

where $W_k = \nabla^2_{uu} L(u_k, n_{1,k}, n_{2,k})$. Eliminating the $d_k^{n_1}$ and $d_k^{n_2}$ in Eq.(18) yields the search direction:

$$(W_k + \Sigma_k) d_k^u = -\nabla_m j\,(u, v_1, v_2) \tag{19}$$

where: $\Sigma_k = (\frac{dz}{du})_k^T V_{1k} Z_k^{-1} (\frac{dz}{du})_k + V_{2k} U_k^{-1}$

$$d_k^{v_1} = -v_{1k} + Z_k^{-1} m_k e - Z_k^{-1} V_{1k} (\frac{dz}{du})_k \; d_k^u$$

$$d_k^{v_2} = -v_{2k} + U_k^{-1} m_k e - U_k^{-1} V_{2k} \; d_k^u$$

3. Comparison of the solution path

We illustrate the solution path of the quasi-sequential, the interior-point quasi-sequential and the simultaneous approach with the following example. Consider the *Rosenbrock* two-dimensional constrained optimization problem:

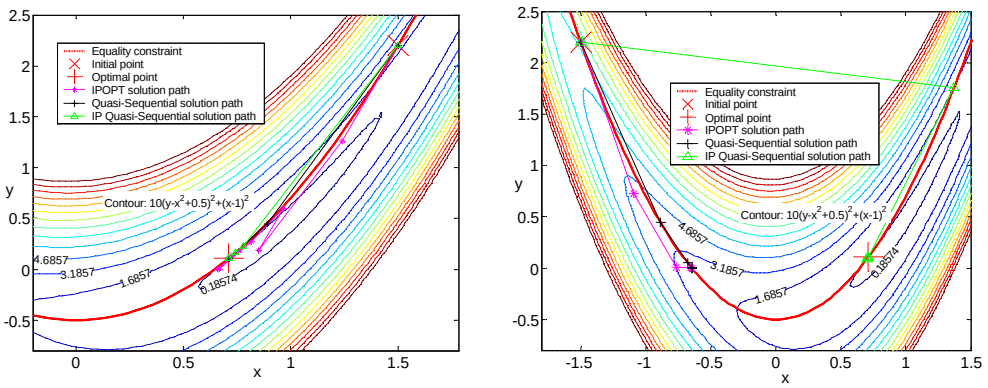
$$\min \ 10(y - x^2 + 0.5)^2 + (x-1)^2 \tag{20}$$

$$\text{s.t} \ \ y = 1.2x^2 - 0.5 \tag{21}$$

$$0 < y < 5 \tag{22}$$

In the sequential approach, the variable x is regarded as the control variable. We solve the problem by the simultaneous, the quasi-sequential approach, and interior-point quasi-sequential approach. The initial point is set at (1.5, 2.2) and (-1.5, 2.2). The optimal point should be at (0.7117, 0.1078), and the minimum objective function value is 0.18574. The solution paths of all three approaches are presented in Fig. 1 where the interior-point quasi-sequential approach is called IP quasi-sequential for short. And the program package IPOPT (Wächter, 2002) is used for the simultaneous approach.

As shown in Fig.1, when the initial point is set at (1.5, 2.2), all of approaches can converge to the optimal point. Moreover, the solution paths are almost the same. But if the initial point is changed to (-1.5, 2.2), only the IP quasi-sequential approach can converge to the optimal point. The IPOPT and quasi-sequential approaches fail, and terminate at the bound of y = 0. The IP quasi-sequential approach shows the different solution path and a better algorithm convergence.



(a) the initial point from (1.5, 2.2) (b) the initial point from (-1.5, 2.2)

Fig. 1. Solution path with respect to different initial point

4. Optimal control of a CSTR

A continuous stirred tank reactor (CSTR), in which an exothermic, irreversible, first order reaction A → B occurs in the liquid phase and the temperature is regulated with external cooling, is considered for optimal control. This example is taken from Pannocchia and Rawlings (2003), Henson and Seborg (1997) with the assumption that the liquid level is also a state variable. Mass and energy balances lead to the following highly nonlinear state model:

$$\frac{dh}{dt} = \frac{F_0 - F}{p \cdot r^2} \tag{23}$$

$$\frac{dc}{dt} = \frac{F_0(c_0 - c)}{p \cdot r^2 \cdot h} - k_0 \cdot c \cdot \exp(-\frac{E}{R \cdot T}) \tag{24}$$

$$\frac{dT}{dt} = \frac{F_0(T_0 - T)}{p \cdot r^2 \cdot h} + \frac{-\Delta H}{r \cdot C_p} \cdot k_0 \cdot c \cdot \exp(-\frac{E}{R \cdot T}) + \frac{2U}{r \cdot r \cdot C_p}(T_c - T) \tag{25}$$

The controlled variables are the level of the tank, h, and the mole fraction c. The third state variable is the reactor temperature, T. The manipulated variables are the outlet flow rate F and the coolant liquid temperature T_c. Moreover, it is assumed that the inlet flow acts as a disturbance.

We consider an operation situation that, at time t = 10 min a disturbance enters the plant, which is an increment of 10% on the inlet flow F_0. Here the objective is defined as to minimize the offset caused by the disturbance by controlling outlet flow rate F and the coolant liquid temperature T_c over a time horizon of t_f = 50 min. Thus the problem is formulated as follows:

$$\min \int_0^{t_f} \left[(h - h^s)^2 + 100(c - c^s)^2 + 0.1(F - F^s)^2 + 0.1(T_c - T_c^s)^2 \right] dt \tag{26}$$

$$s.t.\ \ DAE \text{ model (23)-(25)} \tag{27}$$

$$85 \le F \le 115 (l/\min); 299 \le T_c \le 301 (K) \tag{28}$$

where h^s, c^s ,F^s,T_c^s are the values at the desired steady-state operating point. This leads to a typical constrained dynamic optimization problem. We divide the time horizon t_f into 50 subintervals, and discretize the problem with the 3-point-collocation in each time interval. The two control variables are represented as piecewise constant. Consequently, there are 2×50=100 control variables and 3×3×50=450 state variables after the discretization.

In implementing the three approaches for comparison, the bounds are imposed only on control constraints. All three approaches, the simultaneous, quasi-sequential and IP quasi-sequential, are coded in FORTRAN (the simultaneous approach directly using the IPOPT Fortran package). At the optimal solution, the three approaches provide the same profiles for control variables as shown in Fig.2. The iterations and CPU time are given in Table 1. The results show that the IP quasi-sequential approach requires less iteration numbers, but much computational cost compared with the quasi-sequential approach. We should point out, however, that the comparison is affected by a number of implementation details. In the quasi-sequential approach, the SQP subroutine DNCONG in the IMSL Library is used as the NLP solver. The IP quasi-sequential approach is coded by the authors, where a lot of implementation parameters should be refined. Moreover, the IPOPT used here is an old version; it is being developed or renewed constantly.

Table 1. Comparison of different approaches to the CSTR problem

Disturbance F_0 (*l/min*)	IP Quasi-Sequential (Iter/CPU(s))	Quasi-Sequential (Iter/CPU(s))	IPOPT (Iter/CPU(s))
100? 110	10/0.7031	17/0.5513	44/1.4219

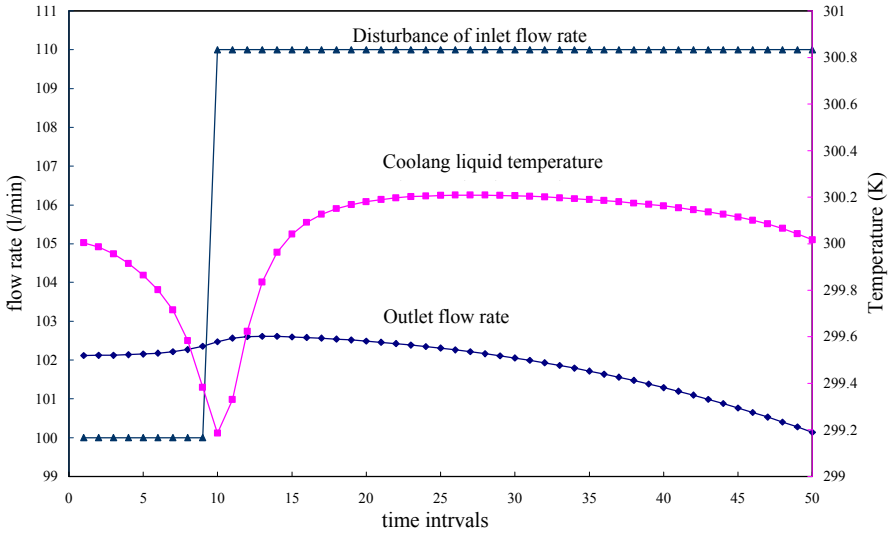


Fig. 2. Optimal trajectories of control variables

5. Conclusions

The quasi-sequential approach would be not very effective for problems with many active constraints, like the most active-set methods for inequality constrained optimization. Here we presented and analyzed an improved quasi-sequential approach, the interior-point quasi-sequential approach, to utilize the ability of interior-point methods for efficient handling a large number of inequalities. The results of the taken examples show that the new approach has advantages of the solution path and higher convergence rate compared with the quasi-sequential approach. But its implementation parameters, such as the update of barrier parameters, should be improved. Finally, the interior-point quasi-sequential approach to handle very large optimization problems with more inequality constraints needs to be studied in the further research.

Acknowledgement

The research was supported by the National Natural Science Foundation of China (No. 20676117).

References

Biegler, L.T., Cervantes, A.M. and Wachter, A., Advances in simultaneous strategies for dynamic process optimization. Chem. Eng. Sci., 57(4): 2002, 575-593.

Biegler, L.T., An overview of simultaneous strategies for dynamic optimization, Chem. Eng. Proc., 46(11): 2007, 1043-1053.

Hong, W., Wang, S., Li, P., Wozny, G. and Biegler, L.T., A quasi-sequential approach to large-scale dynamic optimization problems. AIChE Journal 52(1): 2006, 255-268.

Wächter A. An interior point algorithm for large-scale nonlinear optimization with applications in process engineering. PhD. Dissertation, Carnegie Mellon University, 2002.

Pannocchia, G, Rawlings, J. Disturbance models for offset-free model predictive control. AIChE Journal, 49: 2003, 426-437.

Henson, M.A., Seborg, D.E., Nonlinear process control, Prentice Hall, Upper Saddle River, NJ, 1997.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Branch and Cut Framework for Multi-Stage Stochastic Programming Problems Under Endogenous Uncertainty

Matthew Colvin, Christos T. Maravelias

Department of Chemical and Biological Engineering, University of Wisconsin-Madison, 1415 Engineering Dr, Madison, WI 53706, USA

Abstract

To ensure that decisions in multi-stage stochastic programming (MSSP) formulations do not *anticipate* future outcomes, it is necessary to introduce nonanticipativity constraints (NACs). In the case of endogenous uncertainty, NACs grow very quickly making all but the smallest multi-stage stochastic programming models computationally intractable. To address this challenge, we first present a number of theoretical results that allow us to formulate substantially smaller and tighter MSSP models. Second, we discuss a branch and cut algorithm where necessary inequality NACs are removed from the starting formulation and added only if they are violated. Our theoretical results coupled with the proposed algorithm allow us to generate and solve problems that were previously intractable. The methods were applied to the resource-constrained scheduling of clinical trials in the pharmaceutical research and development pipeline.

Keywords: branch and cut, stochastic programming, endogenous uncertainty.

1. Introduction

Optimization problems under uncertainty can be addressed using different methods, from robust optimization to fuzzy programming, and from simulation-based optimization to stochastic programming (Sahinidis, 2004). In stochastic programming (SP), which we consider in this paper, uncertain parameters are assumed to follow discrete distributions (i.e. each parameter has a finite set of outcomes), and uncertainty is represented via a set of scenarios. In two-stage SP, a subset of decisions are made before the realization of uncertainty (1st-stage decisions) and a subset of decisions play the role of recourse after uncertainty is observed (2nd-stage decisions). In MSSP formulations, uncertainty is observed *gradually* at different stages, leading to a formulation where decisions are also made at different stages. Thus, optimization decisions x_{ts} in MSSP models are defined for all stages, $t \in \mathbf{T}$, and over all scenarios, $s \in \mathbf{S}$. However, decisions made at different scenarios have to be *nonanticipative* of future outcomes. In other words, decisions in two scenarios which at the current stage have identical realizations of uncertain parameters should be the same.

For any pair of scenarios, (s,s'), there exists a stage, $t^{s,s'}$, at which scenarios s and s' become *distinguishable*. Since our knowledge in these two scenarios until stage $t^{s,s'}$ is the same, optimization decisions in scenarios s and s' must be the same, which can be generally expressed using the following logic (nonanticipativity) condition:

$$\left\{t < t^{s,s'}\right\} \Rightarrow \left\{x_{ts} = x_{ts'}\right\} \ \forall t, s, s' \qquad (1)$$

If stage $t^{s,s'}$ is known in advance, the condition in eq (1) can be expressed as: $x_{ts} = x_{ts'}$, $\forall s, s', t < t^{s,s}$, or (Birge and Louveaux, 1997):

$$x_{ts} \sum_{s' \mid t<t^{s,s'}} p_{s'} - \sum_{s' \mid t<t^{s,s'}} p_{s'} x_{ts'} = 0, \quad \forall t, s \tag{2}$$

where p_s is the probability of scenario s. Note that the equalities in eq (2) grow linearly in the number of scenarios and that they can be used during preprocessing for variable and constraint elimination.

When the timing of uncertainty realization is dependent upon optimization decisions, a special case of *endogenous* uncertainty, the condition in eq (1) can be transformed into:

$$-Ry_{tss'} \le x_{ts} - x_{ts'} \le Ry_{tss'} \quad \forall t, s, s' \tag{3}$$

where binary variable $y_{tss'}$ is equal to 1 if scenarios s and s' are distinguishable in stage t, and R is an upper bound on $x_{ts} \ge 0$. The nonanticipativity constraints (NACs) in eq (3) grow quadratically in the number of scenarios, quickly dominating the problem size. Furthermore, a new binary variable has to be introduced to model nonanticipativity.

The goal of this paper is the development of a) theoretical results for the reduction of the NACs in eq (3) and b) a branch-and-cut solution algorithm for MSSP models under endogenous uncertainty in which the timing of uncertainty realization is dependent upon optimization decisions. A thorough discussion of MSSP can be found in Birge and Louveaux (1997), Kall and Wallace (1994), and Ruszczynski (1997). A more thorough discussion of endogenous uncertainty can be found in Jonsbraten et al. (1999).

2. Resource Constrained Scheduling of Clinical Trials

While applicable to a broader class of problems, the methods presented in this paper were developed for and applied to the problem of resource-constrained scheduling of clinical trials in the pharmaceutical research and development (R&D) pipeline and results will be presented in this context.

Given are:

- A fixed time horizon divided into uniform time periods (stages), $t \in \mathbf{T}=\{1, 2, \ldots, T\}$
- A set of candidate drugs, $i \in \mathbf{I}$, with known potential revenues (if successfully developed)
- A set of limited resources, $r \in \mathbf{R}$.
- A set of clinical trials, $j \in \mathbf{J} \subseteq$ {PI, PII, PIII}, for each drug with known cost, duration, resource requirements, and probability of success; clinical trials have to be carried our in sequence PI $\rightarrow$ PII $\rightarrow$ PIII.

Our goal is to determine if and when each clinical trial should be performed so as to maximize the expected net present value (ENPV) of the R&D pipeline. It is assumed that when a clinical trial fails, development of the drug is halted and all investment in the drug is lost. Further it is assumed that the success or failure of each clinical trial is the only source of uncertainty.

The above problem is an optimization problem under endogenous uncertainty because our decisions determine when clinical trials are carried out and thus when uncertainty (clinical trial outcomes) is observed. For each drug $i \in \mathbf{I}$ there are up to four possible outcomes: it fails PI trials (PI-F), it fails PII trials (PII-F), it fails PIII trials (PIII-F), and it passes all trials (PIII-P). Therefore, uncertainty can be expressed in terms of a single parameter, ξ_i, with ordered sample space $\Omega_i \subseteq$ {PI-F, PII-F, PIII-F, PIII-P}. The problem at hand can be formulated as a MSSP model with a single set of independent

variables, X_{ijts}, which are equal to 1 if clinical trial (i,j) is started at stage t in scenario s. Since the timing of uncertainty observation is dependent upon our decisions, inequality NACs of the form of eq (3) must be enforced for all clinical trials, in every stage, and between all pairs of scenarios. A detailed problem statement and MSSP formulation can be found in Colvin and Maravelias (2008).

In a *naïve* MSSP formulation, an instance containing three drugs would require ~290,000 NACs compared with ~4,100 other constraints. For an instance containing six drugs these numbers grow to ~1.2×10^9 and ~250,000, respectively. Clearly, problems of this size cannot be addressed by existing optimization tools.

3. Results for Nonanticipativity Constraints

In this section, we present a number of properties that enable the formulation of substantially smaller and tighter MSSP formulations of problems with endogenous uncertainty in which the timing of uncertainty realization is dependent upon optimization decisions.

Property 1. *It is sufficient to express NACs only for pairs of scenarios* (s,s') *that differ in only a single uncertain parameter* ξ_i. .

This result is based on the work of Goel and Grossmann (2006) and can be shown using transitivity. In the case of the drug development problem, this property says that if the realizations in scenarios s and s' are different for more than one drug, then it is unnecessary to enforce NACs between s and s'. The drug whose outcome is different in scenarios s and s' will be referred to as the *critical drug*.

Property 2. *If a set of decisions must be carried out in a predefined sequence and each decision in this set can lead to the observation of a single outcome for each uncertain parameter, then it is sufficient to express NACs for scenarios* (s, s') *that differ in the outcome of a single parameter* ξ_i *and the outcomes of* ξ_i *in scenarios s and s' are associated with decisions that must be made consecutively in the sequence.*

This property leads to improvements in a smaller class of problems, in which the timing of decisions is unknown but certain decisions have to be made in a specified sequence. For example, in the planning of R&D activities, clinical trials must be carried out sequentially (PI $\rightarrow$ PII $\rightarrow$ PIII) and the outcomes after the completion of PI trials can be split into two sets {PI-F}, which can be observed, and {PII-F, PIII-F, PIII-P} which cannot be observed until additional trials are carried out. Thus, if i^* is the critical drug for scenarios s and s', then it is sufficient to express NACs between s and s' only if the outcomes of ξ_{i^*} in s and s' are consecutive elements in Ω_{i^*}.

Properties 1 and 2 allow us to apply NACs for a reduced set of scenario pairs, which we will refer to as $\mathbf{\Psi}$. The *association* of a decision variable with a specific outcome and thus the differentiation of a specific pair of scenarios leads to the following property that enables the formulation of NACs using substantially fewer variables:

Property 3. *If the stage at which scenarios s* and *s' become distinguishable is a function of a single optimization decision, then the logic condition in eq (1) can be transformed into mixed-integer constraints in terms of decisions* x_{ts} *only.*

In the scheduling of clinical trials, the completion of PI trials can only lead to the observation of outcome PI-F, because if PI is successful then we should carry out more trials before we observe another outcome. The trial that differentiates scenarios s and s'

will be referred to as *critical trial* and denoted by ($i^{s,s'}, j^{s,s'}$). Using Properties 1-3, we can express NACs as follows:

$$-\sum_{t'\le t-\tau_{i^{s,s'}j^{s,s'}}} X_{i^{s,s'}j^{s,s'}t's} \le X_{ijts} - X_{ijts'} \le \sum_{t'\le t-\tau_{i^{s,s'}j^{s,s'}}} X_{i^{s,s'}j^{s,s'}t's}, \quad \forall i,j,t,(s,s') \in \mathbf{\Psi} \qquad (4)$$

where τ_{ij} is the duration of trial (i,j). The summation of $X_{i^{s,s'}j^{s,s'}t's}$ variables is equal to 1 if the critical trial ($i^{s,s'}, j^{s,s'}$) is finished before stage t, which would mean that scenarios s and s' are distinguishable, and therefore the corresponding decisions can be different. Note that NACs are expressed for a subset of scenario pairs (Properties 1 and 2) without introducing new $y_{tss'}$ decision variables (Property 3).
Another implication of the sequencing of decisions is that decisions cannot be carried out before all prerequisites are completed, which leads to the following two results:

Property 4. *If the stage at which scenarios s and s' become distinguishable depends on a single decision which cannot be made before stage* $t^{s,s'}_{\min}$*, then nonanticipativity between scenarios s and s' in stages* $t < t^{s,s'}_{\min}$ *can be enforced using eq. (2).*

Property 5. *Non-anticipativity between s and s' for decisions that precede the decision that leads to the realization of uncertainty that distinguishes s and s' can be enforced using eq (2).*

Properties 4 and 5 lead to the replacement of a large number of double inequalities of the form of eq (3) by equalities of the form of eq (2). This leads to smaller and tighter formulations.
Finally, for all formulations employing eq. (2) with only binary decision variables, we can further reduce the number of equality NACs:

Property 6. *Let* $\mathbf{S}_t$ *be a set of scenarios that are necessarily indistinguishable at stage t, i.e.* $\mathbf{S}_t = \{s: t^{s,s'} < t, \forall s' \in \mathbf{S}_t\}$*. If* $p_s > 0$ $\forall s \in \mathbf{S}_t$ *and* $x_{ts} \in \{0,1\}$*, then NACs among all scenarios in* $\mathbf{S}_t$ *can be enforced using a single equality of the form of eq. (2).*

Nonanticipativity for decisions in scenarios $s \in \mathbf{S}_t$ is typically enforced using $|\mathbf{S}_t|$ equalities (note that eq (2) is expressed for every s). Property 6 says that in the case of binary variables nonanticipativity for decisions in scenarios $s \in \mathbf{S}_t$ can be enforced using a single equation. It is important to note here that Property 6 is widely applicable because eq (2) is used to enforce nonanticipativity in all SP models.
The reduction in the number of NACs due to properties 1, 2, and 4-6 is demonstrated in Figure 1. For an instance with six drugs number of NACs is reduced from ~1.2×10^9 to ~1.5×10^6. Furthermore, the growth in NACs is almost linear in the number of scenarios. Nevertheless, NACs could still represent the majority of all constraints.

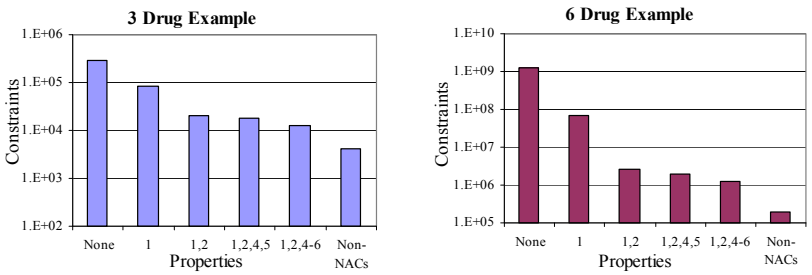


Figure 1: Reduction in NACs due to Properties 1, 2, 4, 5, and 6.

4. Branch and Cut Algorithm

Even though none of the NACs is redundant, it can be shown that only a small fraction of these constraints are active in any feasible solution. In the drug development problem in particular, it can be shown that at most 12.5% of inequality NACs can be violated in any solution, and in practice less than 4% of these constraints were violated. These observations led us to explore the development of a branch and cut (BaC) algorithm in which the initial formulation does not include any NACs represented by inequalities, though equalities (see Properties 4-6) are included. When inequality NACs are found to be violated, they are added into the formulation and kept in all descendant nodes. Note that unlike typical branch and cut algorithms, we removed *essential*, not tightening constraints.

Removing essential constraints from the formulation allows for integer solutions that are infeasible to the full model to be found feasible. To handle this, a number of modifications to the standard BaC algorithm were made. First, the heuristic solution strategy was turned off to prevent better than optimal, but infeasible, solutions to be found and used as lower bound for pruning. Second, bound updating was modified because the lower bound cannot be updated immediately upon finding an integer solution (the solution must first be checked for NACs feasibility). Finally, if an integer solution is found to violate removed NACs, it must be resolved after the addition of violated NACs (possibly leading to a fractional solution), and subsequently partitioned using standard branching.

Having successfully implemented our basic algorithm, we then sought to improve its performance in two ways: node selection rules and selection of nodes where NACs are checked. Note that in our algorithm, local search (which is the default node selection method) loses two key advantages. First, the advantage of having the basis of the previous node is diminished because hundreds or thousands of violated NACs are added, especially in early nodes. Second, with the heuristic solution turned off, a lower bound is not in place to quickly identify unproductive areas of the tree. In later nodes where fewer NACs are added and a lower bound is available, the advantages of local search are important. We found using a best first search for a fixed number of nodes, N^{max}, before using local search provided for the fasted solution times. While the optimal N^{max} shifted for each problem, it was found to be roughly proportional to the number of drugs in our example problems.

Related to the node selection rules is the decision to test for violated NACs at all nodes or at integer nodes only. Checking for infeasibilities at all nodes increases the computational cost per node substantially (because the solution is checked against a large number of NACs), but reduces the total cut additions because a cut added at an early fractional node is carried to all descendent nodes reducing the number of cuts that are added multiple times at different nodes. Testing for infeasibilities at every node outperformed testing only at integer solutions; however, when used in conjunction with our hybrid node selection rule, adding cuts in all early nodes, but only in integer nodes after a fixed depth, D^{max}, seemed to perform better than either pure approach.

5. Computational Results

The proposed algorithm was implemented in Xpress-IVE and tested on a number of instances varying in the number of drugs and scenarios. Model and solution statistics for the full-space MSSP formulation and the proposed algorithm are given in Table 1.

Table 1: Model statistics and computational results.

Example	3-drug	4-drug	5-drug	6-drug	7-drugA	7-drugB
Scenarios	64	144	288	864	2304	3456
Variables	4,512	6,192	14,112	69,216	152,256	237,600
Full-space model						
Constraints	16,994	27,404	67,871	403,204	942,617	1,526,773
Equality NACs	4,989	7,010	14,759	81,930	161,454	289,357
Inequality NACs	7,856	14,838	41,676	267,040	670,295	1,059,648
LP Relaxation ($10^6)	867.5	2,343.60	3,403.90	3,993.80	- *	- *
Solution ($10^6)	827.9	2,166.60	3,364.60	3,974.10	-	-
CPU sec	310.4	418.6	667.1	16,429	-	-
Branch-and-cut: starting model statistics						
Constraints	9,138	12,566	26,195	136,164	272,322	467,125
LP Relaxation ($10^6)	872.3	2,368.40	3,405.80	4,021.10	5,183.20	4,431.00
Solution ($10^6)	827.9	2,166.60	3,364.60	3,974.10	5,055.10	4,239.80
Customized branch-and-cut: computational results						
Nodes Searched	7	8	1	5	26	14
Cut Additions	852	1,832	284	5,096	18,948	16,206
CPU sec	84.3	323.1	438.2	2,781	116,605	280,193

* Model could not be generated.

First, note that the LP relaxation of the starting formulation for the BaC algorithm is almost as tight as the one of the full-space model despite removing a large number of NACs. Second, all instances are solved faster using the BaC algorithm, but more importantly, problems that were unsolvable because the corresponding MSSP model could not be generated can now be addressed.

6. Conclusions

We have presented a set of theoretical results and a novel branch and cut algorithm for the solution of MSSP formulations under endogenous uncertainty in which the timing of uncertainty realization is dependent upon optimization decisions. The proposed methods address the explosion in the number of NACs in stochastic programming formulations. They can be applied to a wide range of problems under uncertainty.

Acknowledgements

The authors of this paper would like to acknowledge support from the National Science Foundation under Grant CTS-0547443.

References

Birge, J.R., Louveaux, F., 1997, *Introduction to Stochastic Programming.* New York: Springer.

Colvin, M., Maravelias, C.T., 2008, A Stochastic Programming Approach for Clinical Trial Planning in New Drug Development, *Computers & Chemical Engineering,* 32, 2626-2642.

Goel, V., Grossmann, I. E., 2006, A Class of Stochastic Programs with Decision Dependent Uncertainty, *Mathematical Programming, Ser. B* 180, 355-394.

Jonsbraten, T. W.; Wets, R.J.B.; Woodruff, D. L., 1998, A class of stochastic programs with decision dependent random elements. *Annals of OR*, 82, 83-106

Kall, P., Wallace, S.W., 1994, *Stochastic Programming*, Chichester, Wiley.

Ruszczynski, A., 1997, Decomposition Methods in Stochastic Programming, *Mathematical Programming,* 79, 333-353.

Sahinidis, N. V., 2004, Optimization under uncertainty: State-of-the-art and opportunities. *Computers and Chemical Engineering*, 28 (6-7), 971-983.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

ProSEC : Modelling and Simulation in 3D of Brazed Aluminium Core-in-Drum Plate-fin Heat Exchangers

Florian Picard[a], David Averous[a], Xavier Joulia[b], Denis Barreteau[b]

[a]*Fives Cryo, 25Bis, rue du fort, B.P. 87, F-88194 Golbey Cedex, France*
[b]*ENSIACET, 118 route de Narbonne, 31077 Toulouse CEDEX 4, France*

Abstract

A model for steady-state simulation in 3D of a Core-in-Drum plate-fin heat exchanger is presented in this paper. The model takes into account both the geometrical description of the PFHE, layer by layer, and the description of the external piping layout of each stream outside the core. Several steps that allow to define the entire Core-in-Drum model are presented in this case study, with in particular the mesh generation, the topological description of stream circulation, calculation constraints and, to finish, detailed results.

Keywords: Plate-Fin Heat Exchanger, cryogenics, 3D hydrothermal modelling, 3D simulation, maldistribution.

1. Introduction to Core-in-Drum Plate-Fin Heat Exchangers

1.1. General points about Plate-Fin Heat Exchangers (PFHE)

PFHE are widely used in process industries such as gas processing and petrochemical industries including cryogenic applications: industrial gas and hydrocarbon separation are the most common, but many processes use this type of equipment, including recovery of natural gas liquids, helium refrigerators and liquefiers, hydrogen purification, ammonia and ethylene processes, nuclear engineering and syngas production. Maximizing thermal efficiency with high reliability is the main purpose for such cryogenic processes. PFHEs promote exchange between many streams simultaneously (cases with more than 12 streams are common) greatly increasing energy efficiency.

1.2. Study case of a Core-in-Drum Exchanger

Most Core-in-Drum heat exchangers are used as vaporizers, reboilers, chillers or condensers to be in general associated with distillation columns. The basic principle of such heat exchangers is to permit exchange between two streams, including phase-change(s), in crossflow configuration. Figures 1 and 2 show respectively the mechanical description and the functioning principle of a Core-in-Drum PFHE.

The cold stream (called "PROP", see figures 1 and 2) enters the shell by a nozzle (item *4*) as a liquid or two-phase and submerges the core (item *6*) partially or completely. The warm stream (here called "OVHD"), fed to the core by an inlet nozzle (item *1*) and a header (item *2*), flows horizontally, away from the inlet, and cools against the vertical cold stream which vaporizes and rises by the thermosyphon effect through devoted free-opened layers, to finish with exiting through several nozzles (items *5*) . The table 1 summarizes the functioning parameters of this study case.

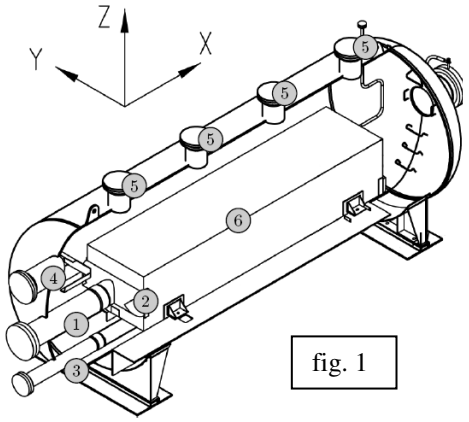


fig. 1

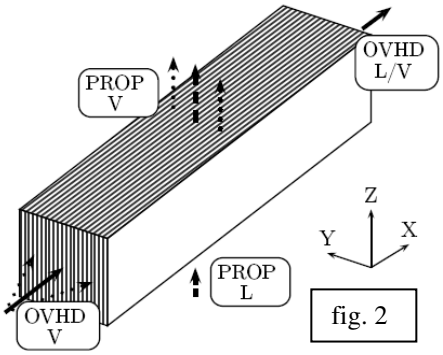


fig. 2

tab. 1

	unity	Warm stream	Cold stream
		OVHD	PROP
Gas	$kg.h^{-1}$	163109	0
Liquid	$kg.h^{-1}$	0	153250
Temperature	°C	-32.9	-35.7
Pressure	Bar	18.27	1.72
Layer		A	B
Number of layers		85	84

Figure 3 illustrates the internal structure of the two layer patterns A and B, which define respectively the OVHD and PROP circuits. The stream OVHD enters each suitable layer of the stacking arrangement by a central distributor, flows through straight fins, and finally exits by an end distributor. The stream PROP is drawn to the top through fins, which are separated by bars along the width of the layer.

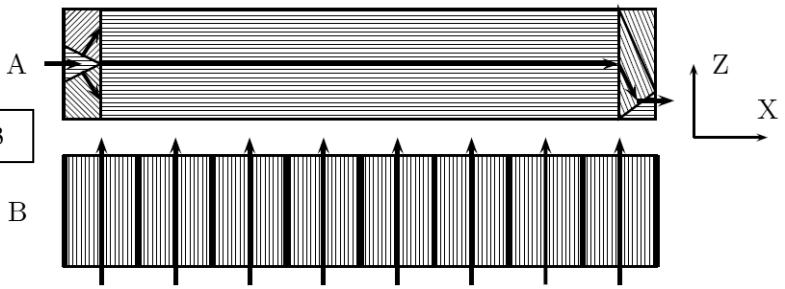


fig. 3

2. Definition of the Core-in-Drum model

The principal interest of this three dimensional model in steady-state is that the influence of flow distribution, induced by distributors and in particular by the cross-flow configuration, is integrated into the total thermal performance of the PFHE. It may result in flow imbalance (maldistribution) along the width of each circuit, disrupting the thermal behaviour of the PFHE because of the high connection between thermal and flow variables.

2.1. Meshing method

Figure 4 illustrates the specific mesh operations for the two layers. Grid is defined in order to take into account the coupling of thermal and hydraulic variables. About 220000 cells have been used to mesh the entire core-in-drum (plates and layers).

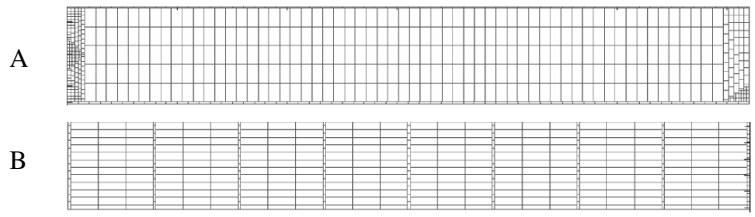


fig. 4

2.2. Flow management

The PFHE can be looked upon as a flowsheet, including specific operation units which are connected together through a flow network. All external components (manifold, nozzles, header,…) and all the meshes used to describe the internal flow circulation in each layer of the stacking pattern are integrated. Thus, the flow circulation for each stream is modelled by a devoted 3D graph, which is closely linked to each others by thermal consideration due to the particular structure of PFHE. Such a complexity needs powerful tools in order to represent the entire path followed by each stream outside the core and between covered layers. Graph theory is very useful for modelling this topological description. Figure 5 shows simply the connexions between the eight circuits of only one layer defining the PROP circulation in the stacking arrangement.

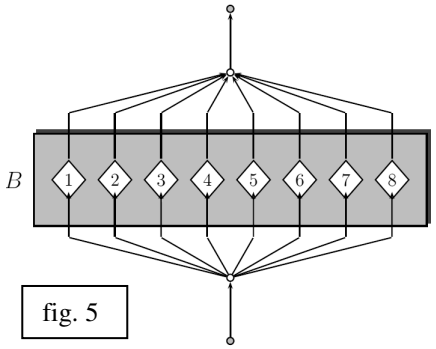


fig. 5

2.3. Calculation

All variables which define streams and all components (plates, bars) in the system are highly interconnected, increasing the complexity of the full model of the PFHE. Several million equations can be gathered in model of such a PFHE. Fortunately, each equation of the model involves few variables. Finding the solution of such a large and sparse system requires adopting specific numerical methods and tools, in particular to preserve memory requirement, improve calculation time and guarantee the quality of the solution. Consequently it would be inefficient to adopt a sequential modular strategy rather than a global approach.

Moreover, it is numerically difficult to resolve an energy conserving temperature and vapor fraction field when consecutive iterations gravitate between single and two-phase regions, because there can be an order of magnitude difference in the values of heat transfer coefficients being used to calculate the energy coefficient matrix for each iteration.

3. Results

Overall results are gathered in the table 2. The mass vapor fraction of the cold stream PROP at the exit of the core is about 0.535, while the warm stream OVHD is partially liquefied with a residual mass vapor fraction of 0.498. Notice that the pressure drop for each stream in each layer is low.

	unity	Warm stream	Cold stream
Flow	$kg.h^{-1}$	163109	153250
Mass vapor fraction	-	0.498	0.535
Temperature	°C	-33.9	-37.1
Pressure drop	mbar	11.32	9.82
Total power	MW	39.62	39.62

Outlet specifications by stream

As it is shown in figure 7, PROP temperature profiles may seem surprising : because the OVHD warm stream enters perpendicularly at the edge of the core (fig. 6), the temperatures there are expected to be significantly higher in this area than the temperatures in the rest of the core. It would seem that heat transfer between OVHD and PROP is reduced in this area.

fig. 6

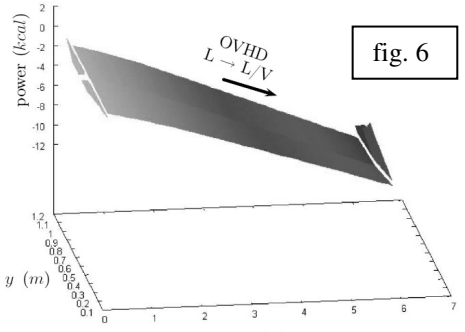
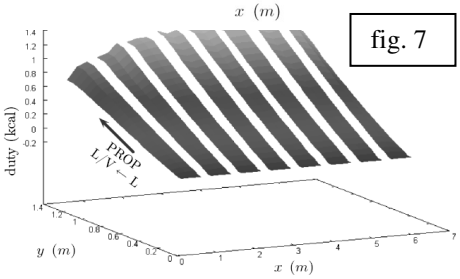


fig. 7



To understand this notable feature, some clue may be obtained by representing the PROP stream profiles of the relative difference between the mass flow rate calculated at a given x and the theoretical average mass flow rate evaluated by assuming a homogeneous distribution of the flow along the entire width (fig. 8). In spite of the geometric uniformity, coolant mass flow rate is not uniform across the core layer. There is a maximum flow maldistribution rate of 25%. Note that the largest flow maldistribution corresponds to the temperature imbalance described in the previous paragraph.

Consequences on the performance may be analyzed in figure 9, in which the heat transfer rate profile for PROP stream is represented. This curve is perfectly similar to the PROP mass flow rate field. Thus, the flow distribution becomes heavily involved in the thermal performance. However, the outlet vapor mass fraction, represented in the same figure, is almost continuous (0.536). The maximal gap between the two extremes of the curve is 2.3%.

fig. 8

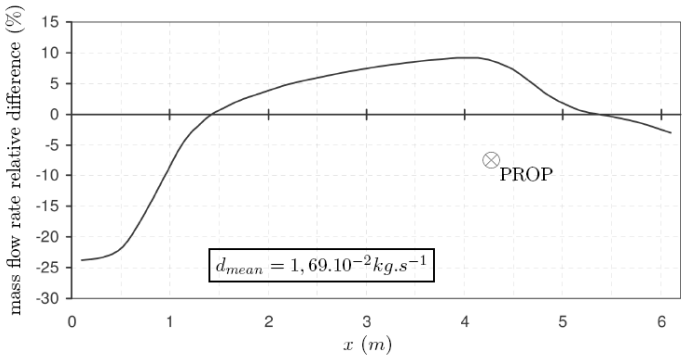
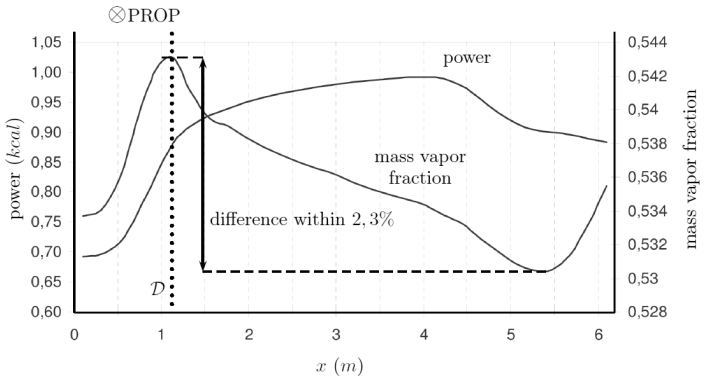


fig. 9



Curves plotted on the figure 10 characterize the width-wise pressure drop contributions due to both gravity and friction drag applied on the PROP stream. The gravitational term is the main constraint imposed on the vertical flow circulation (from 95.2 to 96.7 percent of the overall pressure drop) compared to the friction loss term. Physics impose the equality of the pressure drop between each vein of PROP flow, in such a way that inlet and outlet pressures are respectively uniform on each side of the core. However, thermal properties of each PROP vein evolve individually due to the changing neighboring thermal conditions imposed by OVHD thermal fluctuations. Finally, the pressure drop equilibrium along the width of the layer is obtained by the coupled compensation of frictional and gravitational terms, which allows the regulation of the flow rate.

fig. 10

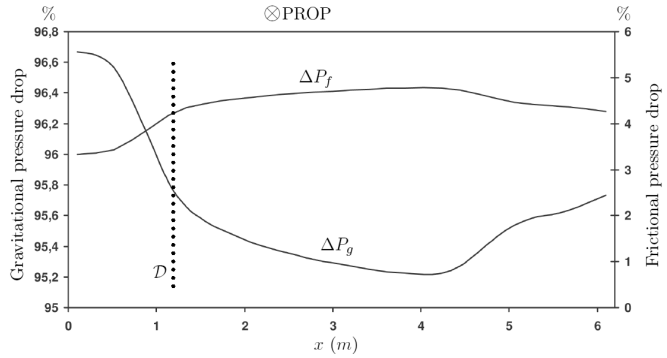
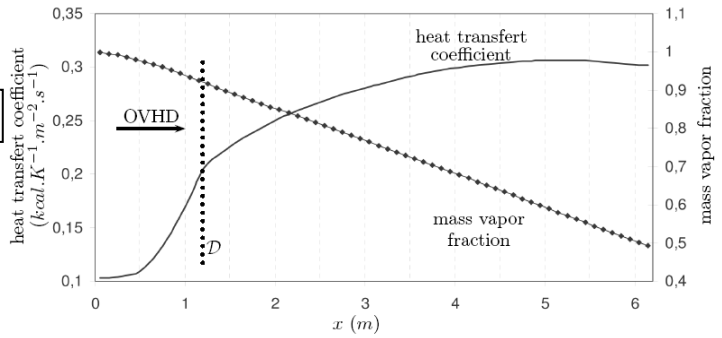


fig. 11

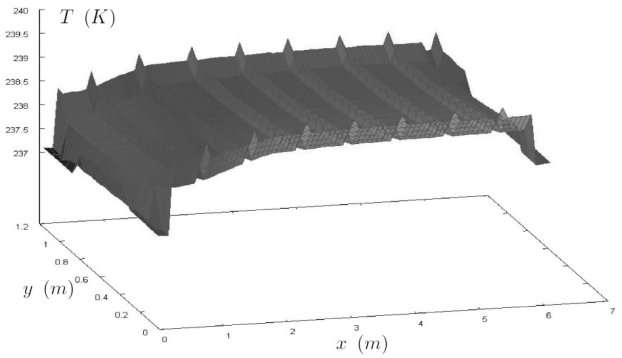


The study of the OVHD stream provides some indications concerning the understanding of the variations of mass flow and heat transfer rates of PROP stream. Figure 11 shows the heat transfer coefficient of the OVHD stream along the flow direction, with respect

to x. Logically, the heat transfer coefficient is correlated with the mass vapor fraction : this factor remains low for values of mass vapor fraction higher than 0.9 and increases regularly with a decrease of the mass vapor fraction. It is possible to visualize two distinct "rates" of heat transfer, marked by an inflexion point (at around 0.9) which is materialized by the dotted line on the figure.

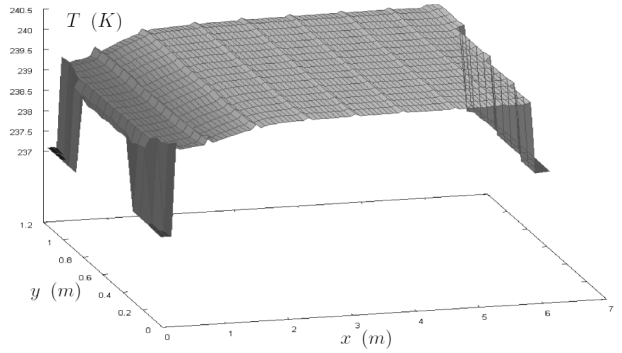
The figure 12 shows the thermal profiles of a parting sheet of the stacking arrangement, sandwiched between two distinct layers A (above) and B (below), in which circulate streams OVHD and PROP respectively. The plate closely retranscribes thermal phenomena induced by the flow circulation in each specific layer geometry. Separating-bars and dead-zones of the two distributors could be considered thermal wells, by which transit preferentially the energy.

fig. 12



The figure 13 illustrates the thermal profile of the top cap-sheet, which is directly in contact with the OVHD stream. It is interesting to observe the influence of dead zones on the two distributors implied in the description of the layer A.

fig. 13



4. Conclusion

3D modeling is a highly improved method for PFHE study. The coupling of thermal and hydraulic parameters on the global functioning of the core can be easily visualized and provides significant clues to understand such a complex system. A specific meshing grid method has been developed to take into account all kind of pattern geometries, which constitute the entire stacking arrangement. Moreover, the complex description of such PFHE involves a perfect management of flow distribution, thanks to graph theory. But calculations of such large system of equations require appropriate numerical resolutions, involving in particular numerical methods for sparse systems.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Sensitivity analysis of non-linear dynamic models: Prioritizing experimental research

Gürkan Sin[a], Anna Eliasson Lantz[b], Krist V. Gernaey[c]

[a]*CAPEC-Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, DK-2800 Kgs. Lyngby, Denmark.*
[b]*Center for Microbial Biotechnology, Department of Systems Biology, Technical University of Denmark, Building 223, DK-2800 Kgs. Lyngby, Denmark*
[c]*BIOENG-Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, DK-2800 Kgs. Lyngby, Denmark.*

Abstract

This study demonstrates the use of (global) sensitivity analysis as a supporting tool for guiding and prioritising experimental research activities. The basic idea is to focus the experimental efforts on those factors identified as the most significant by the sensitivity analysis. Such an approach is ideal for identifying critical regions in the experimental design space (typically multi-dimensional). These critical regions can later be investigated experimentally by means of the well-established Design of Experiments methodology. The sensitivity-analysis based approach promises better use of resources by preventing wastage of valuable experimental time on investigating factors which are not influential on the system behavior. The approach requires a first-principles model of the system on which the sensitivity analysis is performed, and was tested on a dynamic fermentation model describing antibiotic production by *Streptomyces coelicolor*.

Keywords: *Sensitivity, uncertainty, fermentation, first-principles, SRC, experiment*

1. Introduction

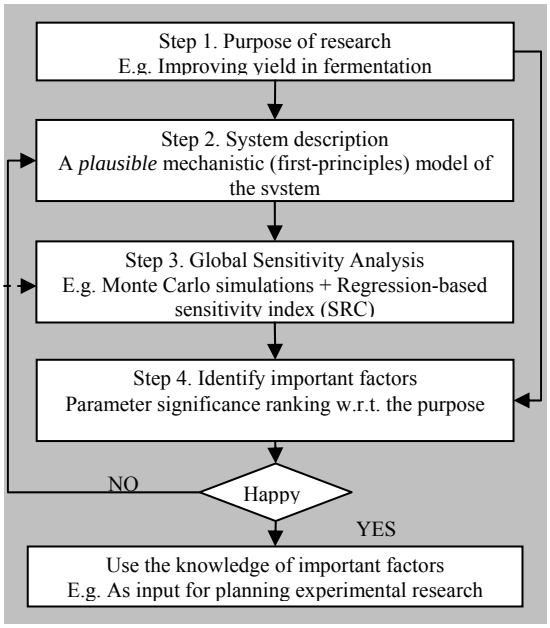
In a general sense, uncertainty exists in the predictions of any numerical model. This uncertainty stems from, among others, the numerous assumptions and simplifications made during the model-building process. Uncertainty analysis is complemented by sensitivity analysis, which aims at quantifying which of the model assumptions/input parameters causes the observed uncertainty in model predictions. The result of sensitivity analysis is a parameter significance ranking on model predictions. Such information is valuable for a number of purposes: (i) identify critical regions in the input parameter space, (ii) defend the credibility of a model, (iii) test hypotheses (Saltelli *et al*., 2008; Helton and Davis, 2003).

Traditionally sensitivity analysis is performed using differential analysis –typically by taking the first order derivative of a model output with respect to a parameter. This method however is local and may not be applicable for non-linear models – representing the majority of the engineering models. In a previous study, Sin *et al*. (2009) evaluated global methods for sensitivity analysis, namely regression-based (i.e. standardized regression coefficients, SRC) and screening-based (i.e. Morris method) approaches. The results showed that the global methods particularly perform better with respect to assessment of the significance of model parameters on the model predictions, especially when the degree of non-linearity of the model in question is high.

Another important finding of this study is that the sensitivity methods identified only few (5 to 7) out of a total of 56 parameters to cause most of the variance (up to 95%) in the predictions. Based on these results, we suggest using sensitivity analysis for better directing experimental resources in such a way that only important factors (e.g. what causes variation in the product quality, what influences the yield etc.) are investigated. The aim of this contribution is thus to develop a methodology that uses global sensitivity analysis methods for establishing priorities in experimental research.
The proposed method is different from the traditional Design of Experiment (DoE) approaches (Mandenius and Brundin, 2008), which focus typically on a few factors and mostly use a statistical model to understand the interactions between the factors. Our method is an integrated approach which systematically screens all possible factors involved in a systems' behavior using a mechanistic (knowledge-based) model.

2. Materials and methods

2.1. Global Sensitivity Analysis for prioritizing experimental research
The framework in which sensitivity analysis is used for prioritizing the experimental efforts is shown in Figure 1.



The first step is obviously specifying the objective of the experimental research work. This could be a specific objective like improving the yield in a fermentation process. The second step is to obtain a mechanistic model describing the system behavior; the only requirement is that the model should adequately describe interactions in the system, e.g. physical, chemical and biological reactions for a fermentation system. In step 3, the global sensitivity analysis is performed using an appropriate method (Helton and Davis, 2003; Saltelli *et al*., 2008; Sin *et al*., 2009).

Figure 1 Sensitivity analysis based methodology for prioritizing experimental research

In step 4, the sensitivity results are interpreted and the important factors affecting the purpose are identified. The results are then cross-checked with the available process knowledge in order to pinpoint any possible discrepancy. In case the evaluation is positive, the experimenter can then move on and use the knowledge in his/her planning of the experimental work.

2.2. Case study: Batch cultivation of Streptomyces coelicolor for antibiotic production
As a case study, batch cultivation of *S. coelicolor* for antibiotic production was used. The model predicts glucose, oxygen, biomass, actinorhodin, and off-gas CO_2 among others, and generally describes the system behavior adequately. The model contains 56

parameters relating to biological, chemical and gas-liquid exchange processes in the system. For more information and details, the reader is referred to Sin *et al*. (2008).

2.3. Regression-based sensitivity analysis

2.3.1. Standardized Regression Coefficient

In this method, the sensitivity measures are in fact the regression coefficients obtained from the linear regression models built on the outputs generated by Monte Carlo (MC) simulations (for more information see below). This method requires scalar output, hence for the outputs of the dynamic models one needs to use proper scalar values, e.g. mean value of a dynamic output. Let the scalar model output matrix be denoted as **y** with a dimension of *KxN,* in which K is the number of outputs and N is the number of MC simulations. With that output matrix one multi-linear regression model was built for each model output k:

$$y_{ik} = b_{0k} + \sum_{j=1}^{M} b_{jk}\theta_{ij} + \varepsilon_{ik} \quad for\ i = 1,2,\ldots N \quad and \quad for\ k = 1,2....K \qquad \text{Eq. 1}$$

y_{ik} is the scalar value for the kth output, b_{jk} is the coefficient of the j^{th} input factor on y_k, θ_{ij} is the value of the j^{th} factor, M is the number of factors, and ε_{ik} is the error of the regression model. The eq.1 is re-written by scaling the outputs and the factors using the corresponding mean (μ) and standard deviation (σ) (Saltelli *et al*., 2008):

$$\frac{y_{ik} - \mu_{y_k}}{\sigma_{y_k}} = \sum_{j=1}^{M} \beta_{jk} \cdot \frac{\theta_{ij} - \mu_{\theta j}}{\sigma_{\theta j}} + \varepsilon_{ik} \qquad \text{Eq. 2}$$

β_{jk} is called the standardized regression coefficient (SRC) of the j^{th} input factor on the k^{th} model output, y_k. Its magnitude is a measure of the sensitivity of the output to the factor.

Table 1 Expert review of uncertainty of input parameters (Sin *et al*., 2008)

Uncertainty Class	**Variation factor**[1]	**Parameters** (legends explained in Sin *et al*. (2008))
1. Low	5%	Y_{SX}, MW_X, MW_S, γ_S, MW_{act}, γ_{ACT}, i_{NRED}, MW_{red}, γ_{RED}, T, pH, pK_{NH}, pK_{HCO3} pK_{H2PO4}, pKw, $k_{f,NH}$, $k_{f,CO2}$, $k_{f,H2PO4}$, $k_{f,W}$, R, V_L
2. Medium	25%	f_{XI}, i_{NX}, i_{PX}, γ_X, μ_{max}, m_S, k_d, t_{lag}, K_La_{O2}, P_{in}, P_{O2}, P_{CO2}, P_{out}, D_{O2}, D_{CO2}, D_{N2}, K_{HO2}, K_{HCO2}, K_{HN2},$V_{G,}$ Q_{Gin}
3. High	50%	K_S, K_O, K_N, K_P, Y_{SACT}, β_{ACT}, α_{ACT}, S_{ACT}^{max}, Y_{SRED}, β_{RED}, α_{RED}, S_{RED}^{max}, S^*_{NH3}, K_La_{NH3}

[1]Stands for the maximum percent of deviation from the mean value of a parameter.

2.3.2. Monte Carlo simulation

MC simulation involves 3 steps: (1) specifying input uncertainty (2) sampling from the input uncertainty and (3) performing simulations with the sampled inputs (Helton and Davis, 2003). For step 1, the input uncertainty was specified following an expert review process and the results shown in Table 1. For step 2, Latin Hypercube Sampling (LHS) was used with the sampling number fixed at 500. For step 3, the model was simulated with 500 different input samples to obtain MC simulations.

3. Results and discussions

We first take a simple purpose in step 1 to evaluate the methodology, by focusing on the factors affecting off-gas CO_2 and dissolved oxygen dynamics in the system. For step 2,

the above-mentioned model is used (Sin *et al*., 2008). The other steps are presented below.

3.1. Step 3 Global sensitivity analysis

3.1.1. Monte Carlo simulations

The MC simulation results (500 in total) of biomass, glucose, oxygen and off-gas CO_2 are shown in Figure 2. Generally speaking the larger the spread of the MC simulations, the higher the uncertainty is on the output in question. From Fig 2, one observes there is considerable uncertainty on the oxygen and off-gas CO_2.This uncertainty on the output is in fact caused by the uncertainty in the inputs (Table 1). The sensitivity analysis now aims at decomposing this uncertainty with respect to factors in the input uncertainty.

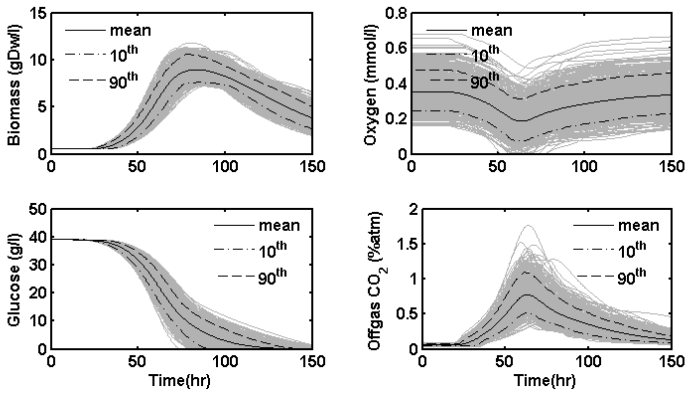


Figure 2 500 Monte Carlo simulation results for biomass, glucose, oxygen and off-gas CO_2 in a batch cultivation of *S. coelicolor* (see legend on the plots)

3.1.2. Standardized regression coefficients (SRC)

The SRC of the factors for off-gas CO_2, oxygen, glucose and biomass are shown in Table 2. The SRC take values between -1 and 1. The absolute value of the SRC indicate how significant each factor is on the output, whereas the sign indicates a positive or a negative impact of the factors on the output in question.

Table 2 Ranking importance of factors on two model outputs: oxygen and off-gas CO_2

	Off-gas CO_2		**Oxygen**		**Biomass**		**Glucose**	
Rank	Parameter	SRC	Parameter	SRC	Parameter	SRC	Parameter	SRC
1	Pout	0.58	P_{O2}	0.59	m_S	-0.67	t_{lag}	0.74
2	Q_{Gin}	-0.54	Pin	0.58	k_d	-0.42	i_{PX}	0.66
3	Pin	-0.53	K_{HO2}	0.57	i_{PX}	-0.37	μ_{max}	-0.63
4	P_{O2}	-0.11	K_La	0.12	t_{lag}	-0.26	K_P	0.54
5	V_L	0.11	μ_{max}	-0.10	Y_{SX}	0.21	pK_{H2PO4}	-0.40
6	R	0.11	K_P	0.06	pK_{H2PO4}	0.19	m_S	-0.36
7	i_{PX}	-0.08	t_{lag}	0.03	K_P	-0.13	K_S	0.12
8	m_S	0.07	γ_S	-0.03	μ_{max}	-0.11	Y_{SX}	0.12
9	P_{CO2}	0.07						

3.2. Step 4-Identification of significant factors

One observes that the majority of the factors in fact has negligible impact on the outputs, while only few have a significant influence (e.g. 9 factors have SRC value higher than ±0.03 on the off-gas CO_2 and 8 factors on the oxygen in the fermenter). The significant factors affecting the off-gas CO_2 and oxygen are mainly related to the aeration process (partial pressure of oxygen and CO_2, air flow-rate, etc) and to a lesser

extent to the microbial metabolism (Sin *et al*., 2009). On the other hand, the significant factors affecting the glucose and biomass in the system are mainly related to microbial metabolism (growth rate, maintenance, etc). This sort of information is crucial in focusing the experimental efforts in line with the purpose of the experiment (see below).

3.3. Step 5-Experiments (verification of the concept)

Since the purpose of the experiment was set to investigate off-gas CO_2 and oxygen in the fermentation system, the information provided above by importance ranking of factors thus helps narrowing down the experimental degrees of freedom to focus only on the aeration process in the system.

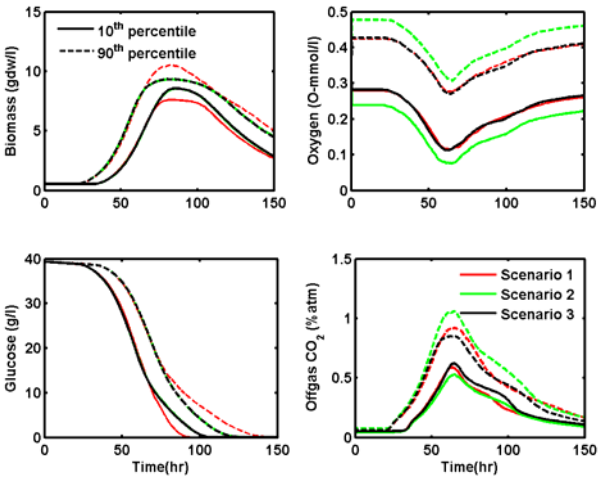


Figure 3 Virtual experiments: uncertainty reduction in three scenarios (see text)

To verify this concept, we do a simulation study using the above described model (Sin et al 2008). We assume that three experiments have been performed with the system (i.e. three hypothetical experiments) that focused on measuring the parameters of (i) the aeration (P_{out}, Q_{Gin}, P_{in}, P_{O2}, P_{CO2}, K_La), (ii) the biomass metabolism (i_{px}, m_S, μ_{max}, K_P), or (iii) both. One also assumes that the experiments are of good quality, hence providing measurements with only a 5% error. With that information, we update the input uncertainty in Table 1 (by assigning class I uncertainty to the measured parameters), and perform 3 sets of MC simulations. The uncertainty in the model outputs for each hypothetical experiment is then evaluated (Figure 3) and summarized in Table 3.

The calculated variance indicates the lack of information on the state of knowledge; hence it is a suitable measure to compare the performance of different *(*in this case *hypothetical)* experiments Compared to the reference case, the largest increase in knowledge (largest reduction in uncertainty) on off-gas CO_2 and oxygen was obtained in scenario 1, while in scenario 2 no incremental knowledge could be obtained on these two outputs. The reason is simply that in scenario 1 the efforts were spent to better measuring the parameters with highest sensitivity on off-gas CO_2 and oxygen, while in scenario 2 less influential factors were measured. Similarly, scenario 3 was able to contribute significant knowledge on the state of off-gas CO_2 and oxygen, however in this case this was at the expense of more experimental efforts (recall that more factors were measured in this scenario).

Table 3. Experimental scenario versus uncertainty estimation in four model outputs

Experiment	Off-gas CO_2	Oxygen	Glucose	Biomass
	σ^2_{CO2}	σ^2_{O2}	$\sigma^2_{glucose}$	$\sigma^2_{biomass}$
Reference	5.6e-3	14.62	2310.8	137.75
Scenario 1	**0.38e-3**	**0.58**	2311.8	137.71
Scenario 2	5.9e-3	14.5	1415.1	47.98
Scenario 3	**0.34e-3**	**0.59**	1391.5	47.15

Important to remark, on the other hand, is that scenario 1 did not bring any additional knowledge on the state of glucose and biomass, whereas scenario 2 did. Again the reason is simply as stated above: the factors measured in scenario 2 have significant influence on biomass and glucose. Summarizing, when one measures factors that are highly influential on the outputs in question (e.g. off-gas CO_2), one obtains significant knowledge on that state. This knowledge can help the experimenter in setting up priorities for experimental research, thereby directing the resources more effectively.

4. Conclusions

This study evaluated global sensitivity analysis as a means for establishing research priorities in experimental work. The concept was successfully evaluated using a simulation study with a system consisting of a batch cultivation of *S. coelicolor* for antibiotic production. The sensitivity analysis indeed identified that only few factors (around 8 to 10 out of 56) were influential on the behavior of the system, while the rest were of little or no significance. Using this information, several *hypothetical* experiments were performed, whose results showed indeed that experiments directed at measuring the few significant factors brought in significant additional information about the state of the system. On the contrary, when one invests resources on measuring less influential factors, one obtains no or negligible amounts of information.

Since the sensitivity analysis based-methodology was evaluated using only simulations, experimental validation would be the logical next step. Note that the methodology is expected to be valuable on its own – given that a mechanistic model describing various features of the system is available – especially since it helps screening a large number of factors involved in typical experimental work as a preliminary analysis prior to performing more detailed experiment design, e.g. DoE. This preliminary analysis is traditionally based on previous experiences and prior knowledge. The benefit of this new approach is that it makes the preliminary analysis systematic and efficient.

References

G. Sin, A.E. Lantz, K.V. Gernaey, 2009. Good modelling practice (GMoP) for PAT applications: Propagation of input uncertainty and sensitivity analysis. Biotechnol. Prog. *DOI 10.1021/bp.166.*

G. Sin, P. Ödman, N. Petersen, A.E. Lantz, K.V. Gernaey, 2008, Matrix notation for efficient development of first-principles models within a PAT framework: Integrated modeling of antibiotic production with *S. coelicolor*, Biotech. Bioeng., 101, 153-171.

J.C. Helton, F.J. Davis, 2003. Latin hypercube sampling and the propagation of uncertainty in analyses of complex systems. Reliab. Engng. Syst. Saf., 81, 23-69.

A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, S. Tarantola, 2008. Global Sensitivity Analysis: The primer. West Sussex, England, John Wiley & Sons.

C.-F. Mandenius, A. Brundin, 2008. Bioprocess Optimization Using Design-of-Experiments Methodology. Biotechnol. Prog. 24, 1191-1203.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Continuous Implementation of the Ideal Time Delay in EMSO

Thiago Corrêa do Quinto[a], Argimiro R. Secchi[b] and Evaristo C. Biscaia Jr.[b]

[a]*TRANSPETRO/DGN/GAS/TO*
[b]*LMSCP/PEQ/COPPE/UFRJ*
Av. Horácio Macedo, 2030 - Centro de Tecnologia - Bloco G - Sala G-115
CEP 21941-914 - Rio de Janeiro – RJ – E-mail address: evaristo@peq.coppe.ufrj.br

Abstract

The concept of time delay arises in many practical chemical engineering applications. Ideal time delay appears every time an advective transport is the solely transport mechanism present in part of the process. Due the distributed nature of pure time delay, this phenomena cannot be directly simulated by a finite dimension lumped system. The main objective of this contribution is the automatic development of rational approximations of ideal time delay transfer functions resulting in lumped systems of specified finite dimension. Different approximations were generated and their dynamic behaviors were analyzed and compared, where stability properties are the main focus of the analysis. Oscillations and inverse response appear in various time delay approximations, therefore these undesirable behaviors were minimized through the inclusion of adjustable *m*-order filters in the approximations. Data buffering is the common strategy considered in many dynamic process simulators to include pure time delays. However, this strategy is not compatible with numerical time integration that uses variable time steps in order to meet specified accuracy, because the delayed data must be interpolated compromising the accuracy of the numerical integration. Different continuous approximations were included in EMSO – Environment for Modeling Simulation and Optimization – and their performances were compared through some practical examples. The orthogonal collocation approach presented better stability properties for high-order approximations and the undesirable oscillations were well-controlled with low-order filters.

Keywords: ideal time delay, rational approximation, dynamic simulation, EMSO.

1. Introduction

The ideal time delay is present in many dynamics systems and has unstable effect in control systems. Your effect in the time response is simply to delay the exciting signal by *v* time units. For sheet lines, webs, fibers, and conveyors, the time delay is the distance between the manipulated variable and the process variable divided by the speed. For plug flow, the entire residence time is a transportation delay, that is, the time delay is the volume divided by the flow rate. The flow in pipelines, sample lines, static mixers, coils, and heat exchanger tubes can be considered to be essentially plug flow (Blevins et al., 2003). Other example of time delayed effects is the transcription of a specific gene in response to some stimulus that may lead to the production of an enzyme that ultimately changes the flux distribution within a metabolic network. However, there is a considerable time delay of maybe 15 or 20 min between sensing the stimulus and the eventual increase in enzyme activity, which is much slower than biochemical

reactions (Mocek et al., 2005). For these reasons, the mathematical modeling of time delay is very important when the original system holds such characteristic, in order to be used in control system design.

A difficulty to the time delay modeling is its Laplace transform $e^{-\nu \cdot s}$ cannot be expressed as a finite-order polynomials rate in *s* domain. Thus, an exact transfer functions representation of a system with time delay is not possible. For instance, first and second order transfer functions with time delay are not low-order transfer functions and when inserted in control systems may produce oscillations even if proportional controllers are used. However, an accurate approximation with a polynomials ratio in s domain, as Padè approximations, can be obtained.

Wang and Hu (1999) proposed a robust stability test for dynamic systems with short delays based on a 6^{th}-order polynomial, but it does not present good results for high frequencies and large time delays. Phillip et al. (1999) showed that Padè approximation for step responses presents squared and absolute errors larger than Bessel approximation, becoming unstable for system of order higher than 4. Guzmán et al. (2008) used the Smith predictor (SP) scheme to cope the problem of controlling delayed systems by an interactive tool. Their work focused on describing the problem of typical PI control for controlling time-delay systems, how the SP can be used to solve such problems, and also how it is possible to face the SP drawbacks against uncertainties, integrating systems, and unstable systems.

In simulations evaluated in the time domain, time delay can be simulated with FIFO queues (first in, first out). In this case, the stack length is directly proportional to the delay time and inversely proportional to the integration time step. Due to large delays inside the feedback loop, queue level dynamics might exhibit oscillations (Gómez-Stern et al., 2002). Other form to simulate the dead time is by series of capacitive systems (or CST – Continuous Stirred Tank), but it requires a considerable computational effort due to the large number of CST necessary to have an acceptable accuracy.

In this work, different continuous approximations, based on Taylor series and orthogonal collocation, are proposed to model the time delay and are applied in dynamic simulations on EMSO – *Environment for Modeling Simulation and Optimization* – (Soares and Secchi, 2003). The step function with time delay, which is a hard function to approximate, is used to evaluate the robustness and performance of the methods.

2. Approximation based on Taylor series

The time delay in the Laplace domain, Equation (1), was approximated in a continuous form by the ratio of two *n*-degree polynomials, the Padè approximation.

$$\frac{H(s)}{G(s)} = e^{-\nu \cdot s} \tag{1}$$

where ν is the time delay. If *f(x)* is an analytic function of *x*, in the neighborhoods of $x = 0$, then it can be expanded in a Taylor series around this point:

$$f(x) = \sum_{i=0}^{\infty} \frac{1}{i!} \cdot \left. \frac{d^i f(x)}{dx^i} \right|_{x=0} = \sum_{i=0}^{\infty} c_i \cdot x^i \tag{2}$$

Approximating Equation (2) by the ratio of two *n*-degree polynomials in *x*:

$$f(x) \cong f(0) \cdot \frac{b_0 + \sum_{i=1}^{n} b_i \cdot x^i}{a_0 + \sum_{i=1}^{n} a_i \cdot x^i} \tag{3}$$

Coefficients a_i and b_i of Equation (3) are determined to reproduce the Taylor series expansion until the order $2n$, with $c_0 = f(0)$ and $a_0 = b_0 = 1$:

$$\left(c_0 + \sum_{i=1}^{2\cdot n} c_i \cdot x^i\right) \cdot \left(1 + \sum_{i=1}^{n} a_i \cdot x^i\right) = f(0) \cdot \left(1 + \sum_{i=1}^{n} b_i \cdot x^i\right) \tag{4}$$

Evaluating the products e equaling the equipotent terms coefficients of x, results in the following linear systems:

$$\begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_n \\ c_2 & c_3 & c_4 & \cdots & c_{n+1} \\ \vdots & \vdots & \vdots & & \vdots \\ c_n & c_{n+1} & c_{n+2} & \cdots & c_{2\cdot n-1} \end{bmatrix} \cdot \begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_1 \end{bmatrix} = -\begin{bmatrix} c_{n+1} \\ c_{n+2} \\ \vdots \\ c_{2\cdot n} \end{bmatrix} \quad and \quad \begin{bmatrix} a_1 & a_0 & 0 & 0 & \cdots & 0 & 0 \\ a_2 & a_1 & a_0 & 0 & \cdots & 0 & 0 \\ a_3 & a_2 & a_1 & a_0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ a_{n-1} & a_{n-2} & a_{n-3} & a_{n-4} & \cdots & a_0 & 0 \\ a_n & a_{n-1} & a_{n-2} & a_{n-3} & \cdots & a_1 & a_0 \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{n-1} \\ c_n \end{bmatrix} = c_0 \cdot \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_{n-1} \\ b_n \end{bmatrix}$$

where, respectively, coefficients a_i and b_i are determined.

Approximating Equation (1) as described above, Equation (5) is obtained.

$$\frac{H(s)}{G(s)} = e^{-\nu \cdot s} = \frac{1 + b_1 \cdot \nu \cdot s + b_2 \cdot \nu^2 \cdot s^2 + b_3 \cdot \nu^3 \cdot s^3 + \cdots + b_n \cdot \nu^n \cdot s^n}{1 + a_1 \cdot \nu \cdot s + a_2 \cdot \nu^2 \cdot s^2 + a_3 \cdot \nu^3 \cdot s^3 + \cdots + a_n \cdot \nu^n \cdot s^n} \tag{5}$$

Multiplying the numerator and denominator of Equation (5) by an auxiliary function, *X(s)*, and using the following inverse of Laplace transform:

$$L^{-1}(s^i \cdot X(s)) = \frac{d^i x(t)}{dt^i}$$

Equation (5) can be written in time domain as:

$$\begin{aligned} x(t) + a_1 \cdot \nu \cdot \frac{dx(t)}{dt} + a_2 \cdot \nu^2 \cdot \frac{d^2 x(t)}{dt^2} + a_3 \cdot \nu^3 \cdot \frac{d^3 x(t)}{dt^3} + \cdots + a_n \cdot \nu^n \cdot \frac{d^n x(t)}{dt^n} &= g(t) \\ x(t) + b_1 \cdot \nu \cdot \frac{dx(t)}{dt} + b_2 \cdot \nu^2 \cdot \frac{d^2 x(t)}{dt^2} + b_3 \cdot \nu^3 \cdot \frac{d^3 x(t)}{dt^3} + \cdots + b_n \cdot \nu^n \cdot \frac{d^n x(t)}{dt^n} &= h(t) \end{aligned} \tag{6}$$

Defining auxiliary state variables:

$$x_i(t) = \frac{d^{i-1} x(t)}{dt^{i-1}}, \qquad i = 1:n \tag{7}$$

Equation (6) can be described by the following first-order linear system of ordinary differential equations:

$$\begin{cases} \dfrac{dx_1(t)}{dt} = x_2(t) \\ \dfrac{dx_2(t)}{dt} = x_3(t) \\ \vdots \\ \dfrac{dx_n(t)}{dt} = g(t) - \left(x_1(t) + a_1 \cdot \nu \cdot x_2(t) + a_2 \cdot \nu^2 \cdot x_3(t) + a_3 \cdot \nu^3 \cdot x_4(t) + \cdots + a_{n-1} \cdot \nu^{n-1} \cdot x_n(t)\right) \end{cases} \tag{8}$$

Subject to the initial condition: $x_i(0) = 0$, $i = 1$ *to* n. The $h(t)$ function is determined by:

$$h(t) = x_1(t) + b_1 \cdot \nu \cdot x_2(t) + b_2 \cdot \nu^2 \cdot x_3(t) + b_3 \cdot \nu^3 \cdot x_4(t) + \cdots + b_{n-1} \cdot \nu^{n-1} \cdot x_n(t) + b_n \cdot \nu^n \cdot \frac{dx_n(t)}{dt} \tag{9}$$

Using the last differential equation of Equation (8) and knowing that $a_n = b_n$:

$$h(t) = g(t) + (b_1 - a_1) \cdot \nu \cdot x_2(t) + (b_2 - a_2) \cdot \nu^2 \cdot x_3(t) + (b_3 - a_3) \cdot \nu^3 \cdot x_4(t) + \cdots + (b_{n-1} - a_{n-1}) \cdot \nu^{n-1} \cdot x_n(t) \tag{11}$$

Therefore, for the simulation of the model, the following generalized system is used:

$$\frac{d\underline{x}(t)}{dt} = \underline{\underline{A}} \cdot \underline{x}(t) + \underline{b} \cdot u(t) \tag{12}$$

Figure 1 compares model predictions for two distinct polynomials with exact unitary step function with ν = 1 time unit. The model is stable for $n < 13$ presenting considerable oscillation before the step and acceptable result after the step, but for high-

degree polynomials the matrix *A* contains high truncate errors besides its high density that demands great computational effort.

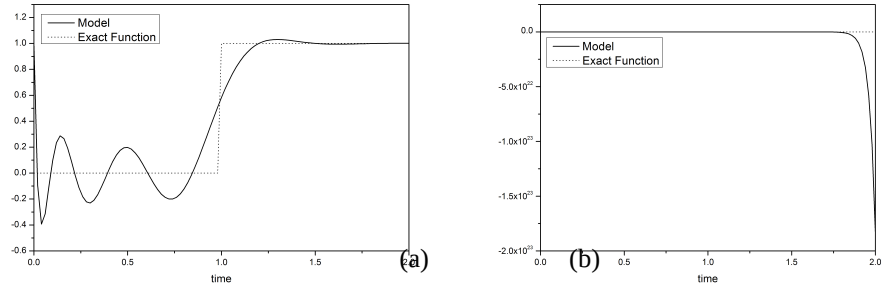


Figure 1: Approximation function of the unitary step with: (a) 6th degree polynomial with 100 discretization points and squared error of 1.8555; (b) 13th degree polynomial and 100 discretization points (unstable).

3. Approximation based on orthogonal collocation

An alternative to solve the problem of high truncate errors in the approach described in the previous section is to apply the orthogonal collocation method (Michelsen and Villadsen, 1972). Choosing the Jacobi polynomials $P_n^{\alpha,\beta}(x)$ of degree *n* and appropriate values of the parameters α and β, the matrix *A* and vector *b*, Equation (12), are obtained from the matrix of first derivatives of the polynomial at the collocation points (including $x = 0$ and $x = 1$). For $n = 5$, $\alpha = 1$ and $\beta = 0$, it becomes:

Matrix A

$$\begin{bmatrix} -12.6 & -4.08 & 1.22 & -0.566 & 0.307 & -0.109 \\ 9.80 & -2.53 & -3.13 & 1.16 & -0.583 & 0.202 \\ -5.18 & 5.54 & -1.14 & -2.97 & 1.18 & -0.385 \\ 4.11 & -3.49 & 5.07 & -0.719 & -3.46 & 0.926 \\ -4.39 & 3.46 & -3.95 & 6.81 & -0.555 & -4.02 \\ 9.94 & -7.68 & 8.23 & -11.6 & 25.7 & -18.5 \end{bmatrix}$$

Roots of $P_5^{1,0}(x)$:

$$\begin{bmatrix} 0.0398 & 0.198 & 0.438 & 0.695 & 0.901 \end{bmatrix}$$

Vector b

$$\begin{bmatrix} 15.8 & -4.92 & 2.96 & -2.43 & 2.64 & -6.00 \end{bmatrix}$$

In Figure 2 it can be seen that the limitation in the polynomial approximation for Taylor series does not occur in this case, thus higher degree polynomials can be used. However, the approximation of the step function present high frequency oscillation besides the even higher density of matrix *A*.

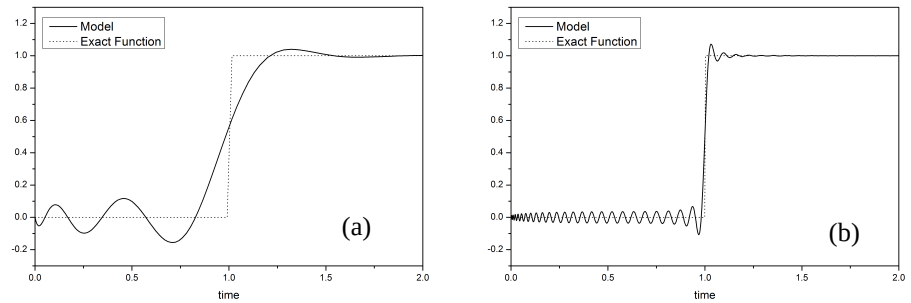


Figure 2: Orthogonal collocation approximation for unitary step function with v = 1 time unit and: (a) 5th degree Jacobi polynomial with squared error of 1.2359; (b) 50th degree Jacobi polynomial with squared error of 1.0344.

4. Approximation based on series of orthogonal collocation

In order to reduce the undesirable oscillations, the system was discretized in *N* intervals with application of Jacobi polynomial of degree 1 in each interval, creating a tridiagonal matrix *A* to solve the ordinary differential equations system:

Matrix A

$$\begin{bmatrix} -4.5 & -1.5 & 0 & 0 & 0 & 0 \\ 13.5 & -7.5 & 0 & 0 & 0 & 0 \\ 0 & 6 & -4.5 & -1.5 & 0 & 0 \\ 0 & -6 & 13.5 & -7.5 & 0 & 0 \\ 0 & 0 & 0 & 6 & -4.5 & -1.5 \\ 0 & 0 & 0 & -6 & 13.5 & -7.5 \end{bmatrix}$$

Root of $P_1^{1,0}(x)$:

[1/3]

Vector b

$$\begin{bmatrix} 6 & -6 & 0 & 0 & 0 & 0 \end{bmatrix}$$

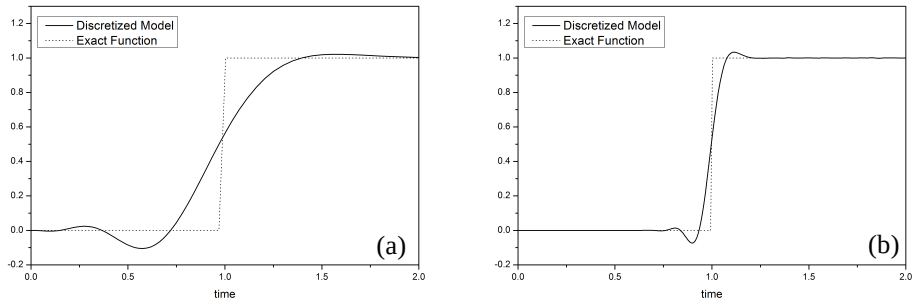


Figure 3: Series of first-degree Jacobi polynomials approximation for unitary step function with ν = 1 time unit and: (a) 3 intervals, with global order of 6 and squared error of 1.1495; (b) 24 intervals, with global order of 48 and squared error of 1.0344.

In Figure 3, a considerable reduction of the oscillation can be observed. In Figure 3a, with 3 intervals, smoother oscillations are obtained when compared with Figure 2a (global order of 11). This result is even more expressive in Figure 3b (global order of 48) when compared with Figure 2b (global order of 51).

5. Approximation based on series of orthogonal collocation with capacitive filter

In this section a filter with capacitive characteristics is applied in the system to reduce even more the oscillation in the previous section. A first-order filter is added into the system by a γ factor with range from 0 to 1, where $\gamma = 1$ prevail the orthogonal collocation and $\gamma = 0$ the capacitive filter.

In Figure 4, different time delayed simulations are plotted and compared with ideal time delay. The results show a damping in the oscillations of the previous cases, but an increase of the squared errors. Therefore, for system without inherent oscillation (Figures 4a, 4b and 4c), filtering is a resource for system with abrupt changes, reducing the oscillation generated by the approximation, otherwise the use of filter is not desired, because it attenuates natural oscillation (Figure 4d) and reduces the accuracy.

6. Conclusion

The time delay approximation based on discretized orthogonal collocation is more robust and present better performance than the usual Padè approximations. The use of capacitive filters is beneficial for systems without inherent oscillations and with abrupt

changes. The computational costs of the proposed approximation are reduced when using series of first-degree Jacobi polynomials.

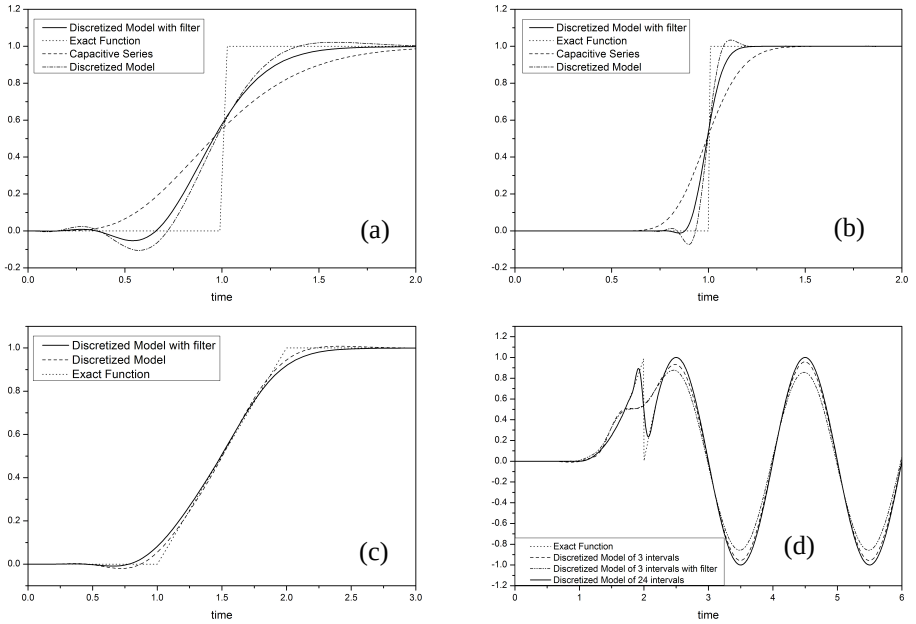


Figure 4: Time delayed simulations: (a) 6^{th} global order approximation of step function with capacitive series and discretized orthogonal collocation without and with filter (squared error of 1.2592, and $\gamma = 0.8$); (b) 48^{th} global order (with filter: squared error of 1.2669, and $\gamma = 0.94$); (c) discretized orthogonal collocation in ramp function $(t-1-(t-2)\cdot\Phi(t-2))\cdot\Phi(t-1)$ with global order of 6 and squared errors of 0.1354 (without filter) and 0.2302 (with filter) ; and (d) function $(t-1)2\cdot(1-\Phi(t-2))+(\sin(\pi\cdot(t-2)))\cdot\Phi(t-2)$ with squared errors of 1.3116 (3 intervals) and 1.3006 (24 intervals), both without filter, where $\Phi(t)$ is step function.

References

Blevins, T.L., G.K. McMillan, W.K. Wojsznis, M.W. Brown, 2003 Advanced Control Unleashed - Plant Performance Management for Optimum Benefit, ISA..

Gómez-Stern. F., J.M. Fornés, F.R. Rubio; 2002, Dead-time compensation for ABR traffic control over ATM networks, Control Engineering Practice, 10, p. 481–491.

Guzmán, J.L., P. García, T. Hängglund, S. Dormido, P. Albertos, M. Berenguel; 2008, Interactive tool for analysis of time-delay systems with dead-time compensators, Control Engineering Practice, 16, p. 824–835.

Michelsen, M.L. and J. Villadsen; 1972, A convenient computational procedure for collocation constants, The Chemical Engineering Journal, 4 (1), p. 64-68.

Mocek, W.T., R. Rudnicki, E.O. Voit; 2005, Approximation of delays in biochemical systems, Mathematical Biosciences, 198, p. 190-216.

Philipp, L.D., A. Mahmood and B.L. Phlipp; 1999, An Improved Refinable Rational Approximation to the Ideal Time Delay, Transactions on circuits and systems I-Fundamental Theory and applications, 46 (5), p. 637-640.

Soares, R.P., A.R. Secchi; 2003, EMSO: A new environment for modelling, simulation and optimisation, Computer Aided Chemical Engineering, 14, p. 947-952.

Wang, Z., H. Hu; 1999, Robust Stability test for dynamic systems with short delays by using padé approximation, Nonlinear Dynamics, 18, p. 275-287.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A New Algorithm for Global Optimization: Molecular-Inspired Parallel Tempering

Silvia Ochoa[a,b], Jens-Uwe Repke[a] and Günter Wozny[a]

[a] *Chair of Process Dynamics and Operation, Berlin University of Technology, Sekr. KWT 9, Strasse 17. Juni 135, Berlin 10623, Germany*
[b]*Research Group on Modelling and Control (GIMOC), University of Antioquia, Medellín, Colombia.*

Abstract

A new stochastic algorithm for global optimization is proposed in this work, which is inspired in the Parallel Tempering algorithm, and is denoted as Molecular Inspired Parallel Tempering (MIPT) because it mimics the behavior of charged molecules in solution, taking into account the different forces that are exerted on each molecule. Molecules are classified into explorers and refiners. Explorers have lower friction values and are subject to repulsion forces, covering a wider region of the search space for finding the global optimum. Refiners have higher friction values, intending to make a finer search in the neighborhood of already found local minima. MIPT efficiency is tested through four challenging case studies, showing not only its capability to find the global optimum in all cases with a 100% of success ratio, but also with a comparable computational effort than other established well known stochastic methods.

Keywords: Global Optimization, Molecular-Inspired Parallel Tempering, Stochastic Optimization.

1. Introduction

The scope of Global Optimization (GO), as stated by Neumaier (2004), is to find the absolutely best set of admissible conditions to achieve an objective under given constraints. In general, deterministic or probabilistic approaches can be used for solving the GO problem. In this work, a probabilistic method is proposed, which is based on the Parallel Tempering (PT) algorithm (Earl and Deem, 2005). PT algorithms simulate various non-interacting replicas of the original system at different temperatures. Replicas at higher temperatures are able to explore wider regions, whereas those at lower temperatures are able to perform a finer sampling in local regions. A special characteristic of PT algorithms is that they exchange temperatures between different replicas, providing a global character by allowing escaping from local minima and exploring a wider search space. PT methods have been successfully applied in different fields (simulation of polymers, proteins, and general optimization problems) for more than 20 years; however, there are still open topics regarding the algorithm, and it has been suggested that temperature may not always be the best parameter to temper, and that PT can be conducted by other parameters such as pair potentials or chemical potentials. In Section 2, a Molecular Inspired Parallel Tempering (MIPT) algorithm for solving GO problems, is presented, which uses a friction factor as parameter for tempering. MIPT is easy to implement, does not need derivative information, and includes two important characteristics, global character and local refinement, which according to Törn and Žilinskas (1989), are important features that should be present in

GO methods. In Section 3, four different examples solved using MIPT are presented, and a comparison between MIPT and other well-established methods is provided.

2. Molecular-Inspired Parallel Tempering Algorithm (MIPT)

MIPT mimics the behavior of charged molecules in solution. Figure 1 shows a charged two-molecule system, in which each molecule is affected by three different forces: repulsion ($\boldsymbol{F}_{\text{rep}}$), random ($\boldsymbol{F}_{\text{Bm}}$), and friction ($\boldsymbol{F}_{\text{f}}$). Repulsion force ($\boldsymbol{F}_{\text{rep}}(i)$) exerted by molecule j over molecule i (both with the same electrostatic charge) is calculated by Coulomb's law, as inversely proportional to the square of their distance (d_{ij}). The random force ($\boldsymbol{F}_{\text{Bm}}(i)$), is responsible for the brownian motion of the molecule and it is expressed by means of a normalized Gaussian distribution with zero mean and standard deviation one ($\xi(0,1)$). Finally, the friction force ($\boldsymbol{F}_{\text{f}}(\text{i})$) has an opposite direction to the net force ($\boldsymbol{F}_{\text{net}}=\boldsymbol{F}_{\text{rep}}+\boldsymbol{F}_{\text{Bm}}$) and it is proportional by a γ-factor, to the velocity υ of the molecule. The γ-factor is a friction coefficient inversely proportional to the molecules temperature, and it is the parameter used in this approach for tempering the algorithm.

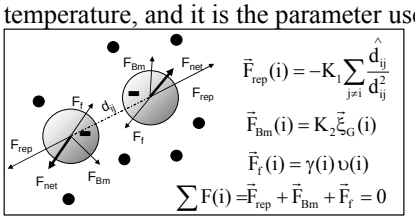


Figure 1 Forces acting on a system of electrically charged molecules. Black and gray particles represent molecules of solvent and charged solute molecules, respectively.

$$\Delta\vec{x}(i) = \frac{\vec{F}_{rep}(i) + K_2\vec{\xi}_G(i)}{\gamma(i)} \qquad (1)$$

The net force ($\boldsymbol{F}_{\text{net}}$) acting on each molecule is responsible for its displacement in the space of decision variables; such displacement is given by (1). The flow diagram of MIPT algorithm is presented in Figure 2. A key feature of MIPT is that molecules are classified into Refiners and Explorers. Refiners are always feasible points constrained to higher friction values, forcing the search to a narrow region around their current position, providing the method with a local character. These are molecules that make shorter displacements in the search region, in order to refine the search in a neighborhood that contains local optima. Explorers have lower friction values, are allowed to be unfeasible and are affected by the repulsive effect, which force them to move towards unexplored zones, providing the method with a global character. Explorers make larger moves escaping from local optima that are already being explored by refiners. The optimization starts by setting up the input data to the algorithm: the number of molecules (n_m) and decision variables (n_{dv}), the stopping criteria and the tuning parameters of the algorithm. MIPT tuning parameters include the γ-values, K_1 and K_2, the β parameter in the Metropolis condition and the minimum number of explorers (*minNexp*). It is suggested that $n_{\text{m}} \geq 2 \cdot n_{\text{dv}}$, in order to cover a wider search region. The probability of finding the global optimum increases with the number of molecules, but there is a compromise between success and computational effort. The next step is the random generation of the initial *nm* molecules (starting positions *x0(i)*). The molecules are ranked according to their objective function values and their feasibility. Then, $\gamma(i)$ is calculated and the algorithm opens two branches, one for the refiner molecules and the other for the explorers. The refiners branch starts setting the

repulsion force to zero (F_{rep}=0). Then, the random force F_{bm}(i) is calculated from a Gaussian distribution. Equation (1) is used for calculating the new position (x_{new}) of the molecule. The objective function value (f_{new}) for x_{new} is evaluated, and the Gibbs free energy change $\Delta G(i)$ with respect to the current objective function value (f_{opt}) is calculated according to (2). After calculating $\Delta G(i)$, the feasibility condition is checked again. If the new position is an unfeasible point, $x_{new}(i)$ is rejected. On the contrary, if, x_{new} is feasible and also $\Delta G(i)<0$, the new position is accepted and position for molecule i is updated to the new value ($x_{opt}(i)=x_{new}(i)$).

$$\Delta G(i) = \frac{f_{new}(i) - f_{opt}(i)}{\max(f_{opt}) - \min(f_{opt})} \qquad (2)$$

The Explorers branch begins calculating $F_{rep}(i)$ as a function of the distance between molecules. Calculation of $\Delta G(i)$ is carried out using (2). Then, the Metropolis acceptance criterion is evaluated in order to reject or accept the new position $x_{new}(i)$. If the Metropolis criterion is satisfied, the new position will be accepted and updated. Furthermore, if the new position is not only feasible but also has an optimal value lower than the objective function value for the worst refiner (that with the higher f_{opt}), it would be set as refiner, in addition, the worst refiner would be set as explorer, if and only if, the minimum number of explorers (*minNexp*) has been already reached. Finally, the stopping criteria are checked, and the algorithm stops if any of them has been met.

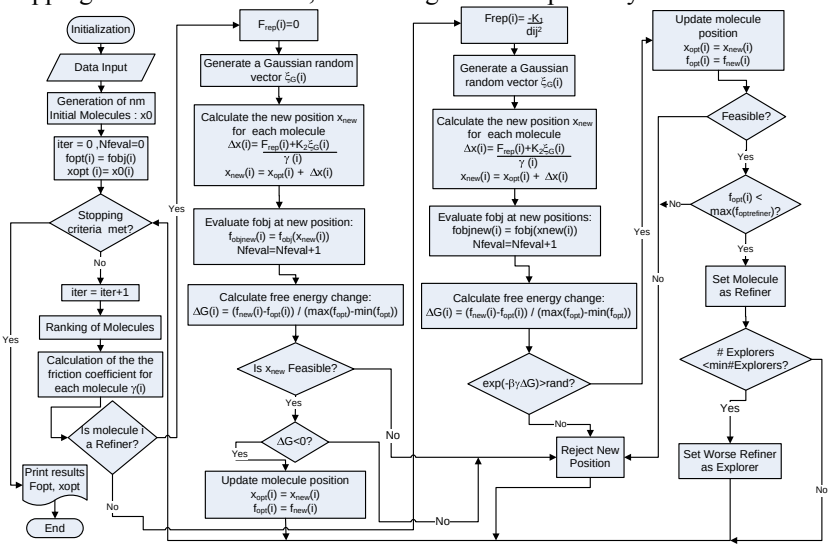


Figure 2 Molecular-Inspired Parallel Tempering (MIPT) Algorithm for Global Optimization.

3. Examples

In this section, the capability and efficiency of the MIPT algorithm for solving different GO problems is tested using four case studies. MIPT results were obtained using a PC with 1.66GHz Intel Core 2 Duo processor and 1 GB RAM.

3.1. Extended Dixon-Szegö Set

The first example is a set of 14 standard continuous GO test problems taken from the Dixon-Szegö collection. This set of functions was chosen because it is diverse enough to cover many kinds of difficulties that arise in global optimization (Hedar and

Fukushima, 2006). A description of the 14 test problems including the already known global minima values can be found elsewhere (Hedar and Fukushima, 2003). Figure 3 shows a comparison between the results obtained using MIPT, the new version of the GLOBAL method (Csendes et al., 2008) and the Continuous Grasp (C-GRASP) method (Hirsch et al., 2007). In this work, each test problem was run 100 times using MIPT (the stopping criterion was the same as the reported in Hirsch et al. (2007)), achieving 100% of success ratio. Second column in Figure 3 shows the number of decision variables for each problem. The first column of each method shows the average number of function evaluations (*Nfeval*) while the second reports the *CPU time*. It is important to notice that the comparison should be made in terms of *Nfeval*, because the three methods were run in different computers and large differences might be observed. MIPT has shown a better performance in solving the Branin, Easom, Rosenbrock-2, Shubert, Hartman-3 and Rosenbrock-5 problems, whereas the GLOBAL achieved better performance for the remaining problems except for the Zakharov-5 and the Goldstein-Price. It is remarkable that the MIPT method lies within the range of average values of the other methods, showing to be competitive in solving problems with several local minima.

Test Function	ndv	MIPT		GLOBAL		C-GRASP	
		Nfeval	*CPUTime*	*Nfeval*	*CPUTime*	*Nfeval*	*CPUTime*
Branin	2	**379**	0.0579	1023	0.058	59857	0.0016
Easom	2	**388**	0.0611	1604	0.0916	89630	0.0042
Goldstein-Price	2	607	0.0803	923	0.0516	**29**	0
Rosenbrock-2	2	**1775**	0.1807	6274	0.4117	1158350	0.0132
Shubert	2	**395**	0.0613	1399	0.093	82363	0.0078
Hartman-3	3	**890**	0.1289	3608	0.3208	20743	0.0026
Shekel-10	4	10566	1.4695	**1815**	0.1614	4701358	3.5172
Shekel-5	4	10689	2.0058	**1489**	0.1313	5545982	2.3316
Shekel-7	4	14663	2.4590	**1684**	0.1461	4052800	2.3768
Rosenbrock-5	5	**348670**	47.1290	374685	24.7559	6205503	1.752
Zakharov-5	5	6658	0.9751	8227	0.5369	**959**	0
Hartman-6	6	22676	5.4186	**16933**	1.788	79685	0.014
Rosenbrock-10	10	3980100	345.1947	**1908469**	131	20282529	11.4388
Zakharov-10	10	65252	12.4934	**47288**	3.1428	3607653	1.0346

Figure 3 Comparative Results for case study 1: MIPT vs. GLOBAL and C-GRASP.

$$\max V_{PK}(X_1, X_2, X_3, X_4, X_5, Y_1)$$

s.to: $\boldsymbol{f}(\boldsymbol{x}, p) = 0$

$$V_{in} - V_{HK} = 0 \;;\; V_{HK} - V_{PFK} - V_{Pol} = 0$$

$$V_{PFK} - V_{GAPD} - 0.5V_{Gol} = 0;\; 2V_{GAPD} - V_{PK} = 0$$

$$2V_{GAPD} + V_{PK} - V_{HK} - V_{Pol} - V_{PFK} - V_{ATPase} = 0$$

$$V_{PK} \le 2V_{in};\; Y_{k0} \le Y_k \le Y_{k0};\quad k = 1,2,3,4,5,8$$

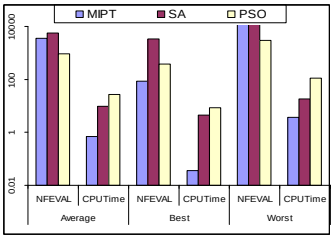


Figure 4 Case Study 2: Optimization Problem (left), Results for MIPT, SA and PSO (right).

3.2. Nonlinear Constrained Optimization Problem

The second case study is an example of the steady state optimization problem of biochemical reaction networks, which addresses the maximization of the flux of Piruvate Kinase (V_{PK}), enzyme directly responsible for ethanol production in the *Saccharomyces cerevisiae* pathway (Xu et al., 2008). This case study (Figure 4-left) consists on a nonlinear constrained optimization problem, where f represents the nonlinear steady-state model of the process, X_i the metabolites concentrations, Y_k the enzymes activities, and X_{i0} and Y_{i0} are the basal steady-state values (corresponding to an objective function value V_{Pk0}=30.1124). Xu solved the problem applying a standard iterative IOM and a modified iterative IOM approaches, finding an objective function value V_{PK}=64.828V_{PK0} and V_{PK}=64.829V_{PK0}, respectively. The same problem was solved in this work using MIPT, Simulated Annealing (SA) and Particle Swarm Optimization

(PSO). These three methods reached the same objective function value, V_{PK}= 65.022 V_{PK0}, better than that obtained by the IOM approaches. A comparison on the performance for MIPT, SA and PSO in terms of number of function evaluations (*Nfeval*) and *CPU* time for the average, best and worst cases from 100 different runs randomly initialized, is presented in Figure 4-right. Values for the IOM approaches are not reported in the original work. As can be seen, MIPT has the best performance in terms of CPU time, whereas the number of function evaluations required for reaching the global optimum lies between those needed by SA and PSO.

3.3. Mixed Integer Nonlinear Problem MINLP

The optimization problem addressed here (Figure 5-left) consists on a MINLP involving 7 decision variables. A recent work by Yiqing et al., (2007) reports results for this case study, comparing Genetic Algorithms (GA), a Simulated Annealing based algorithm (M-SIMPSA), the original PSO algorithm, and an improved PSO algorithm denoted as R-PSO. Furthermore, R-PSO algorithm was used in two variants: R-PSO_unc which updates continuous and discrete variables simultaneously, and the other, denoted as R-PSO_c that makes the mixed variables updating at difference pace. Figure 5-right shows the comparison between the mentioned stochastic algorithms and MIPT, in terms of the average number of function evaluations (*Nfeval*) and the success ratio (*NRC*), which is the percentage of runs that converged to the global optimum (f_{opt}=4.579582), in 100 executions. As shown in Figure 5, MIPT algorithm not only has reached the global optimum for all runs (NRC=100%), but also has required the less number of function evaluations, showing its capacity for dealing with MINLP problems.

$$\begin{aligned}
&\min f(x_1,x_2,x_3,y_1,y_2,y_3,y_4)\\
&\quad=(y_1-1)^2+(y_2-2)^2+(y_3-1)^2-\ln(y_4+1)\\
&\quad+(x_1-1)^2+(x_2-2)^2+(x_3-3)^2\\
&\text{s.to: } y_1+y_2+y_3+x_1+x_2+x_3\le 5\\
&\quad y_3^2+x_1^2+x_2^2+x_3^2\le 5.5\\
&\quad y_1+x_1\le 1.2;\ y_2+x_2\le 1.8\\
&\quad y_3+x_3\le 2.5;\ y_4+x_1\le 1.2\\
&\quad y_2^2+x_2^2\le 1.64;\ y_3^2+x_3^2\le 4.25\\
&\quad y_2^2+x_3^2\le 4.64;\ x\le 0;\ y\in\{0,1\}^4
\end{aligned}$$

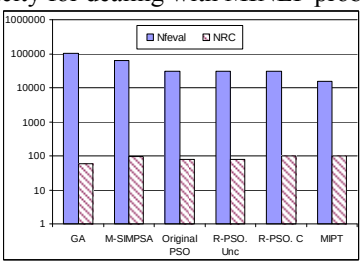


Figure 5 Case Study 3: Optimization Problem (left) and Stochastic Methods Comparison (right).

3.4. Optimization Problem in a Large Scale System

This case study consists on a two-distillation-column system for purification of a beer stream in an ethanol production process. The model consists on a total of 2411 equations, involving 241 state variables. Five control loops for controlling the pressure in both columns, and the level in the reboilers and in the partial condenser of the rectification column were included. The initial conditions for the model are the steady state values. In case of disturbance, the heat duty of both reboilers (Q_{R1} and Q_{R2}) and the heat duty of the partial condenser (Q_{C2}) can be used to optimize the economic operation of the process. The cost function (F_{opt}) to be minimized is given by equation (3), where Δt_{opt} is the optimization horizon, F_{EB1}, F_{EB2} are the ethanol flow rate in the bottoms of each column, F_{WD2} is the water flow rate in the distillate of the second column. The factors that multiply each term of (3) are penalization terms related to the ethanol price (for the three first weights), the cost of steam to the reboilers and the cost of cooling water in the partial condenser. *f* represents the nonlinear dynamic model of the process, u_{min} and u_{max} are the upper and lower bounds for the decision variables in (3) and x_{ED2} is the mass fraction of ethanol in the distillate of the second column (desired product). In order to compare the performance of different optimization algorithms in optimizing the

economic operation of this large scale system, a step disturbance corresponding to a 25% increase of the feed to the system was done. An optimization horizon of two hours was considered. The maximum number of function evaluations used as stopping criteria was 200. The results obtained using MIPT, SA, PSO and a Gradient based method (GRAD) are summarized in Figure 6-left, where the average, best and worst values from 5 randomly initialized runs for each method are compared. It can be seen that MIPT algorithm has reached the same average and best optimal values than SA and PSO, but requiring less CPU Time. In contrast, the GRAD method had the worse performance for the average and worst cases. Additionally, Figure 6-right compares the behaviour of the ethanol losses (taking as the sum of F_{EB1} and F_{EB2}), x_{ED2}, F_{opt} (instantaneous values shown), for the process with optimization (denoted as after optimization, only MIPT results are shown), and without (i.e. Q_{R1}, Q_{R2} and Q_{C2} were kept at their steady state values after the disturbance). MIPT has shown to be an efficient algorithm that can be also applied in optimization of large scale systems.

$$\min_{Q_{R_1},Q_{R_2},Q_C} \left(\int_{t_{opt}}^{t_{opt}+\Delta t_{opt}} 0.63\left(F_{EB_1}+F_{EB_2}\right)+0.63\,F_{WD_2}+6.33\left(Q_{R_1}+Q_{R_2}\right)+0.31Q_C \right) \tag{3}$$

$$\text{s.to. } f(\dot{x},x,u,d,t)=0;\ x_i(t_0)=x_{0i};\ u_{min}\le u\le u_{max};\ x_{ED_2}\ge 0.9$$

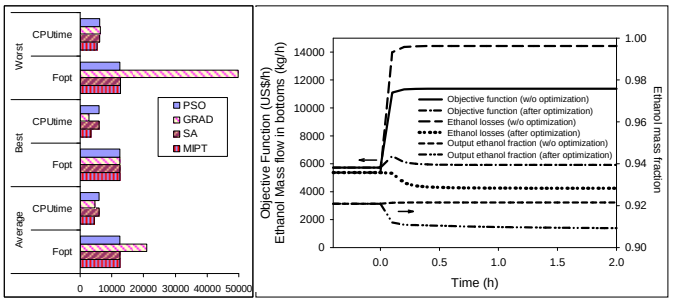


Figure 6 Case Study 4: Comparative Results for MIPT, SA, GRAD and PSO (left) and Comparison with vs. without optimization (right).

4. Conclusions

A new stochastic algorithm for Global Optimization, denoted as Molecular-Inspired Parallel Tempering (MIPT) has been proposed. Through four challenging case studies, MIPT performance was compared to well-established optimization algorithms, reaching a better performance in most of the cases. It was shown that MIPT was able to reach the global optima for a wide range of problems, including constrained NLP, MINLP and highly multimodal functions; with a high success rate and a reasonable number of function evaluations, showing to be suitable for solving global optimization problems.

References

T. Csendes, L.Pál, J.O.H. Sendín, J. R. Banga, 2008, Optim Lett 2:445–454.
D. Earl and M. Deem, 2005, Phys Chem Chem Phys.7:3910-3916.
A-R.Hedar and M.Fukushima, 2003, Optim. Methods Softw, 18, 265–282.
A-R.Hedar and M.Fukushima, 2006, European Journal of Operational Research, 170,329–349.
M. J.Hirsch,C.N.Meneses, P.M.Pardalos, M.G.C.Resende, 2007,Optimization Letters, 1:201–212.
A. Neumaier, 2004, Acta Numerica,13, 271-369.
A. Törn and A. Žilinskas, 1989, Global Optimization, Springer verlag.
G.Xu, C. Shao, Z. Xiu, 2008, Computers and Chemical Engineering, 32, 1546–1568.
L. Yiqing, Y. Xigang, L. Yongjian, 2007, Computers and Chemical Engineering, 31,153–162.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Molecular-scale modeling of the degradation of phenol in Advanced Oxidation Processes reaction media

Bruno Ramos, Antonio Carlos S. C. Teixeira

LSCP/CESQ - Department of Chemical Engineering, Escola Politécnica da Universidade de São Paulo. Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil.

Abstract

Modeling, in its diverse length and time scales, has been pointed as one of the most remarkable features responsible for the advances in Chemical Engineering science. The smallest physically relevant scale modeling deals with frozen individual molecules; an approach known as *quantum modeling* (QM): length scales of the order of Angstrons (10^{-10} m) and no time effects taken into account. Modeling in this scale is useful to obtain a large amount of data on properties which can be further applied into other modeling approaches or serve as parameters to process design and optimization. In this work, Density Functional and Semiempirical methods were used to study the degradation of phenol promoted by hydroxyl radicals, as those generated in Advanced Oxidation Processes media. The results are discussed and compared with experimental data collected by the group. A method is presented for modeling ChE interesting reactions and obtaining kinetic constant values.

Keywords: molecular modeling, phenol, AOP, kinetic

1. Introduction

One of the major tasks of a chemical engineer is to make the bridge between molecules and a desired product with specific function and structure. In fact, the change in the focus – from the process to the product – is a clear illustration of this fact. The future of Chemical Engineering, among all the driving forces that push industry through the 21st century (i.e. market demands, efficient consumption of resources, minimization of environmental impact, better use of research and human resources etc.), lies within a multidisciplinary and multiscale approach – or 3PE approach (triplet molecular processus – product – process), as labeled by J. C. Charpentier (Charpentier, 2007). According to that author, one of the major goals of Chemical Engineering, in this globalization and sustainability scenery, is the implementation of multiscale simulation and computational modeling applications to real world situations, from molecular to production scale. In order to accomplish this, some tools, divided in three fields, are fundamental: molecular modeling, scientific instrumentation and non-invasive measurement techniques, and powerful computational tools. It is only through simultaneous advances in these fields that it will be possible to understand and describe the relationships between events in nano- and micro-scale.

2. Methods

2.1. Energy evaluation

The energies were evaluated by using two different approaches: the Density Functional Theory (DFT) and the semiempirical model AM1. All calculations were carried out in Gaussian 03 (Frisch *et al.*, 2003).

Density Functional Theory (DFT)

The Density Functional Theory embraces a series of electronic structure methods that use the electron density as basic variable, and their origin trace back to the Thomas-Fermi model, in 1927 (Cramer, 2004). Basically, DFT aspires to predict *exactly* properties of many-electron systems without recourse to the wave function, using only (explicitly or implicitly) in the ground-state electron density. In this work, the functional chosen was the hybrid combination known by the acronym B3LYP (Scuseria & Staroverov, 2005):

$$E_{XC}^{B3LYP} = (1-a)E_X^{LSDA} + aE_{XC}^{\lambda=0} + bE_X^B + cE_C^{LYP} + (1-c)E_C^{LSDA}$$

This functional is a combination of the LSDA (local spin density approximation) exchange functional, the non-interactive exchange-correlation functional, the Becke exchange functional, the Lee-Yang-Parr correlation functional and the LSDA correlation functional; scaled by three empirically determined constants *a*, *b* and *c*. The explicit form of most of these functionals involves integrals whose arguments are electron densities or electron density gradients, and therefore an iterative procedure is necessary to evaluate the electron density and, consequently, the energy.

Semiempirical (AM1)

The semiempirical methods start out from the *ab initio* formalism and then introduce rather drastic assumptions to speed up the calculations, typically by neglecting many of the less important terms in the *ab initio* equations. In order to compensate for the errors caused by these approximations, empirical parameters are incorporated into the formalism and calibrated against reliable experimental or theoretical reference data. The most popular semiempirical methods are based on the MNDO model (modified neglect of diatomic overlap). The Austin Model 1 (AM1) was developed as an improvement of the basic MNDO architecture, in order to account for a correct evaluation of hydrogen bonding energies (Thiel, 2005) and, since the modeled system is rich in hydrogen bonds, this was the chosen semiempirical method for energy evaluation.

2.2. Geometry Optimization

The geometries modeled were optimized using the Eigenvector Following (EF) algorithm. The EF procedure is essentially based upon the widely known Newton-Raphson method (Cramer, 2004).

2.3 Frequency Calculation

In order to evaluate the thermodynamical properties of the chemical systems, the numerical values of the vibrational frequencies are necessary. They are also useful as an indication of the nature of a transitional state. The molecular frequencies depend upon the second derivative of the energy with respect to the nuclear positions (i.e. the Hessian matrix). A description on the followed procedure can be found at Ochsterski (1999).

2.4. Aqueous solution modeling

Polarized Continuum Model (PCM)
The Polarized Continuum Model is one of the Self-Consistent Reaction Field (SCRF) methods, which are – in essence – adaptations of the Poisson method for *ab initio* calculations. Basically, it simulates the presence of the system into a continuum of constant dielectric (Cramer, 2004). The chosen model, known by the acronym SCI-PCM (self-consistent isodensity PCM) simulates the molecule inside an isosurface of electron density (s.a. the surfaces constructed using Van der Waals atomic radii), and uses a numerical evaluation of the electrostatic potential. This potential enters the Hamiltonian for the system and, thus, returns the solvation energy.

Explicit solvation models
Instead of working with potentials, it is common to construct a structure with some of the solvation molecules around the system under scoop. These structures are known as supramolecules, and often offer a better accounting of the effects of hydrogen bonds.

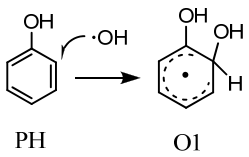
2.5. Thermodynamical and kinetic properties evaluation

In Statistical Mechanics there is a function with an equivalent status of the wave function in Quantum Mechanics. That is, a function that characterizes the system. It is called *partition function*, and it is used as an input to calculate the thermodynamical properties, i.e., enthalpy, Gibbs free energy, chemical potential etc.

The kinetic constant is evaluated according to the Transition State Theory (Ochterski, 2001).

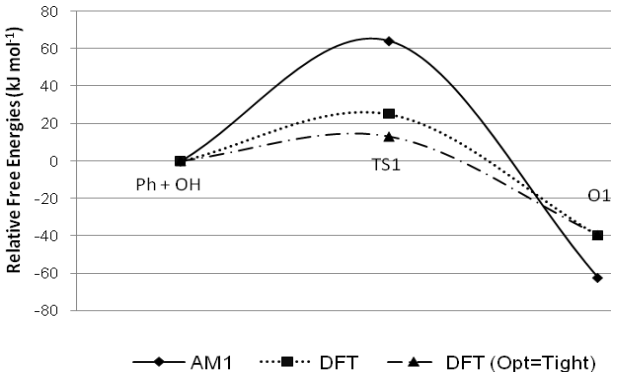
3. Results and Discussion

3.1. Reaction
The reaction was modeled as follows, using the software *Gaussian 03* for calculating and the software *GaussView 3* to generate the inputs and visualize the outputs.



PH O1

3.2. Free energy profile as a function of the theory employed
A thermodynamical evaluation at the optimized geometries of reactants (Ph + OH), product (O1) and transition state (TS1) retrieved the profile shown below. As can be seen by the relative energies, the semiempirical method has shown a larger activation free energy (defined as the difference between TS and reactant free energies) and predicted a larger exoergic behavior than the DFT functional. The transition state was also optimized by using a tighter convergence criteria. The activation free energy was found to be smaller than the one found with the standard set of convergence thresholds. By comparison with the experimental work conducted by Bonin *et al.* (2007), which found an activation free energy of ca. 16 kJ mol^{-1}, the DFT optimization with tighter criteria was found to describe better the system modeled.



3.3. Aqueous environment modeling

The aqueous environment was simulated by three different techniques: *i)* the PCM implicit model; *ii)* the supramolecular explicit model; and *iii)* by Molecular Dynamics and Monte Carlo techniques. The results suggest that the implicit solvation model is inadequate to simulate the behavior of the system in aqueous solution, possibly by not accounting correctly for the hydrogen interactions. Table 1 resumes the results.

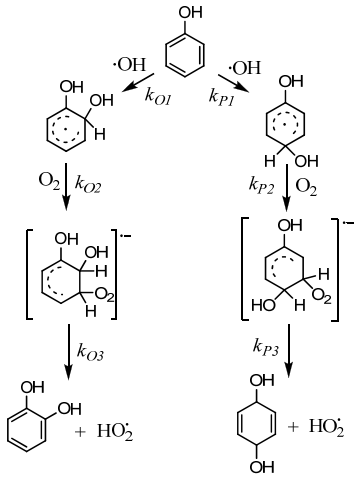
Table 1 – Solvation energies according to different solvation models.

	Total energy variation due to solvation (Hartree/particle)						
	ΔU_{solv} (PCM)	ΔU_{solv} (3W)	ΔU_{solv} (4W)	ΔU_{solv} (3W2)	ΔU_{solv} (4W2)	ΔU_{solv} (MD)	ΔU_{solv} (MC)
Phenol	-0,0301	-11,4007	-16,1264	-	-	-96,8851	-1,310
O1	0,0915	-12,0882	-	-24,1740	-32,2284	-	-

The results show a clear distinction between PCM and explicit models. Particularly for the radical O1, the PCM model predicts a lower stability in aqueous phase than in gas phase. This is an intuitively and physically erroneous conclusion. The solvation molecules stabilize the charge imbalance of the solute molecules and, therefore, improve the stabilization of the system.

3.4. Kinetic constants and comparison to experiment

The kinetic constants were evaluated for the whole mechanism of reactions leading to the formation of catechol and hydroquininone from phenol, as shown below:



For this sequence of reactions, the set of kinetic constants is shown in Table 2, and Figure 1 illustrates a simulation of these mechanisms plotted against experimental data, for a batch reaction of a media initially containing 12 mmol L^{-1} phenol and constant [·OH] and [O_2] concentrations, as taken from experiment.

Table 2 – Kinetic constants calculated and experimental values (when available)

	Kinetic constant (M^{-1} s^{-1} ou s^{-1})	
	Calculated	Experimental
k_{O1}	$3{,}33 \times 10^{10}$	$9{,}0 \times 10^{8}$ Edalatmanesh, 2008 $2{,}0 \times 10^{10}$ Friedrich, 2007
k_{O2}	$1{,}79 \times 10^{36}$	-
k_{O3}	$1{,}97 \times 10^{10}$	-
k_{P1}	$7{,}47 \times 10^{08}$	$2{,}0 \times 10^{8}$ Edalatmanesh, 2008 $6{,}0 \times 10^{9}$ Friedrich, 2007
k_{P2}	$8{,}56 \times 10^{33}$	-
k_{P3}	$4{,}17 \times 10^{10}$	-

All the kinetic constants were evaluated using DFT thermodynamical results, as was seen that DFT shows a better correlation to reality than AM1. The analysis of Figure 1 leads to the observation that phenol and cathecol general behaviour are in good agreement with experimental data: both theoretical and experimental profiles predict the complete decay of the pollutant at 500 seconds, and a maximum peak of cathecol of 4 mmol L^{-1}. The kinetic constant values found were within the margins of experimental data collected, as can be seen from Table 2.

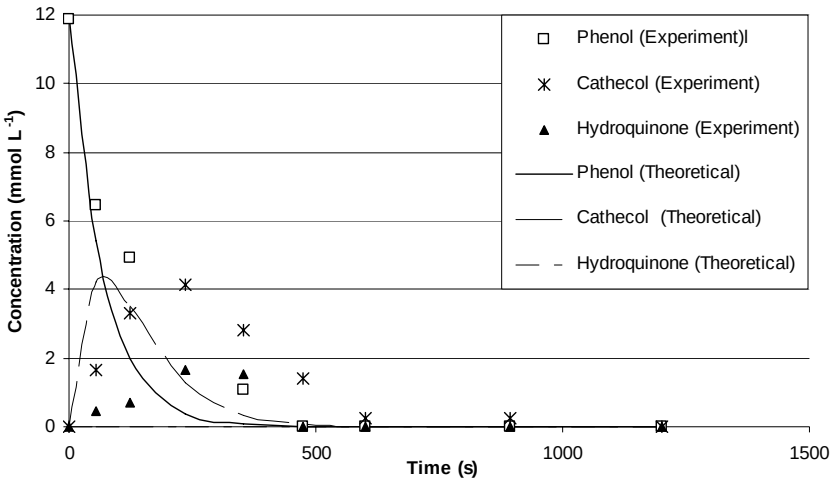


Figure 1 – Concentration profile simulation applying calculated kinetic constants in comparison to experimental data.

3. 4. Conclusion

From the results found in this work, it can be suggested that:

i) DFT plays a better role than AM1, in describing the mechanics of AOP reaction media;

ii) Implicit solvation models, such as SCI-PCM are inadequate to describe the solvation processes in the studied systems: an explicit solvation sphere is fundamental for the correct accounting of the highly stabilizing hydrogen interactions;

iii) DFT kinetic constants are in good agreement with experimental values obtained and can be used to simulate concentration profiles and provide qualitative insights on the behavior of the reactions.

References

J. C. Charpentier, 2007, The triplet "molecular processes-product-process" engineering: the future of chemical engineering? Chem. Eng. Sci., 57, p. 4667-4690.

M. J. Frisch; J. A. Pople (eds.), 2003, Gaussian 03, Revision B-03. Gaussian, Inc., Wallingford CT.

J. C. Cramer, 2004, Essentials of computational chemistry. 2ed. John Wiley & Sons Ltd: New York.

G. E. Scuseria; V. N. Staroverov, 2005, Progress in the development of exchange-correlation functionals. Theory and applications of computational chemistry. p. 669-724.

W. Thiel, 2005, Semiempirical quantum-chemical methods in computational chemistry. Theory and applications of computational chemistry. p. 559-580.

J. W. Ochterski, 1999, Vibrational analysis in Gaussian. Gaussian, Inc.

J. W. Ochterski, 2001, Thermochemistry in Gaussian, Gaussian, Inc.

J. Bonin *et al.*, 2007, Reaction of hydroxyl radical with phenol in water up to supercritical conditions. J. Phys. Chem. A, 111, p. 1869-1878.

M. Edalatmanesh *et al.*, 2008, Kinetic modeling of aqueous phenol degradation by UV/H_2O_2 process. Intl. J. Chem. Kinet., 40, p. 34-43.

L. C. Friedrich *et al.*, 2007, Modelagem cinética da oxidação do fenol pela reação de Fenton. XXXª Reunião Anual da Sociedade Brasileira de Química.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A NOVEL APPROACH FOR THE PREDICTION OF PSD IN ANTISOLVENT MEDIATED CRYSTALLIZATION

Massimiliano Grosso [a], Omar Galan [b], Roberto Baratti [a], Josè A. Romagnoli [a*]

[a] *Dipartimento di Ingegneria Chimica e Materiali, Università degli Studi di Cagliari, Piazza d'Armi, I-09123, Cagliari, Italy*
[b] *Department of Chemical Engineering, Lousiana State University, USA*
** On leave from Department of Chemical Engineering, Lousiana State University, USA*

Abstract

A phenomenological model for the description of antisolvent mediated crystal growth processes is here presented. The crystal size growth dynamics is supposed to be driven by a deterministic growth factor coupled to a stochastic component. The evolution in time of the particle size distribution is then described in terms of a Fokker-Planck equation. Validations against experimental data are presented for the NaCl-water-ethanol anti-solvent crystallization system.

Keywords: Antisolvent crystallization, Particle size distribution estimation, Fokker-Planck equation.

1. Introduction

Antisolvent aided crystallization is an advantageous technique of separation where the solute is highly soluble or heat sensitive. The driving force in crystal formation is the super-saturation that establishes the thermodynamic equilibrium for the solid-liquid separation. The development of rigorous mathematical models describing the dynamic of crystal growth in crystallization processes is based-on population balances. At the core of the structured population dynamics, the number of crystals in a fed-batch crystallizer is increased by nucleation and decreased by dissolution or breakage. Structured population balances models provide detailed information regarding the crystal size distribution in the crystallization unit. However, they demand a great deal of knowledge on the complex thermodynamic associated with the solute and solvent properties to be adequately incorporated in the population balances. Some important contributions in this subject have been reported in the literature (Worlitschek and Mazzotti, 2004; Nowee et al., 2008).

An alternative approach to deal with particulate systems characterized by MCS and CSD is the Fokker-Planck Equation (FPE). The FPE is just an equation of motion for the distribution function of fluctuating macroscopic variables. The study of stochastic system as the Brownian motion resulted in the Fokker-Planck Equation (Risken, 1984).

In this work we propose a phenomenological model for the description of the crystal growth process, which is based on a novel stochastic approach for the prediction of MCS and CSD in a bench-scale fed-batch crystallization unit where anti-solvent is added to speed-up the crystal formation process. The crystal growth is modeled as a stochastic process ruled by a by a classic logistic equation of common use in theoretical ecology (Grosso et al., 2007) coupled with a geometric Brownian motion. The resulting

model is a continuity equation (Fokker-Planck Equation, FPE) for the probability density of the crystal size. A comparison with experimental data will be presented for validation of the FPE approach for the NaCl-water-ethanol anti-solvent crystallization system

2. Mathematical Model

The crystals are classified by their size, L and the growth of each individual crystal is assumed to be independent by the other crystals and governed by the same deterministic model. In the proposed model, fluctuations and unknown dynamics not captured by the deterministic term are modeled as a random component that has an additive action on the growth dynamics (Gelb, 1988). This random component can be thought as a Geometric Brownian Motion, GBM, (Ross, 2003) where the intensity of the fluctuations depends linearly by the crystal size. The stochastic model can thus be written as a Langevin equation of the following type:

$$\frac{dL}{dt} = \nu(L,t;\theta)L + L\eta(t) \qquad 1$$

In Equation 1, $\nu(L,t;\boldsymbol{\theta})$ is the expected rate of growth (the deterministic model introduced below), L is the size of the single crystal, t is the time, $\boldsymbol{\theta}$ is the vector parameter defined in the model, and $\eta(t)$ is a random term assumed as Gaussian additive white noise i.e., $E[\eta(t)]=0$ and $E[\eta(t)\,\eta(t+\tau)]=2D\delta(\tau)$, where D is the additive noise intensity. Equation 1 implies that the crystal size L behaves as a random variable, characterized by a certain probability density function (PDF) $\psi(L,t)$ depending on the state variables of the system, i.e. the size L and time t. From an applicative point of view, the PDF coincides with the Particle Size Distribution (PSD) that can be experimentally observed.

It should also be noted that, when the GBM assumption holds, the PDF is a lognormal distribution, at least in the limit case of constant ν value (Ross, 2003). If ν depends also on L, some distortions from the ideal lognormal case can be however expected. This feature is qualitatively observed for many (although not all) crystalline substances (Eberl et al, 1990) and in the present case.

The Langevin Equation 1 can be further manipulated as follow:

$$\frac{1}{L}\frac{dL}{dt} = \nu(L,t;\theta) + \eta(t) = \frac{dy}{dt} = \mu(y,t;\theta) + \eta(t) \qquad 2$$

The new random variable, y, can be thus described in terms of its probability density distribution, $\psi(y,t)$, at any instant of time t and should follow the Fokker-Planck Equation, FPE (Risken, 1996):

$$\frac{\partial\psi(y,t)}{\partial t} = -\frac{\partial}{\partial y}\left(\nu(y,t;\theta)\psi(y,t)\right) + D\frac{\partial^2\psi(y,t)}{\partial y^2} \qquad 3$$

In equation 3, the "pseudo-diffusion" coefficient D takes into account the fluctuations intensity in the particle growth process (Matsoukas and Lin, 2006). As it regards the deterministic part of the model, a Logistic growth process is assumed (c.f., Tsoularis and Wallace, 2002).

$$v(y,t;\theta)=ry[1-\left(\frac{y}{K}\right)] \quad 4$$

This choice is mainly motivated to the requirement for a simple model with a parsimonious number of adjustable parameters, i.e. the growth rate, *r*, and the asymptotic equilibrium value *K*. The present growth model can be regarded as the simplest model taking into account mild nonlinearities. In spite of its simplicity, this model provides the main qualitative features of a typical growth process: the growth follows a linear law at low crystal size values and saturates at a higher equilibrium value. Using the selected growth model, the FPE can be written as:

$$\frac{\partial \psi}{\partial t}=-\frac{\partial}{\partial y}\left(ry\left(1-\frac{y}{K}\right)\psi\right)+D\frac{\partial^2\psi}{\partial y^2} \quad 5.$$

along with the boundary conditions (Risken, 1996):

$$-v(y)\psi(y,t)+D\frac{\partial\psi(y,t)}{\partial y}=0 \qquad \text{at } y=0,\forall t$$
$$\frac{\partial\psi(y,t)}{\partial y}=0 \qquad \text{as} y\to\infty \quad \forall t \quad 6$$

Finally the evolution in time of the probability density is described in terms of a linear, partial differential equation depending on the parameters *r* (linear growth rate), *K* (crystal size asymptotic value) and *D* (diffusivity). Such parameters are considered constant for each experimental condition but they can change with the anti-solvent flow rate. Indeed, antisolvent crystallization kinetics is affected by the antisolvent feed rate and, the parameters appearing in the model should be related to this operating condition. The numerical integration of Equation 5 was accomplished by exploiting the routine DASPG, from IMSL, in order to solve the resulting implicit system of algebraic ordinary differential equations obtained by means of collocation on finite elements.

3. Experiments

The experimental data used in the present work are those reported in Nowee et al. (2008). They refer to an aqueous solution of Sodium Chloride (NaCl) where Ethanol was added at different constant feed rates. Four experiments were selected under three constant antisolvent feeding profiles – one experiment at a feed rate of 0.83 $ml \cdot min^{-1}$, (low feed rates, hereafter we refer to this experiment as run A), two at a medium rate of 1.64 $ml \cdot min^{-1}$ (run B) and one at a higher rate of 3.2 $ml \cdot min^{-1}$ (run C). The first acquisition time, t_0, for each run will be the initial condition used for the FP model calibration. The PSD was measured at different time instants using Mastersizer 2000 particle size analyzer.

Model calibration for the estimation of parameters in the model is carried out separately for every run. The parameters to be estimated are: $\boldsymbol{\theta} = [\log(D), r, K]$ ($\log(D)$ is used instead of *D* in order to reduce the statistic correlation between the parameters). It should be noted that direct measurements of the Particle Size Distribution are available at *N* different spatial locations and at *M* different time values for every operating condition, i.e. anti-solvent flow rate. Parameter inference is accomplished by using the least square criterion, thus searching the minimum of the objective function:

$$\Phi(\boldsymbol{\theta}) = \Phi(r, K, D) = \sum_{j=1}^{M}\sum_{k=1}^{N}\left(\psi_{\mathrm{mod}}\left(y_k, t_j; \boldsymbol{\theta}\right) - \psi_{\exp}\left(y_k, t_j\right)\right)^2 \qquad 7$$

In Equation 7, $\psi_{mod}(y_k,t_j)$ is the probability density function evaluated through numerical integration of Equation 5, at time t_j and size coordinate y_k, while the distribution $\psi_{exp}(y_k,t_j)$ is the experimental observation of the PSD for the size coordinate y_k at time t_i. The model calibration is thus carried out by comparing N point observations of the distribution (N ~ 40), monitored at M different times (M between 6 and 10, depending on the experimental run). The minimum search is carried out by the Levenberg-Marquardt method. As it regards the supplementary run at intermediate feeding rate, this was used for an a-posteriori model validation.

4. Results

Figure 1 reports the point estimations of the adjustable parameters of the model *vs* the corresponding feed rate. The error bar for the 95% confidence interval is also shown. Point and confidence estimations are carried out with traditional inference methods (e.g. Bates and Watt, 1987). It should be noted that the uncertainty in the estimation is rather small and the parameters significantly change with V_0, especially at the highest feed rate

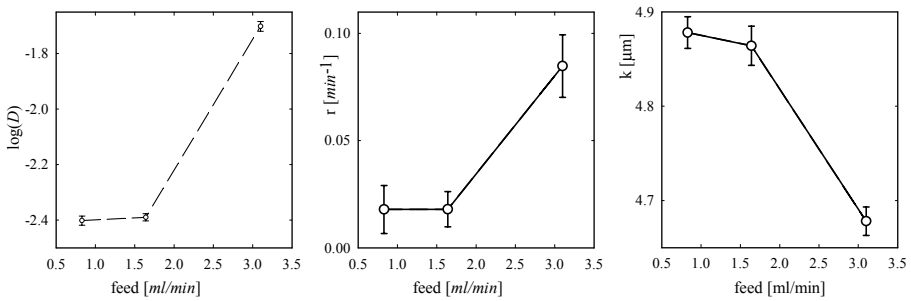


Figure 1: Parameter estimation together with the 95% confidence interval *vs* V_0.

Figure 2 reports the comparison of the PSD experimentally observed and the model prediction for the experiment at low feed rates, at the first and final acquisition time. Similar results are found at the higher feed rates, It is evident that there is a quantitative agreement between experiments and phenomenological model. In particular, the model is able to correctly capture the log-normal shape of the experimental PSD.

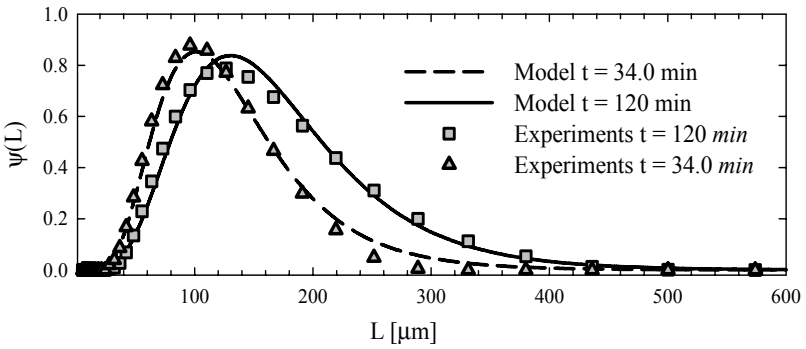


Figure 2: Comparison between model and the experimental probability density functions for the feeding rate V_0 = 0.82 ml/min: triangles are the experimental observations at t= 34 min, dashed line is the corresponding model prediction; square points are the experimental observations at t= 240 min, solid line is the corresponding model prediction.

Comparison between model and experiments is also carried out by reporting the time evolution of the experimental observations and the corresponding model prediction for first moments of the distribution, i.e., the mean, μ and the variance σ^2:

$$\mu(t) = \int_{\Omega} L\,\psi(L,t)dL \qquad o^2(t) = \int_{\Omega} [L-\mu(t)]^2\,\psi(L,t)dL \qquad 8$$

Figures 3 and 4 show, respectively, the mean and variance experimentally observed (square points) compared with the theoretical predictions (solid line) for the three runs as a function of time. The agreement is rather good at each time and the FPE model, driven by its deterministic part (the logistic growth term), correctly describes the increasing trend of the average crystal growth. In Figure 3b and 4b we also report experimental data provided by the second experimental run (cross points), (not used for the parameter inference), demonstrating that the model has also good predictive capabilities

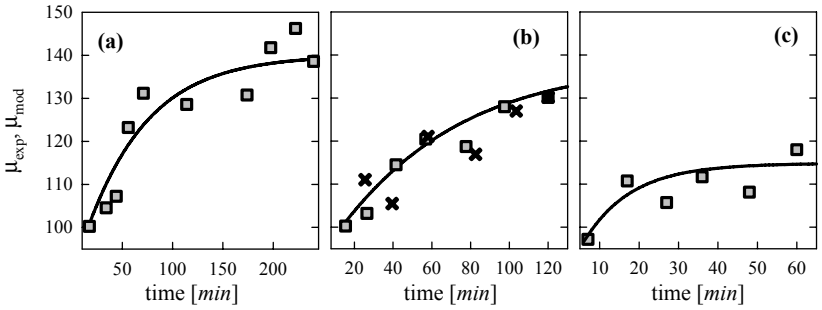


Figure 3: Mean of the Crystal Size Distributions for the three feeding rates Square points are the experimental observation, solid lines are the model predictions

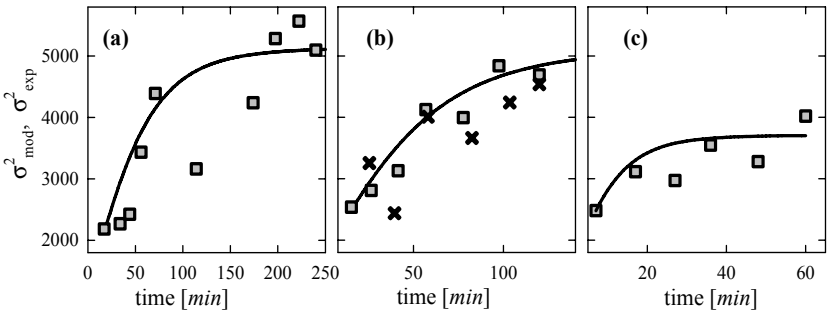


Figure 4: Variance of the Crystal Size Distributions for the three feeding rates Square points are the experimental observation, solid lines are the model predictions

5. Conclusions

A stochastic phenomenological model for the description of antisolvent crystal growth processes is here proposed. The crystal size is considered as a random variable, whose probability density evolution in time is described in terms of a Fokker-Planck equation. The model is tested on data provided in a bench-scale fed-batch crystallization unit where anti-solvent is added to speed-up the crystal formation process. The FPE formulation appears as a powerful predictive tool, as confirmed by the excellent agreement with the experiments..

Acknowledgments

Jose Romagnoli kindly acknowledges the Regione Sardegna for the support through the program "Visiting Professor 2008".
Massimiliano Grosso acknowledges PRIN-2006091953 project for its financial support

References

Eberl D.D., Srodon J., Kralik M, Taylor B.E., 1990, Ostwald ripening of clays and metamorphic minerals, Science 248, 474-477

Bates and Watts, 1988, Nonlinear regression analysis, Wiley and Sons, New York

Gelb A., 1988, Applied Optimal Estimator, M.I.T. Press, Cambridge (Massachusetts)

Grosso M., Cella R., Baratti R., 2007, Development of a probabilistic model for the plant growth in the nursery phase, 3rd International Symposium, Food And Agricultural Products: Processing And Innovations, September 24-26, 275

Nowee S.M., Abbas A., Romagnoli J.A., 2008, Antisolvent crystallization: Model identification, experimental validation and dynamic simulation, Chem. Eng. Sci. doi:10.1016/j.ces.2008.08.003

Matsoukas T. and Lin Y., 2006, Fokker-Planck equation for particle growth by monomer attachment, Phys. Rev. E, 031122

Risken H., 1996, The Fokker-Planck equation: Methods of solutions and Applications, Springer-Verlag, Berlin

Ross S.M., 2003, An Elementary Introduction to Mathematical Finance: Options and other Topics, Cambridge University Press, Cambridge

Tsoularis A. and Wallace J., 2002, Analysis of logistic growth models, Mathematical Biosciences, 179, 21-55

Worlitschek, J., and M. Mazzotti, 2004, Model-Based Optimization of Particle Size Distribution in Batch-Cooling Crystallization of Paracetamol, Crystal Growth & Design, 891-903.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Heterogeneous anaerobic biofilm reactor models Application to UASB, EGSB and AFB reactors

Mauren Fuentes, Pío A. Aguirre and Nicolás J. Scenna

INGAR Instituto de Desarrollo y Diseño (CONICET-UTN)-Avellaneda 3657 (3000) Santa Fe, Argentina

Abstract

In a previous work, a methodology developed for modeling anaerobic fluidized bed (AFB) reactors was presented. The aim of this work is to extend this methodology for modeling upflow anaerobic sludge blanket (UASB) and expanded granular sludge bed (EGSB) reactors, and compare and discuss model hypotheses and simulation results. A set of experimental data obtained by Kato et al. (2003), during the start-up of a bioreactor operating as UASB and EGSB reactor configurations, is used for model validation. A good agreement was obtained among experimental and predicted values. A simulation-based sensitivity analysis of model parameters such as the specific rate of granule rupture and the axial dispersion coefficient is performed. A decrease in the granule diameter is predicted for values of the specific rate of granule rupture higher than 1×10^{-7} m s^2 kg^{-1}. At low values of the axial dispersion coefficient, a decrease in the bioreactor efficiency is predicted. Simulation results are more sensitive when the bioreactor operates with a UASB configuration.

Keywords: dynamic modeling and simulation, anaerobic biofilm reactors: UASB, EGSB and AFB, three-phase (solid-liquid-gas) systems.

1. Introduction

Anaerobic reactors based on biofilm development such as UASB, EGSB and AFB reactors have become popular high-rate wastewater treatment systems. The anaerobic digestion is a multi-step process consisting of hydrolysis of complex organic substrates such as proteins, lipids, and carbohydrates into soluble amino acids, fatty acids, and sugars followed by the fermentation to acetate, formate, hydrogen, and carbon dioxide, which are finally utilized by methanogenic microorganisms to form methane. There are clear differences between UASB and AFB reactors, e.g. the former is based on the ability of microorganisms to form dense aggregates by autoimmobilization, and operates at a lower fluid superficial velocity than the second one, which is loaded with inert support particles for biofilm attachment. The conception of the EGSB reactor can be considered as an improvement of the UASB reactor. Inside the EGSB reactor, the higher upflow velocities, which are caused by a high recycle rate and a high height/diameter ratio, cause the sludge expansion through the whole reactor thus improving its contact with the effluent and reduce the unit dead volume.

A better understanding of the microbial and hydraulic mechanisms that regulate the system has contributed to the development of more compact and efficient units. The heterogeneous sludge distribution in UASB reactors does not make easy the application of most mathematical models developed for completely mixed anaerobic digestion systems, since these models assume homogeneous biomass distribution and hydrodynamic pattern within the reactor. The UASB reactor has being modeled as a

sum of compartments that are assumed to have ideal (mixing or plug) flow patterns, and the non-idealities of flow are described through the linkage among these compartments by bypassing and back mixing flows (Pontes and Pinto, 2006), and as a compartment using a one-dimensional dispersed plug flow model (Kalyuzhnyi et al., 2006). The adoption of some ideas from solid dynamics in a three-phase fluidized bed reactor can be used to describe the multiple solid-liquid-gas interactions that take place in UASB and EGSB reactors. In a previous work, a heterogeneous model for three-phase fluidized bed systems was developed (Fuentes et al., 2008a, 2009) and validated (Fuentes et al., 2007, 2008b). The aim of this work is to extend this methodology for modeling UASB and EGSB reactors. Some hypotheses related to the bioparticle model and hydrodynamics are discussed. Due to space restrictions, only one example is presented. It is based on the experimental results obtained by Kato et al. (2003) operating a bioreactor with UASB and EGSB configurations. A simulation-based sensitivity analysis of model parameters such as the specific rate of granule rupture and the axial dispersion coefficient is performed. A complete explanation (hypotheses and equations) of the model and other application examples could be presented in an extended version of the paper.

2. Mathematical Model

The methodology here used is based on three major modeling tasks: (1) the anaerobic digestion model, (2) the bioparticle model, and finally (3) the hydrodynamics model.
UASB and EGSB reactors are modeled as three-phase gas-solid-liquid systems. The solid phase consists of bioparticles (granules) composed by active and non-active biomass. The liquid phase is composed by the chemical species in solution (substrates, products, enzymes, ions, and water) and (active and non-active) suspended biomass, which are assumed to be as solutes. The gas phase is formed by the gaseous products from degradation stages. The global model structure is represented in Fig. 1.

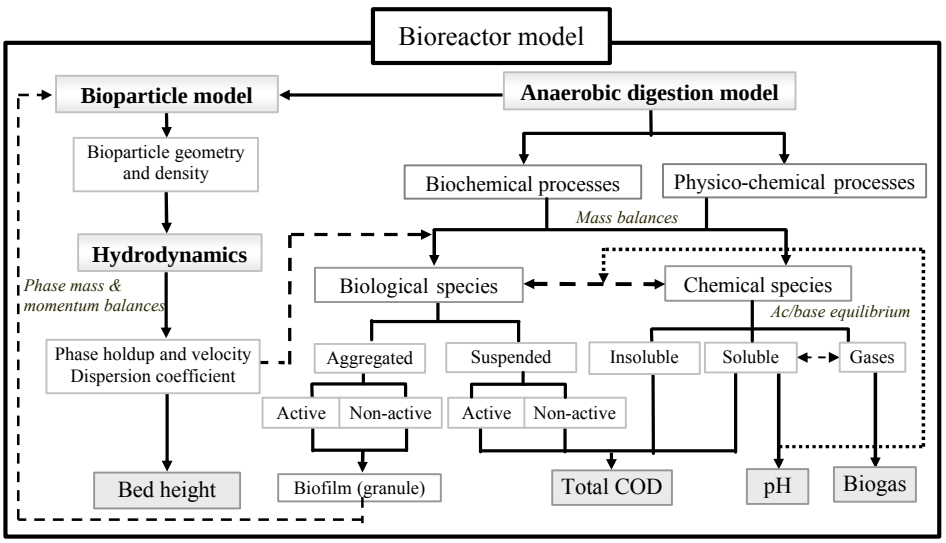


Fig. 1. Bioreactor (UASB and EGSB) model structure (modified from Fuentes et al., 2009).

A heterogeneous one dimensional (axial) dynamic model is proposed to compute the differential mass and momentum balances for the three phases and phase components

(Fuentes et al., 2008a, 2009). The bioreactor model resulted in an integral–partial derivative and algebraic equation (IPDAE) system.

2.1. Anaerobic digestion model

The anaerobic digestion model involves the biochemical (growth-uptake, death, hydrolysis and disaggregating) and physico-chemical (system charge balance for calculating pH, gas-liquid mass transfer) processes which take place in the bioreactor. The anaerobic degradation scheme proposed by Angelidaki et. al (1999) is here used. Microorganism growth kinetic expressions and kinetic and physico-chemical parameters have been previously described (Fuentes et al., 2007). The specific growth and death rates are assumed to be the same for suspended and aggregated biomass. The non-active biomass is considered as particulate material subjected to hydrolysis.

2.2. Bioparticle model

Generally, the granule composition has been described from the ash content, moisture and (active and non-active) biomass fraction. These parameters can vary during the granule formation, and thus, the dry density of granules varies. As first modeling approach, homogeneous biomass distribution on granules, constant wet biomass density and spherical geometry are assumed.

For modeling the disintegration of aggregates, some physical aspects related to the structure and compactness of granules need to be considered. Hydrodynamics affects biomass processes inside a bioreactor. These effects are more pronounced in fluidized beds than in UASB reactors, attending to the operational fluid velocity ranges (Huang and Wu, 1996; Nicolella et al., 2000). Tiwari et al. (2006) studied the influence of some factors such as temperature, alkalinity, nature and strength of substrate, and cation concentration on granule formation in UASB reactors. They concluded that: (1) a careful temperature control and an adequate alkalinity is required for generation and maintenance of granules; (2) the nature and strength of substrate in conjunction with intra-granular diffusion to a large extent determines the microstructure of the granules; and (3) the divalent cations such as calcium and iron may enhance granulation by ionic bridging and linking exo-cellular polymers. However, their presence in excess may lead to cementation due to precipitation leading to increased ash content and mass transfer limitation.

Here, the specific rate of granule rupture k_r has been defined for calculating the biomass transport phenomena. This parameter explains how the process of granulation is affected by all these environmental and operational conditions, and differs of the specific rate of biomass hydrolysis, which is related to the enzyme attack on non-active biomass. Since several factors are involved in the stability and formation of granule, and no other modeling approach has been published, the rate of granule (disaggregating) rupture is modeled as a first-order function on energy dissipation parameter (which is an upflow velocity function) and mass concentration of each microbial species present in the biomass aggregates. Parameter k_r has been assumed the same for all biological species. A sensitivity analysis of model results related to this parameter is presented in Section 3.2.

2.3. Hydrodynamics model

Characteristics of phase mixture and flow patterns are expressed by the hydrodynamic model. Generally, the flow pattern in the UASB has been described as a plug flow reactor (D_z=0) because of the low turbulence provided by the upflow velocity and gas

production. The EGSB behaves as a continuous-stirred tank reactor (D_z? 8) due to the effects of the high recirculation rate. Liquid film mass transfer resistance seems to be negligible in the EGSB but important in the UASB reactor. This could be a consequence of the different hydrodynamic conditions in the two reactors. Here, the dispersion coefficient D_z is assumed as a function of granule characteristics (terminal settling velocity and diameter) and the liquid and gas phase velocities. A sensitivity analysis of model results related to this parameter is presented in Section 3.2.

The mathematical model was implemented and solved using the process modeling software tool gPROMS (Process Systems Enterprise Ltd). The backward finite difference method (BFDM) was used to solve the partial differential equations (PDEs). Using second-order BFDM over a uniform grid of 20 intervals resulted in 3718 equations and 1140 differential variables. The total CPU time required to solve the case study described in the following section is about 37 s on an 800MHz Pentium IV PC.

3. Example

The experimental results obtained by Kato et al. (2003) are here used to calculate the application example. This work was carried out in a pilot EGSB reactor (2 dm inner diameter and 50 dm reactor height) fed with the effluent from a full-scale UASB treating municipal wastewater. The bioreactor was monitored in three different periods by varying the upflow velocity according to Table 1. During the first step-type disturbance, bioreactor operated with a UASB configuration (recirculation flow equal to zero). Average experimental values of influent chemical oxygen demand (COD) and volatile suspended solids (VSS) are shown in Table 2.

Table 1. Reactor operational conditions.

Step	Influent flow (L/d)	Recirculation flow (L/d)	Upflow velocity (dm/d)
1 (0 to 113 day)	945	0	300
2 (114 to 206 day)	945	945	600
3 (207 to 331 day)	945	1890	900

Table 2. Experimental data of the influent stream and values used in the model.

Step	Temperature (ºC)		pH		Total COD (mg/L)		Soluble COD (mg/L)		VSS (mg/L)	
	Exp.	Mod.	Exp.	Mod.	Exp.	Mod.*	Exp.	Mod.*	Exp.	Mod.*
1	29.7±1.7	30	6.9±0.1	6.9	126±53	108	56±22	56	42±27	40
2	30.3±2	30	6.7±0.2	6.9	180±50	156	79±31	78	54±29	62
3	32±1.9	32	6.9±0.2	6.9	156±48	136	55±19	56	63±30	62

* Values of 30, 30, 0, and 40 %VSS were assumed to model carbohydrate, protein, lipid and suspended biomass composition in the influent stream, respectively. Soluble COD was measured as acetic acid concentration.

3.1. Simulation results

Table 2 summarizes the values of the influent stream used for simulations. As observed in Table 3 and Fig. 2, a good agreement between experimental and predicted values of total and soluble COD and VSS concentration of the effluent stream is obtained.

Predicted values of granule diameter and solid holdup are depicted in Fig. 3 (a) and (b), respectively. Although the solid phase seems to be homogeneously distributed along the reactor height (axial dimensionless length Z^*), bioparticles with a higher diameter, and thus, with a higher terminal settling velocity, are settled at the reactor bottom.

Figure 3 (c) and (d) represents the pH and gas holdup profiles, respectively. As expected, a higher pH variation along Z^* is observed during the first step-type

disturbance, when the reactor operates with UASB configuration. A plug flow behavior is observed for the gas phase (Fig. 3d).

Table 3. Experimental and predicted values of the effluent stream.

Step	pH	Total COD (mg/L)		Soluble COD (mg/L)		VSS (mg/L)	
	Pred.*	Exp.	Pred.*	Exp.	Pred.*	Exp.	Pred.*
1	7.05	68±29	60	42±21	30	13±8	20
2	6.90	87±21	68	55±22	39	8±3	20
3	7.03	79±26	71	44±15	36	24±17	25

* The values were measured at the end of each step-type disturbance.

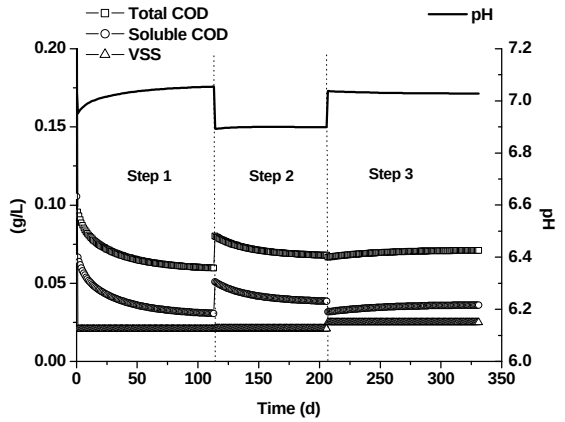


Fig. 2. Predicted values of total and soluble COD, VSS concentration and PH.

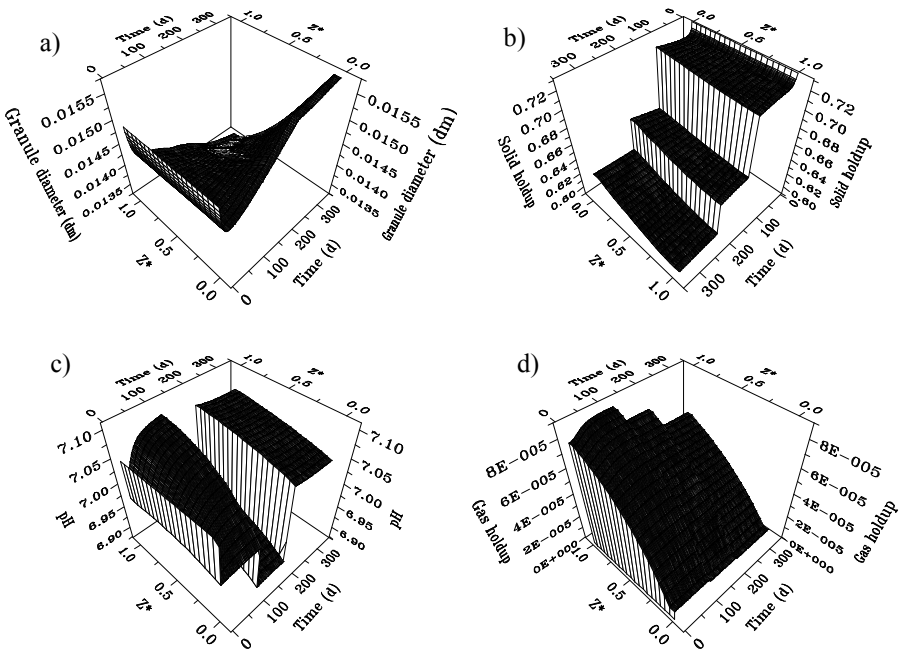


Fig. 3. Predicted values: (a) granule diameter, (b) solid holdup, (c) ph, and (d) gas holdup.

3.2. Sensitivity analysis

Figure 4 shows the sensitivity analysis of model predictions related to the axial dispersion coefficient D_z during Step 1 when the bioreactor operates with a UASB configuration. As expected, a decrease in D_z causes marked profiles along Z^*. Simulation results are less sensitive to this variation at higher upflow velocities, i.e. when operational conditions are similar to EGSB configurations. A decrease in D_z values from 800 $dm^2\ d^{-1}$ to plug flow condition (D_z=0) causes a decrease of 3% in the biogas yield and COD removal efficiency.

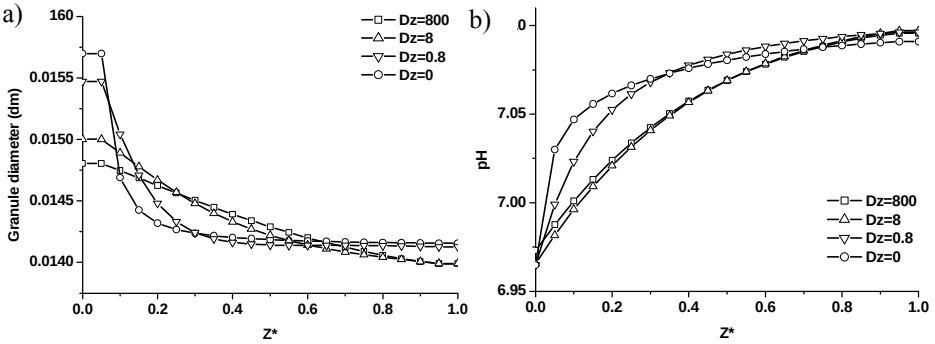


Fig. 4. Sensitivity analysis of model predictions related to the dispersion coefficient D_z ($dm^2\ d^{-1}$): (a) granule diameter, and (b) pH.

A reference value of the specific rate of granule rupture k_r equal to 2×10^{-21} dm d^2 g^{-1} (1.5×10^{-7} m s^2 kg^{-1}) is here used. Values up to $200k_r$ were evaluated. Simulation results for total COD concentration in the effluent stream is practically the same for all k_r values. However, the solid holdup decreases when parameter k_r increases, and hydrodynamic conditions vary. A decrease in the granule mean diameter, and thus, in the terminal settling velocity, causes an increase in the height of reactor. In practice, a higher concentration of suspended biomass and small granules can leave the bioreactor.

References

Angelidaki et. al (1999), A comprehensive model of anaerobic bioconversion of complex substrates to biogas, Biotechnol. Bioeng., 63(5), 363-372.

Fuentes et al. (2007), Anaerobic digestion of carbohydrate and protein-based wastewaters in fluidized bed bioreactors, Lat. Am. App. Res., 37(4), 235-242.

Fuentes et al. (2008a), Hydrodynamic aspects in fluidized bed bioreactor modeling, Chem. Eng. Proc., 47, 1530–1540.

Fuentes et al. (2008b), Anaerobic biofilm reactor modeling focused on hydrodynamics, Chem. Eng. Comm., 195(6), 600-621.

Fuentes et al. (2009), Global modeling and simulation of a three-phase fluidized bed bioreactor, Comp. Chem. Eng., 33, 359-370.

Huang and Wu (1996), Specific energy dissipation rate for fluidized bed bioreactors, Biotechnol. Bioeng., 50, 643-654.

Kalyuzhnyi et al. (2006), Dispersed plug flow model for upflow anaerobic sludge bed reactors with focus on granular sludge dynamics, J. Ind. Microbiol. Biotechnol., 33, 221-237.

Kato et al. (2003), Post-treatment of UASB effluent in an expanded granular sludge bed reactor type using flocculent sludge, Wat. Sci. Tech., 48(6), 279-284.

Nicolella et al. (2000), Wastewater treatment with particulate biofilm reactors, J. Biotechnol., 80, 1-33.

Pontes and Pinto (2006), Analysis of integrated kinetic and flow models for anaerobic digesters, Chem. Eng. J., 122, 65-80.

Tiwari et al. (2006), Influence of extrinsic factors on granulation in UASB reactor. Appl. Microbiol. Biotechnol., 71, 145-154.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Sequencing Batches in a Real-World Pipeline Network Using Constraint Programming

Luiz Carlos Felizari, Lúcia V. R. de Arruda, Ricardo Lüders, Sérgio L. Stebel

CPGEI – Federal University of Technology – Paraná
Av. Sete de Setembro, 3165, 80230-901, Curitiba – PR, Brazil
E-mail address: felizari@cpgei.cefetpr.br

Abstract

This paper presents a new approach for sequencing of batches in a complex pipeline network by using Constraint Programming (CP) techniques. These batches are composed by oil derivatives to be transferred by the pipeline network. Although many papers found in the literature usually deal with reduced pipeline topologies, the network considered in this work is quite complex. It involves nine areas, including three refineries, one harbor and five distribution centers interconnected by fifteen multiproduct pipelines. Although this paper only focus on sequencing of batches, the final task is to provide a complete transfer scheduling for the pipeline network. This scheduling is accomplished by means of three key elements: assignment, sequencing and timing. The proposed CP model considered for sequencing of batches introduces particular features of the problem that can be used by specialized solvers. This approach aims to consider time delays introduced by the pipeline network during product transfers. Preliminary results show that a suitable ordering of batches can be obtained for improving inventory levels. Besides, results show a better performance when compared to a heuristic approach which can fail for some particular cases.

Keywords: Pipeline Network, Scheduling, Sequencing, Constraint Programming.

1. Introduction

Pipeline networks are a reliable and economical transfer way for oil and gas, especially when there is a large quantity of products to be transferred or great distances must be overcome. The distribution and transport of oil refined products involve decisions that should be taken within time periods ranging from hours to months. They include decisions about production, planning and scheduling of operations, inventory management and network utilization rate among others. Extensive reviews of optimization approaches for scheduling of pipelines have been addressed by Cafaro and Cerdá, Magatão *et al.*, Rejowski and Pinto, Relvas *et al.* and Moura *et al*. In general, most of the published works make use of decomposition strategies to reduce the combinatorial complexity for solving scheduling optimization problems in real-world pipelines. In a former work by Neves-Jr *et al.*, the same real-world pipeline network presented in this paper was modeled through MILP techniques using a decomposition approach. The scheduling problem was decomposed into three sub-problems: assignment of resources, sequencing of batches and timing allocation. The solution of assignment and sequencing problems uses constructive heuristics based on specialist knowledge about the pipeline network operation, and solution of timing problem is obtained by a continuous-time MILP model. Since solution of assignment and sequencing are hard to be obtained by a specialist, we propose to reformulate the

sequencing problem by using high-level constructs of Constraint Programming (CP). The main argument of developing MILP/CP approaches is to try to use their complementary strengths in an efficient manner (Magatão *et al*., 2005). The paper is organized as follows. The problem description and the model formulation are presented in sections 2 and 3 respectively. Computational results and analysis are provided in section 4 and conclusions are addressed in section 5.

2. Problem Description and Proposed Approach

The pipeline network represented in Fig. 1 involves nine areas, including three refineries, one harbor and five distribution centers interconnected by fifteen multiproduct pipelines. Besides, some of them can have their flow direction reverted according to operational procedures. Such complexity is dealt by considering three key elements of scheduling: assignment of resources, sequencing of activities and timing allocation for resources (Reklaitis, 1992). Each one is solved by an appropriate technique in a multilevel hierarchical structure (Fig. 2). In our former paper, assignment of resources and sequencing of activities are carried out by a heuristic algorithm. The assignment block takes into account production and demand information received from the company planning, inventory management information as well as actual inventory levels in each area. Based on this information, the assignment block calculates time windows for delivering (or receiving) each batch in each source (or destination) area. In addition, this block determines batches' priorities to be satisfied inside the pipeline network according to the previous calculated time windows. By satisfying these time windows, inventories are kept within operational levels. However, after timing allocation for resources, the obtained scheduling has shown that time windows may not be satisfied. This can be explained by the heuristic approach used for assignment and sequencing, since it does not consider time delays introduced by the pipeline network. For instance, it does not considerer the minimum time necessary to transfer a batch from its source to a final destination along to the pipeline even without bottlenecks. The main idea of this paper is present a new approach to prevent time window violations by using a Constraint Programming (CP) model for batches ordering. Constraint Programming is largely used to represent a scheduling problem in a more "natural" and compacting manner. An important CP feature is given by constraint operators, like OR, AND, IMPLICATION and EQUIVALENCE (Magatão *et al*., 2005). By developing a more formal representation of the sequencing problem, the quality of the pipeline network scheduling can be significantly improved.

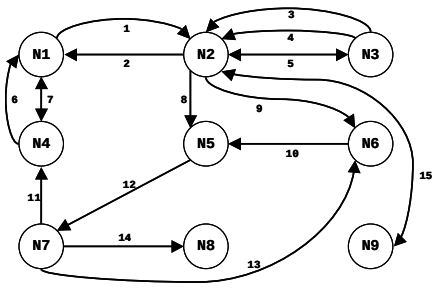


Fig.1 – Pipeline Network

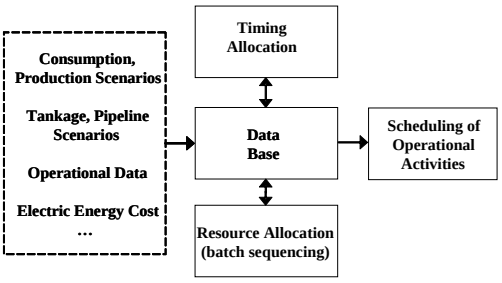


Fig.2 – Optimisation Structure

3. Mathematical Formulation

The proposed mathematical model must represent a sequence of batches that should be pumped in order to determine a minimum time window in each source and destination terminals. It is based on the OPL constraint programming framework supported by ILOG Solver and employs some specific scheduling features available in the ILOG Scheduler. Probably, the most fundamental scheduling concepts in OPL are related to built-in applications such as activity, unary and discrete resources (ILOG, 2002). The proposed CP model is described in the following.

3.1. Objective Function

The objective function in Eq. (1) is composed by time windows violations that should be minimized both in sources ($ao_{b,n}$; $do_{b,n}$) and destinations ($ad_{b,n1}$; $dd_{b,n1}$). The objective function is weighted by operational cost factors (kto; ktd).

$$\min \sum_{b\in B}\sum_{\substack{n\in N\\ n=no_b}}\left(ao_{b,n}+do_{b,n}\right)*kto+\sum_{b\in B}\sum_{\substack{n1\in N\\ n1=nd_b}}\left(ad_{b,n1}+dd_{b,n1}\right)*ktd \qquad (1)$$

3.2. Constraints

The mathematical model is composed by a set of constraints, which are expressed by equations (2) to (15). Eq. (2) establishes time start and finish of a pumping activity $op_{b,n,n1,d}$ in a source node, where vol_b and vaz_b are the volume and pumping flow rate of batch b, respectively. In addition, Eq. (3) defines the duration of activity $op_{b,n,n1,d}$ from second to last pipeline segment through the designated route. The parameter vp_d represents the volume of pipeline segment d.

$$op.end_{b,n,n1d} = op.start_{b,n,n1,d} + vol_b / vaz_b \qquad b\in B, n, n1\in N(n\neq n1), n=no_b, d\in D \qquad (2)$$

$$op.end_{b,n,n1,d} = op.start_{b,n,n1,d} + vp_d / vaz_b \qquad b\in B, n, n1\in N(n\neq n1), d\in D \qquad (3)$$

Delivery and receipt constraints are described by Eqs. (4) and (5), respectively. In each origin area, minimum pumping start time (ted_b) and maximum pumping start time (tec_b) of batch b must be respected. Similarly, the minimum receiving start time (trd_b) and maximum receiving start time (trc_b) must be guaranteed at the final product destination.

$$ted_b - ao_{b,n} \le op.start_{b,n,n1,d} \le tec_b + do_{b,n} \qquad b\in B, n, n1\in N(n\neq n1), n=no_b, d\in D \qquad (4)$$

$$trd_b - ad_{b,n1} \le op.end_{b,n,n1,d} \le trc_b + dd_{b,n1} \qquad b\in B, n, n1\in N(n\neq n1), n1=nd_b, d\in D \qquad (5)$$

Linear disjunction (6) and equivalence constraint (7) establish a precedence condition for pumping activities on shared pipelines. The binary variable $pd_{b,b1,d}$ indicates whether batch b precedes batch $b1$ at a pipeline segment d ($pd_{b,b1,d}$ = 1), or otherwise ($pd_{b,b1,d}$ = 0). The number of shared pipeline segments to batches $b1$ and $b2$ is given by the factor $kpd_{b,b1}$.

$$\left(\sum_{b\in B}\sum_{b1\in Bd\in D}\sum pd_{b,b1,d} = kpd_{b,b1}\right)\vee\left(\sum_{b\in B}\sum_{b1\in Bd\in D}\sum pd_{b,b1,d} = 0\right) \qquad (6)$$

$$pd_{b,b1,d} \Leftrightarrow op_{b,n,n1,d} \textbf{ precedes } op_{b1,n,n1,d} \qquad b,b1\in B, n, n1\in N(n\neq n1), d\in D \qquad (7)$$

Eq. (8) expresses a precedence between distinct activities of a batch during the pipeline transitions along the route.

$$op_{b,n,n1,d} \textbf{ precedes } op_{b,n1,n2,d1} \quad b, \in B, n, n1, n2 \in N(n \neq n1 \neq n2), d, d1 \in D\,(d \neq d1) \tag{8}$$

Constraint (9) determines a precedence between activities of distinct batches at intermediate nodes.

$$\begin{aligned} &op_{b,n,n1,d} \textbf{ precedes } op_{b1,n,n1,d} \Leftrightarrow op_{b,n1,n2,d1} \textbf{ precedes } op_{b1,n1,n3,d2} \\ &b,b1 \in B, n,n1,n2,n3 \in N(n \neq n1 \neq n2 \neq n3), d,d1,d2 \in D(d \neq d1 \neq d2) \end{aligned} \tag{9}$$

A pumping sequence in a source area should satisfy (Eq. 10). For this case, p_b is a product associated to batch b ($p_b \in P$).

$$\begin{aligned} &op_{b,n,n1,d} \textbf{ precedes } op_{b1,n,n1,d} \\ &b,b1 \in B, n,n1 \in N(n \neq n1), d \in D,\ p_b, p_{b1} \in P(p_b = p_{b1}) \end{aligned} \tag{10}$$

Constraint (11) uses a specific OPL construct "**requires**". It enforces that activity $op_{b,n,n1,d}$ needs a *unary resource* from the set of *unary resources* ($d \in D$). Note that this constraint enforces assignment of an activity to a specific pipeline segment as well as a sequencing of activities that have been assigned to this pipeline.

$$op_{b,n,n1,d} \textbf{ requires } d \quad b \in B, n, n1 \in N(n \neq n1), d \in D \tag{11}$$

Constraint (12) uses another fundamental concept provided by OPL for scheduling applications: *discrete resources*. In source areas, *discrete resources* may be used to model a set of pumps. Pumping activities request pump resources to be done. For instance, batches can be pumped from N7 by pipelines 11, 13, and 14. At this node there exist a limited number of pumps and only one batch is allowed to be delivered from N7 in a specific time interval.

$$op_{b,n,n1,d} \textbf{ requires } (1)\ pump_n \quad b \in B, n, n1 \in N(n \neq n1), d \in D \tag{12}$$

Eq. (13) defines a sequencing of batches according to a previous pipeline assignment. Variable ord_b indicates a relative pumping order of batch b into a list of batches. Note that only integer values are assigned to ord_b.

$$\begin{aligned} &op_{b,n,n1,d} \textbf{ precedes } op_{b1,n2,n3,d} \Leftrightarrow ord_b < ord_{b1} \\ &b,b1 \in B, n,n1,n2,n3 \in N(n \neq n1 \neq n2 \neq n3), d \in D \end{aligned} \tag{13}$$

Global constraints (14) and (15) use a predicate "**alldifferent**" which expects an array of discrete variables and holds it if all elements of this array are given a different value (ILOG, 2002).

$$\textbf{alldiferent}\ (op.start_{b,n,n1,d}) \quad b \in B, n, n1 \in N(n \neq n1), d \in D \tag{14}$$

$$\textbf{alldiferent}\ (op.end_{b,n,n1,d}) \quad b \in B, n, n1 \in N(n \neq n1), d \in D \tag{15}$$

3.3. Search Procedure

In Eq. (16), a search component consists of assignment of values for all variables ord_b, using an OPL instruction "**generate**". This generation procedure receives a discrete variable, or an arbitrary array of discrete variables, and generates values for all these variables (ILOG, 2002).

$$\textbf{generate}\ (ord_b) \quad b \in B \tag{16}$$

4. Computational Results

In this section, we compare a sequence of batches generated by the CP model with a sequence previously obtained in the assignment block (rule-based heuristic algorithm). Sequences (due to CP model and due to assignment block) are then processed by the timing block which obtains the respective timing for each sequence through a MILP model. Further details about MILP model can be found in Neves-Jr *et al*. Table 1 presents computational details for seven real-world instances (S1 to S7) considering scheduling horizons of approximately 30 days. The resource allocation block takes production/consumption plans of more than 10 different products during a month. As shown in Table 1, the sequencing model found a solution in few CPU seconds, even for scenarios with more than 80 batches. The reduced number of binary variables is function of batch conflicts in the network. For instance, variable $pd_{b,b1,d}$ is generated when batches $b1$ and $b2$ are pumped from nodes N1 and N3, respectively, and share the pipeline 8. Otherwise, the variable $pd_{b,b1,d}$ will become redundant. Similarly, the set of constraints is highly dependent of batches relationship. Forbidden product sequences, specified *a priori* by the assignment block contributes to a reduction of the number of precedence constraints and, consequently, to lower computational effort. In Table 1, this can be seen for scenarios S5 (78 bathes and 3268 constraints – 14.7s) and S6 (81 batches and 2854 constraints – 14.2s). The second part of Table 1 shows results for heuristic and Constraint Programming approaches (H/CP). A performance metric used to compare results is the total number of batches delivered or received beyond time windows computed in the assignment block. According to section 2, time windows violation is related to violations on inventory level in refinery and terminal tank farms, but it is undesirable. For instance, at the scenario S1, for the CP model, only 6 batches have their time windows violated against 19 batches to the specialist heuristic. However, in scenario S2, CP model provided a worse result than one obtained by a heuristic solution (from 4 to 5 constraint violations), but not so far.

Table 1. Results for operational scenarios

Scenario			**S1**	**S2**	**S3**	**S4**	**S5**	**S6**	**S7**
Number of Batches			49	67	67	72	78	81	84
Binary Variables			96	82	88	214	274	115	255
Continuous Variables			1652	1974	2001	2182	2381	2100	2484
Constraints			1703	2168	2184	2809	3268	2854	3442
Execution Time [s]			5.45	9.13	9.28	11.64	14.7	14.2	17.03
Time Window Violation	Delivery	Early (H/CP)	7/5	2/3	13/9	3/2	3/2	4/0	4/3
		Late (H/CP)	0/0	0/0	0/0	0/0	3/1	1/0	7/0
	Receipt	Early (H/CP)	2/0	1/0	0/0	3/2	1/0	0/0	0/0
		Late (H/CP)	10/1	1/2	12/6	4/1	3/0	0/0	6/3
	Total Violation (H/CP)		19/6	4/5	25/15	10/5	10/3	5/0	17/6
Makespan [days] (H/CP)			31/31	34/35	33/31	31/33	32/32	27/29	32/33

The proposed CP formulation was implemented and solved in ILOG OPL Studio 3.6.1 using ILOG Solver 5.3 and ILOG Scheduler 5.3 on Intel-Core2Duo 2.13GHz with 2GB of RAM.

5. Conclusions

We have developed a Constraint Programming (CP) model for sequencing of batches in a pipeline network. The incorporated approach has been tested with real-world scenarios, that covered monthly time horizons involving more than 80 batches. In order to achieve feasible solutions in acceptable computational times, a decomposition approach was used. This decomposition relied on a multi-level hierarchical structure involving the assignment, the sequencing, and the timing sub-problems of scheduling. Since the ordered sequence of pumped products has an important impact over the final short-term scheduling, the activities ordering heuristic used by the assignment part can fail for particular cases. Compared to the approach discussed in a previous work addressed by Neves-Jr *et al.*, this paper has aimed to reformulate the sequencing part using the high-level constructs of Constraint Programming (CP). In general, the proposal CP model obtained satisfactory solutions with a low computational effort (up to 70% reduction in the number of time windows violation). Furthermore, the model results can present a better level of inventory during the time horizon. Finally, for future work over the sequencing model, a more detailed formulation would enable to obtain optimised sequences with reduced product interfaces.

Acknowledgements

The authors acknowledge financial support from ANP and FINEP (PRH-ANP / MCT PRH10 UTFPR), PETROBRAS (grant nº 0050.0017859.05.3).

References

D. C. Cafaro and J. Cerdá, 2008, Dynamic scheduling of multiproduct pipelines with multiple delivery due dates, Computers and Chemical Engineering, v.32, pp. 728-753.

L. Magatão, L. V. R. Arruda and F. Neves-Jr, 2004, A mixed integer programming approach for scheduling commodities in a pipeline, Computers and Chemical Eng., v.28, pp. 171-185.

L. Magatão, L. V. R. Arruda and F. Neves-Jr, 2005, Using CLP and MILP for scheduling commodities in a pipeline, Computer Aided Chemical Eng., v.20B, pp. 1027-1032.

R. Rejowski Jr. and J. M. Pinto, 2008, A novel continuous time representation for the scheduling of pipeline systems with pumping yield rate constraints, Computers and Chemical Eng., v.32, pp. 1042-1066.

S. Relvas, H. A. Matos, A. P. F. D. Barbosa-Póvoa, J. Fialho, 2007, Reactive Scheduling Framework for a Multiproduct Pipeline with Inventory Management, Ind. Eng. Chem. Res., v.46, pp. 5659-5672.

A. V. Moura, C. C Souza, A. A. Cire, T. M. T. Lopes, 2008, Heuristics and Constraint Programming Hybridizations for a Real Pipeline Planning and Scheduling Problem, in 11th IEEE International Conference on Computational Science and Engineering, pp. 445-462.

F. Neves-Jr, L. V. R. Arruda, L. Magatão, S. L. Stebel, S. N. Boschetto, L. C. Felizari, D. I. Czaikowski, R. Rocha, P. C. Ribas, 2007, An efficient approach to the operational scheduling of a real-world pipeline network, in proceedings of 17th European Symposium on Computer Aided Process Engineering, Bucareste, Computer Aided Chemical Engineering (ESCAPE 17), Amsterdan: Elsevier, v. 24, pp. 697–702.

G. V. Reklaitis, 1992, Overview of scheduling and planning of batch process operations, Proceedings of the NATO, Antalya, Turkey, pp. 660-675.

ILOG OPL Studio 3.6.1, 2002, User's Manual, ILOG Corporation, France.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal Experimental Design for the Determination of Protein Ion-Exchange Equilibrium Parameters

Tilman Barz, Verena Löffler, Harvey Arellano-Garcia, Günter Wozny

Chair of Process Dynamics and Operation, Sekr. KWT-9
Berlin Institute of Technology, D-10623 Berlin, Germany

Abstract

Despite the fact that batch experiments directly establish equilibrium conditions, they still have some main disadvantages, which make the determination of equilibrium parameters a challenging task. These drawbacks are among others the lack of direct measurement of the key variables, the impossibility of setting a desired equilibrium point without a perfect process model, and the high laboratory effort for the preparation and analysis of the experiments. To overcome these problems, a systematic approach based on optimal experimental design is proposed. Here, the inclusion of practical restrictions in the optimization problem formulation is crucial such as the maximal protein amount available, the maximal solubility of protein and salt, or the special characteristics of peak area determination during the measurement of protein concentrations using an analytical HPLC. The efficiency of the developed parallel/sequential design approach is demonstrated by theoretical and experimental results for the multi-component (two proteins/salt mixture) system.

Keywords: optimal experimental design, steric mass ion-exchange, adsorption equilibrium.

1. Introduction

The exact determination of adsorption equilibrium parameters still remains a challenging task. The knowledge of the thermodynamic equilibrium between the mobile phase, the stationary phase, and the components of the mixture to be separated is the fundamental basis for e.g. modeling and simulation of chromatographic applications. In this work, optimal batch experiments are designed in order to determine protein ion-exchange equilibrium parameters. In comparison to conventionally planned experiments, the optimal experimental design (OED) method allows minimizing the experimental effort to be spent for the preparation and analysis of experiments. Furthermore, a quantitative measure for the quality of a certain set of experiment decision variables is provided, which gives insight about the quality and identifiability of the estimated parameters. For this purpose, the size of the parameter covariance matrix is used as a common criterion. Moreover, different metrics are adopted in order to quantify the parameter quality (Franceschini & Macchietto, 2007). However, prerequisite for the model based experimental design is that an appropriate and valid model without structural uncertainties is at hand. Furthermore, in order to cope with parameter uncertainties, different strategies such as parallel, sequential, and parallel/sequential design of experiments can be adopted (Franceschini & Macchietto, 2007). The sequential strategy is commonly adopted, which is based on an iterative

refinement of the estimated parameter set, and thus, it is reliable even for bad initial guesses of the parameters to be estimated. In this work, it will be shown that for the considered batch experiments here, a strictly sequential strategy is not feasible for two reasons: first, it would boost the experimental effort at a maximum, and secondly, it also does not guarantee identifiability of the parameter set for the very first batches. The chosen case study in this work is the separation of bio-molecules from milk whey products. Under the commercially interesting products are the proteins β-Lactoglobulin A and B (βLgA, βLgB), which can be separated by anion-exchange chromatography (Pedersen, Mollerup, Hansen & Jungbauer, 2003). Generally, the wide range of retentivities of the whey proteins makes it difficult to separate them under isocratic conditions. Thus, the application of gradient elution to preparative- and large scale chromatography is the most common mode of operation (Natarajan, Ghose & Cramer, 2002). In gradient elution, a modulator concentration in the liquid phase is adjusted in order to influence the eluent strength, and thus the retentivities of the eluites (Gu, Truei, Tsai & Tsao, 1992). For the treatment of βLgA and βLgB, Cl^- ions (salt) are used as modulator in order to optimize the protein separation. Therefore, the salt is considered as an additional component in the equilibrium model.

1.1. Estimation and Modeling of Protein Ion-Exchange Equilibria

A general overview about typical methods for the determination of adsorption isotherms is given in (Seidel-Morgenstern, 2004). Using batch experiments means a tedious and fault prone laboratory work, but in comparison to alternative methods, no additional effects like transport or kinetic phenomena have to be considered in order to determine the adsorption equilibrium (Seidel-Morgenstern, 2004). For the explicit consideration of non-isocratic operation in chromatographic protein separation, Brooks and Cramer developed a steric mass (SMA) ion-exchange equilibrium formalism, which explicitly accounts for the steric hindrance of salt counterions upon protein binding in multicomponent equilibria. The SMA formalism can be represented in analogy to the variable coefficient in multicomponent Langmuir isotherms (Brooks & Cramer, 1992):

$$q_j^* = f\left(c_j^*, \boldsymbol{\theta}\right); \quad \text{with } \boldsymbol{\theta} = \left[\Lambda, k_{1,j}, \nu_j, \sigma_j\right]^T \text{ and } j = \{2,...,N_C\} \tag{1}$$

where N_C=3 is the number of components, c_j^* and q_j^* represent the liquid and the adsorbed equilibrium concentration, respectively. For the first component, here the salt, the parameters are fixed, and thus it does not need to be determined. Furthermore, since the proteins βLgA and βLgB have a similar structure, the steric factor of the proteins, σ_i, can be assumed to be equal for both proteins (σ_2=σ_3). Thus, there are 6 adsorption model parameters, $\boldsymbol{\theta}$, to be estimated by the available measurement.

2. Experimental Design problem formulation

The experimental design aims to adjust the initial amount of solutes and adsorbent in the batch in order to establish a certain equilibrium point. However, the decisions in order to obtain an specific equilibrium point are not obvious, which is basically due to the diminution of the initial protein concentration to the equilibrium concentration (Seidel-Morgenstern, 2004). It is obvious that using a model these decisions can be calculated, but unfortunately the model's accuracy will then depend directly on the quality of the initial estimate of the equilibrium parameters, $\boldsymbol{\theta}$. Moreover, critical for the formulation of the experimental design problem is the inclusion of practical restrictions in the optimization problem formulation. Therefore, the manual steps for preparing and analyzing batch experiments are first briefly discussed. Following this, a suitable model for the OED problem formulation is derived.

2.1. Experimental Setup, Batch Preparation and Measurements

For each batch preparation four different liquid volumes are arranged, where several decisions have to be made (see also Fig 1):

- volume, V^{sl}, and adsorbent amount, ε_{tot}^{sl}, for the adsorbent slurry
- volumes, V_2^{std} , V_3^{std}, and concentrations, $c_{2,2}^{std}$, $c_{3,3}^{std}$, for the protein standards
- and finally the concentration, $c_{1,1}^{std}$, for the salt standard, and the required volume, V_1^{std}, so as to get a desired total batch volume, V^{bt}

Furthermore, some additional parameters which result from the special pre-treatment of each batch preparation are considered, where c_1^{el} and $q_1^0 = \Lambda$ represent a constant free salt concentration, and the amount of salt in the beginning. The adsorbent is loaded with the maximal amount of salt at the start.

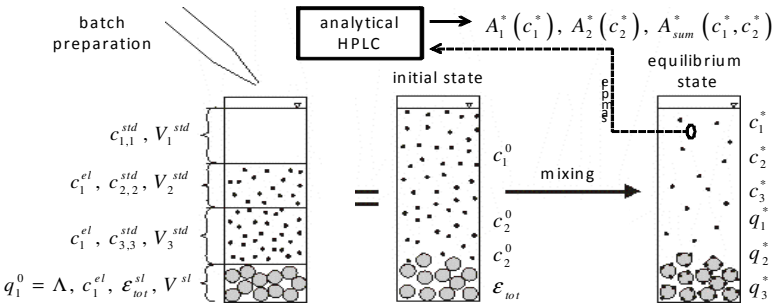


Figure 1: Preparation, initial and equilibrium state of one single batch experiment.

The major drawback lies on the fact that for batch experiments only liquid concentration measurements are available. The adsorbed equilibrium amount is determined based on the difference between the initial, c_i^0, and the equilibrium liquid concentrations, c_i^*, after mixing (Fig. 1). Here, no measurement of the salt ion concentration is carried out mainly because of the fact that the small batch size (4 ml) makes it impractical to implement a conductivity measurement. Consequently, information about neither the liquid salt, c_1^*, nor the adsorbed salt equilibrium concentration, q_1^*, is available. The analysis of the protein concentration is done by peak area integration of the photometric outlet signal and a linear calibration curve using an analytical reverse phase HPLC. Here, a typical problem in the HPLC analysis is encountered, which is the no-complete-base-line separation of the two peak areas, A_2^* and A_3^*, obtained for the equilibrium liquid concentration, c_2^* and c_3^*. Thus, the separation of the connected areas is done manually in order to calculate the concentrations c_2^* and c_3^*. By including the calibration curve into the model, both the total area, A_{sum}^*, and the alienated areas, A_2^* and A_3^*, are considered. According to the fault prone separation by hand, a relatively high and a much smaller measurement accuracy is assigned to A_{sum}^* and A_2^*, A_3^*.

2.2. Analysis of the Degree of Freedom and Mathematical Problem Formulation

For the preparation of one initial batch, all the referred parameters above are rather over-determining the decision space for the experimental design (e.g. a smaller standard protein concentration as well as a smaller standard volume can give the same initial concentration). The degree of freedom for the design of one single batch experiment is: $DoF=N_C+1$. Thus, some parameters are set to constant values such as the adsorbent amount. The standard protein concentrations are set to their maximal solubility. The adjustable variables are the three volumes of the protein standards and the slurry, as well as the salt concentration. Doing so, the relatively high effort for the preparation of

one single experiment does not increase significantly for the preparation of multiple experiments. The objective function for the OED of multiple experiments reads,

$$\begin{aligned} &\min_{\mathbf{u}} \ \varphi\left\{\boldsymbol{\theta}\mathbf{V}\left(\boldsymbol{\theta},\mathbf{u}\right)\right\} = \operatorname{trace}\left[\boldsymbol{\theta}\mathbf{V}\left(\boldsymbol{\theta},\mathbf{u}\right)\right] \\ &\mathbf{u}_k = \left[V_2^{std}, V_3^{std}, V^{sl}, c_{1,1}^{std}\right]^T \ ; \ \ \mathbf{u} = \left[\mathbf{u}_1,...,\mathbf{u}_{N_{Exp}}\right]^T \\ &\boldsymbol{\theta} = \left[\Lambda, k_{1,j}, \nu_j, \sigma_j\right] ; \ \sigma_2 = \sigma_3 \ ; \ \ j \in \{2,3\} \\ &\mathbf{y}_k^{meas} = \left[A_2^*, A_3^*, A_{sum}^*\right] \end{aligned} \quad , \tag{2}$$

where the parameter-covariance matrix, $\boldsymbol{\theta}\mathbf{V}$, represents the inference regions of each of the model parameters. A measure of this region is φ, in this specific case, the trace of the matrix, which is also known as the A-criterion. The decisions of the design problem, $\mathbf{u}$, are chosen in order to maximize the expected information content from the measured data generated by these experiments. For an estimate of the parameter-covariance matrix, $\boldsymbol{\theta}\mathbf{V}$, the inverse of the Fischer-matrix, $\mathbf{F}$, is used.

$$\boldsymbol{\theta}\mathbf{V}\left(\boldsymbol{\theta},\mathbf{u}\right) \approx \mathbf{F}^{-1}\left(\boldsymbol{\theta},\mathbf{u}\right) = \left(\sum_{k=1}^{N_{Exp}} \mathbf{F}_k\left(\boldsymbol{\theta},\mathbf{u}_k\right)\right)^{-1} \tag{3}$$

For one single experiment, $\mathbf{F}_k$ is calculated from the sensitivities of the measured variables, $\partial\mathbf{y}_k^{meas}/\partial\boldsymbol{\theta}$, weighted with the inverse measurement-covariance matrix, $\mathbf{MV}_k$.

$$\mathbf{F}_k = \left(\frac{\partial \mathbf{y}_k^{meas}}{\partial \boldsymbol{\theta}}\right)^T \cdot \mathbf{MV}_k^{-1} \cdot \left(\frac{\partial \mathbf{y}_k^{meas}}{\partial \boldsymbol{\theta}}\right) \tag{4}$$

By this means that previously carried out experiments, $k=\{1,\ldots,f\}$ are considered by a constant Fischer-Matrix, $\mathbf{F}_k$, which is then part of the sum over all experiments, but with their respective experimental design variables, $\mathbf{u}_k$, not included in the decision space, $\mathbf{u}$. Nevertheless, during the adsorption process each of the three components has to fulfill the mass balance. Together with the equilibrium model and additional equations describing the batch preparation and measurements, a total number of $N_{eq}=16*N_{Exp}$ algebraic model equations is obtained. The degree fo freedom is $DoF=(NC+1)*N_{Exp}$. Additionally, linear constraints restricting the total amount of proteins available, as well as the desired batch volume, and the maximal salt concentration, $c_{1,1}^{std}$, are considered. Following the sequential optimization approach, the model equations are solved in a sub-routine during the optimization run, and thus, only the experimental design parameters, $\mathbf{u}$, and the linear constraints are considered in the optimization routine. The resulting optimization problem is (linearly) constrained, small, dense and strongly nonlinear in the objective function and can be solved efficiently using the solver NPSOL (Gill, Murray, Saunders & Wright, 1985).

3. Theoretical and Experimental Results

Prior to the solution of the OED problem, one has to decide on both the strategy to follow (sequential vs. parallel) and the number of experiments, N_{Exp}, to be planned. At first glance, a simultaneous design of experiments seems to be the first choice for the described problem as it corresponds to the standard procedure in regular lab work, where several batches are produced and analyzed in parallel. The alternative is the sequential design approach, where only one experiment is planned. Another option is a combination of both strategies, where parallel planned experiments are designed sub-sequentially.

3.1. Theoretical results regarding the sequential and parallel design of experiments

The following points are of special interest for the decision of the design strategy:

- Prerequisite for the solution of the OED problem is that for the selected number of experiments, identifiability of the model parameters, $\boldsymbol{\theta}$, is reached. Thus, the number of experiments to be planned depends on the quality of the available measurement, $\mathbf{y}^{meas}$.
- Deviations between the initially assumed parameter estimates, $\boldsymbol{\theta}^0$, (used to solve the OED problem) and the (later estimated) "true" parameter set, $\boldsymbol{\theta}^*$, can strongly influence the quality of the experimental design and consequently the inference region of $\boldsymbol{\theta}^*$ (e.g. for large deviations between $\boldsymbol{\theta}^0$ and $\boldsymbol{\theta}^*$, an iterative design and parameter estimation is advantageous).
- Parallel planned experiments can save time and experimental effort.
- Finally, it could be argued that the simultaneous consideration of multiple experiments leads to the maximum information content by perfectly "orthonormal" design conditions, whereas sequentially planned experiments do not exploit the decision space, $\mathbf{u}$, as much efficient as possible.

Figure 2, left, shows theoretical studies about the course of the A-criterion for different design strategies with fixed parameter $\boldsymbol{\theta}^0$. It can be seen that for a minimum of four batch experiments, identifiability of all six parameters is reached. Surprisingly, the sequential design strategy (4+1+…+1) and even the nearly parallel one (10+10) does not affect the course of the A-criterion.

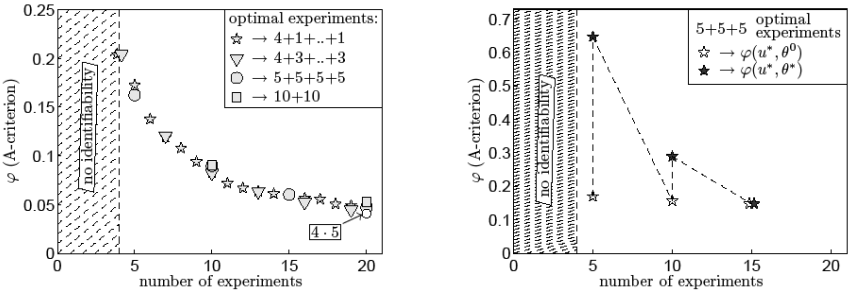


Figure 2: Course of the experimental design criterion over the number of experiments planned, here a combination of sequential and parallel design strategy is chosen, (e.g. 4+1+…+1 means: 4 parallel planned experiments and subsequently several single planned experiments). *Left figure*: theoretical results for one initial parameter set, $\boldsymbol{\theta}^0$=*const.*; *right figure*: experimental results, where the measurement data is used in order to update the parameter estimate, $\boldsymbol{\theta}^0 \rightarrow \boldsymbol{\theta}^*$, after each set of carried out optimal experiments.

Most remarkable is the result when exactly the same optimal design variables, $\mathbf{u}_k$, which were used previously for the first five experiments, are then used again in the subsequent four experiments (4*5 strategy). The A-criterion reaches the same value as for the other strategies. This can only be justified by the low information content of the measurement (see section 2.1). The decrease of the A-criterion for a rising number of experiments is because of an increased accuracy of the measured data (due to repeated measurements) rather than because of new information (e.g. analysis of the adsorption behavior at a different equilibrium point).

3.2. Experimental results for a combined sequential and parallel approach

Figure 2, right, shows the experimental results applying the (5+5+5) strategy. After each five experiments, the parameters are updated using the available measurement and

solving a parameter estimation problem, $\boldsymbol{\theta}^0 \rightarrow \boldsymbol{\theta}^*$. The updated parameters, $\boldsymbol{\theta}^*$, are then used for the subsequent OED. It can be noticed that the A-criterion is strongly affected (degradation) by an imprecise initial estimate of the parameters. With an increasing number of experiments, the quality of $\boldsymbol{\theta}^0$ also increases, and thus, this effect on the A-criterion is minimized. Figure 3 shows both conventionally planned experiments at two different salt ion concentrations (40 batches), and the optimal planned experiments (15 experiments). For the 15 already optimally planned experiments and a much smaller protein consumption, a value of φ_{opt}=0.193 is obtained instead of φ_{conv}=516.0. These values correspond to an average relative standard deviation of the parameters (σ_θ=11.7% and 615%) taken from $\boldsymbol{\theta}\mathbf{V}$, and thus highlighting the benefits of the applied method.

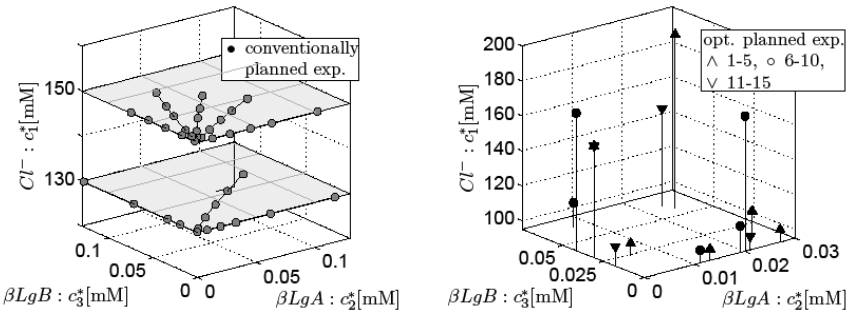


Figure 3: Liquid concentrations for the equilibrium points of conventionally planned experiments, and those which result from the solution of an OED problem.

4. Conclusions

This work proposes a systematic procedure based on a detailed model and the inclusion of design and analysis of experiments for a reliable determination of protein ion-exchange equilibrium parameters. The results show a significant improvement in terms of lab work load, number of experiments required, as well as the necessary protein amount. Moreover, it could be demonstrated experimentally that a benefit of the optimal design strongly depends on the quality of the initial guess of the parameters to be estimated. Thus, for instance, without any information about the parameter range, a sequential approach for OED is more appropriate than an entirely parallel one.

References

Brooks, C. A. & Cramer, S. M. (1992). Steric mass-action ion exchange: Displacement profiles and induced salt gradients. *AIChE Journal, 38* (12), 1969-1978.

Franceschini, G. & Macchietto, S. (2007). Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science , 63* (19), 4846-4872.

Gill, P. E., Murray, W., Saunders, M. A. & Wright, M. H. (1985). Some issues in implementing a sequential quadratic programming algorithm. *ACM Signum Newsletter, 20* (2), 13-19.

Gu, T., Truei, Y. H., Tsai, G. J. & Tsao, G. T. (1992). Modeling of gradient elution in multicomponent nonlinear chromatography. *Chemical Engineering Science, 47* (1), 253-262.

Natarajan, V., Ghose, S. & Cramer, S. M. (2002). Comparison of linear gradient and displacement separations in ion-exchange systems. *Biotechnology and Bioengineering, 78* (4), 365-375.

Pedersen, L., Mollerup, J., Hansen, E. & Jungbauer, A. (2003). Whey proteins as a model system for chromatographic separation of proteins. *J.of Chromatography B, 790* (1-2), 161-173.

Seidel-Morgenstern, A. (2004). Experimental determination of single solute and competitive adsorption isotherms. *Journal of Chromatography A, 1037* (1-2), 255-272.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Integrating Purchase Contract Decisions with Inventory Management Optimization in the Supply Chain

M. Analia Rodríguez[a] and Aldo Vecchietti[a]

[a]*INGAR (CONICET – UTN), Avellaneda 3657, Santa Fe 3000, Argentina*

Abstract

This work addresses the optimization of material inventory management and purchase problems by linking two important players in the supply chain: material suppliers and product manufacturers. The context modeled considers a medium term horizon planning which presents a seasonal demand and a variable restricted capacity from the suppliers. Traditional mathematical constraints and disjunctive programming are used for formulating the problem. The execution performance shows the formulation efficiency and the results analyzed support the importance of solving the problem.

Keywords: inventory optimization, purchase decisions, supply chain.

1. Problem Description

The inventory management is an important problem to address in every industry due to the trade-off between customer satisfaction and capital invested. From an inventory cost perspective, the best condition would be a response-based supply chain where a zero inventory strategy is handled. However, it is rarely possible to achieve that target in practice (**Bowersox et al., 2007**). When products are made-to-stock planning the material purchase depends on the production plan while when products are made-to-order material purchase is more dependent on the demand variations. As a consequence, inventory management is much simpler in the first place than in the second one. Demand is difficult to predict and could lead to stock-out situations since suppliers could have insufficient capacity to provide the appropriate material to all manufacturers. In several industries, such as the production of paper, furniture, textile and food, among others, the use of material family is a regular practice. Companies use a set of possible raw materials (material family), in order to satisfy product demand, manufacturing a product with similar characteristics and properties. This production policy gives flexibility to production decisions allowing a set of possible formulations for the same final product. In this situation if there is no stock of the requested material according to the customer order, they satisfy the order anyway using one material of better quality, from the family, increasing the production costs, commonly called upgrade costs. In other words, the two main alternatives are whether to anticipate the demand buying the corresponding raw material increasing inventory costs; or whether to use better quality and expensive materials to satisfy the demand increasing production costs. For these reasons, decisions concerning material purchase and inventory management are really relevant and should not be considered as decoupled subjects. Another common practice in the industry is to sign provision contracts with suppliers with the purpose of decreasing costs and provision uncertainties (**Park et. al, 2006**). Diverse materials are purchased from various manufacturers signing different contract types. Contracts model

different situations in terms of prices, minimum and maximum quantities and payment conditions. Materials can be provided by various suppliers with diverse costs.
In this paper products made-to-order are studied with a very short lead time expected by the customers, modeling trade-off decision between increasing stock and upgrading production. This context is even more complicated because product demand is influenced by seasonal variations. An optimization approach to solve material purchase and inventory management problems for the supply chain is presented. The main novelty of this work is that the optimization of the inventory management is integrated to material purchase decisions handling different contract types with suppliers.

2. Model Formulation

2.1. Objective function
The objective function (1) defines the actualized profits maximization over the time horizon. The costs involved are the material purchase and the inventory costs while the revenues are calculated by multiplying the product demand and the product price.

$$Max \sum_{t} \frac{\sum_{p} Demand_{pt} \cdot price_{pt} - \left(\sum_{j} \sum_{c} \sum_{k} m_{jckt} + \sum_{f} savg_{ft} \cdot COSTavg_{ft} \cdot MS \right)}{(1+RR)^{t}} \quad (1)$$

Product demand $Demand_{pt}$ is a parameter used to calculate company revenues by multiplying it by the product price $price_{pt}$. In the second term, the money outflow is determined as the positive variable m_{jckt}. It is calculated considering material k purchased from supplier j to be paid in period t according to payment policy of contract c. This variable includes also the fixed cost assumed when selecting the corresponding contract c with supplier j. In the next term, the inventory carrying cost is calculated. For this purpose, $savg_{ft}$ is a positive variable that represents the average stock of family f in period t. The parameter $COSTavg_f$ corresponds to the average cost of family f and parameter MS indicates a percentage of the raw material average costs. This percentage estimates the average stock costs including the following major components: capital, storage, obsolescence and quality costs. As a general practice, these expenses are considered as a single inventory cost rate, represented as a percentage of the material price per unit time. In order to actualize the profits projected in the time horizon, the return rate RR corresponds to the own capital cost which is calculated with the Capital Asset Pricing Model.

2.2. Model constraints
The main constraints are presented in this section.

$$s_{ft} = s_{f(t-1)} + \sum_{j} \sum_{k \in FK_{fk}} q_{jk(t-1)} - d_{f(t-1)} \qquad \forall f \in F, \forall t \geq t_2 \in T \quad (2)$$

The initial stock of family f in period t, s_{ft}, is calculated by (2) as the initial stock of family f in period t-1 plus the quantity ordered, $q_{jk(t-1)}$, of materials k of family f to all suppliers selected in t-1 minus the material family consumption in the previous period $d_{f(t-1)}$. Note that material consumption is calculated considering the number of units of final product demand and the material consumption of each product as shown in (3).

$$d_{ft} = \sum_{p \in PL_{plf}} \sum_{l \in PL_{plf}} Demand_{pt} \cdot \alpha_{plf} \qquad \forall f \in F, \forall t \in T \quad (3)$$

A product p is formed by components l, each of which consumes one material family. It is possible that more than one component uses a certain family f. The formula (3) calculates material consumption d_{ft} according to product demand multiplying the

$Demand_{pt}$ with the parameter α_{plf} which indicates the quantity of family *f* corresponding to component *l* of product *p*, where PL_{plf} relates each component *l* of *p* with the family *f*.

$$\sum_{f} s_{ft} \leq SC \qquad \forall t \in T \qquad (4)$$

Storage capacity is actually finite; equation (4) limits the quantity in stock according to stock capacity parameter *SC*. Other constraints include the initial stock in the horizon planning, which is not null; and material average stock definition.

2.3. Modeling purchase decisions

The decision process involves several levels. Each level is modeled as a disjunction which is then associated with another level using logic equations and discrete variables.

$$\begin{bmatrix} y1_{jft} \\ \sum_{k \in FK_{fk}} q_{jkt} \leq \sum_{k \in FK_{fk}} Qmax_{jkt} \end{bmatrix} \vee \begin{bmatrix} \neg y1_{jft} \\ \sum_{k \in FK_{fk}} q_{jkt} = 0 \end{bmatrix} \qquad \forall j \in J, \forall f \in F, \forall t \in T \qquad (5)$$

Disjunction (5) expresses the selection of which families *f* are bought from each supplier *j* in period *t*. This choice represents the first level in the decision process using boolean variables $y1_{jft}$. If this variable is false then family *f* is not ordered from supplier *j* in period *t*. If it is true, the total quantity ordered must be lower than the maximum capacity of supplier *j*. In general, this quantity diminishes in high season periods. However, the reduction is not the same for all materials in the family. Cheaper materials availability tends to decrease more than the expensive ones (See model results).

$$\begin{bmatrix} y2_{jkt} \\ q_{jkt} \leq Qmax_{jkt} \end{bmatrix} \vee \begin{bmatrix} \neg y2_{jkt} \\ q_{jkt} = 0 \end{bmatrix} \qquad \forall j \in J, \forall k \in FK_{fk}, \forall t \in T \ (6)$$

In the following level given by disjunction (6), variable $y2_{jkt}$ indicates if material *k* is selected from family *f*, according to FK_{fk}, in period *t*. The set FK_{fk} defines which materials correspond to each family. In this term, the quantity ordered of material *k* in period *t* cannot be larger than the supplier capacity for that material and period. If a material *k* is not selected from family *f* then no quantity can be ordered from supplier *j* in *t*. Note that more than one material could be selected for each family. This situation will occur when the most convenient material presents lower availability ($Qmax_{jkt}$) than required. So, there is a commitment between buying material in advance increasing inventory costs and selecting materials of better quality affecting purchase value.

$$\bigvee_{\substack{c \in C \\ (c,t,t') \in TP_{ctt'}}} \begin{bmatrix} y3_{jckt} \\ q_{jkt} \geq Qmin_{cj} \\ w_{jckt} = q_{jkt} \cdot PC_{jkt} \cdot (1 - \delta_{jc}) + FC_{jc} \\ w_{jckt} = m_{jckt'} \end{bmatrix} \begin{array}{l} \forall j \in J, \forall k \in K, \forall t \in T \qquad (7) \\ \delta_{jc_1} = 0; \delta_{jc_2} > 0; \delta_{jc_3} > \delta_{jc_2}; \delta_{jc_4} < 0 \end{array}$$

Disjunction (7) selects the contract type using variable $y3_{jckt}$ for each material *k*, supplier *j* in period *t*. The first equation constrains the quantity of material that at least must be ordered from the supplier. This quantity $Qmin_{cj}$ depends on the contract and the supplier also. Note that while $Qmax_{jkt}$ represents an upper bound for variable q_{jkt}, in this case, parameter $Qmin_{cj}$ determines a lower bound for the order quantity.

The second constraint calculates the costs w_{jckt} of purchasing material *k* according to the amount bought q_{jkt}, price PC_{jkt} and corresponding contract discount ($\delta_{jc} > 0$) or interest rate ($\delta_{jc} < 0$). This cost also includes the fixed cost FC_{jc} to be faced whenever contract *c* is signed with supplier *j*. In the third constraint, positive variable $m_{jckt'}$ determines the amount to be paid for buying material *k* from supplier *j* in period *t'* according to the

payment policy of contract c. In this case, the set $TP_{ctt'}$ establishes that the purchase ordered in period t must be paid in period t' according to contract c. Note that when $t=t'$ there is no financial benefit but when $t<t'$ the financial benefits is given by the difference between them. In the first three contracts, period t is equal to t', however, the fourth one has a longer payment term. Although variable $m_{jckt'}$ could be calculated directly in the second constraint avoiding the use of w_{jckt}, this formulation is considered worthy in order to distinguish the cost concept given by the second equation and the effective money outflows calculated in the last one.

The first contract modeled the purchase with neither requirements nor benefits. For that reason, the minimum quantity requested is zero and the cost is calculated using the regular price with no discount. The second contract type presents a minimum quantity to order given by $Qmin_{c2j}$. It also has a discount applied to the total amount ordered. The third one has a bigger minimum quantity to order, $Qmin_{c3j}$, and a bigger discount too. With the aim of longer business relations with the suppliers, the model constrains the selection of this contract type only in case this material k has already been ordered from supplier j in the previous period. In the last contract type, the minimum quantity required $Qmin_{c4j}$ is the highest and due to the payment term offered, the second equation considers an interest rate ($\delta_{jc4} < 0$). This contract type is the only in which there is a possibility of paying later than the moment the actual purchase is made.

There are logical constraints related to material selection and some others equations guaranteeing coherence in the decision process not shown due to space reasons.

3. Model Results

The formulation was posed in GAMS system with the aim of showing model performance and solution suggested. It was executed over a PC having an Intel Pentium D 2.8 GHz processor. Disjunctions were modeled using LogMIP (**Grossmann et al., 2005**). Table 1 summarizes execution performance. The objective value was 5686.51.

Table 1. Solution performance.

	Equations	Postive Variables	Discrete Variables	Execution Time
Solution performance (Optimal solution)	2593	1777	822	13.906 sec

The main result is represented by variable q_{jkt} shown in Table 2. The lack of material availability during high season forces the company to accumulate materials in previous periods. This increases inventory handling and financial costs but not production costs.

Table 2. Material quantity decision.

q_{jkt}	j_1k_1	j_1k_2	j_1k_4	j_1k_5	j_1k_8	j_2k_1	j_2k_4	j_3k_1	*Quantity ordered*	**Qmax active** (Equation (5))
t_1	**500**	538	**400**			**450**	**420**		**2308**	
t_2	**200**	**300**	**200**	281	***85***	**200**	**200**		**1466**	***Qmin active*** (Disjunction (7))
t_3			**150**	**200**	175				**525**	
t_4		**200**	185	***110***	120			**200**	**815**	

According to these results, in many occasions equation (5) actually restricts the solution while in only two cases the first equation of disjunction (7) is also active. Stock capacity is not a limitation since SC is equal to 5000 and the maximum quantity ordered is 2308.

Considering that material costs are incremented from k_1 to k_{10} as shown in Figure 1, it is remarkable that there is priority to use cheaper materials from each family although they are ordered beforehand due to suppliers availability (see Figures 2a-2c). With the aim of clearing this, material ordered in each time period is compared with material use due to product demand. As an example, in Figure 3 the solution for family f_1 is presented. It shows that the biggest quantities ordered are placed in the first periods while consumption is actually in the third one. The main reason is that material availability of

the least expensive materials (k_1 and k_2) is limited in the third period and additionally, purchase costs for all materials are incremented in peak periods too (Figure 1). Material k_1 is the most ordered, which is the cheapest from family f_1. Material k_2 is requested in minor quantities and material k_3 is not ordered.

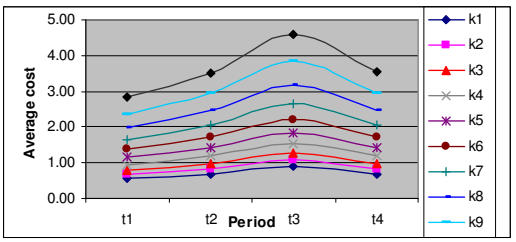


Figure 1: Material average costs

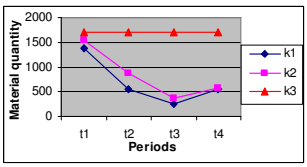


Figure 2 a: Suppliers availability for f_1.

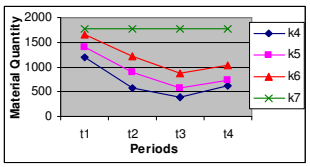


Figure 2 b: Suppliers availability for f_2.

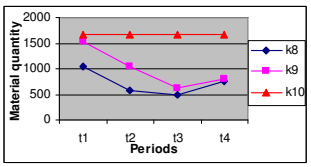


Figure 2 c: Suppliers availability for f_3.

As mentioned before, buying materials in advance causes an increase in stock and financial costs. However, there are two main reasons for using this strategy. In some cases, suppliers cannot provide enough material to satisfy the demand in peak periods. In others, although suppliers can provide certain material from the family, its cost is too expensive in high season thus rising production costs.

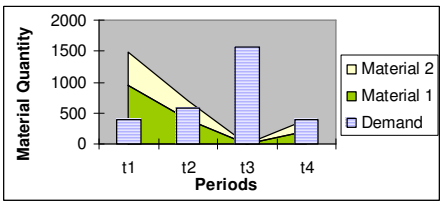


Figure 3: Demand and material purchase for family f_1.

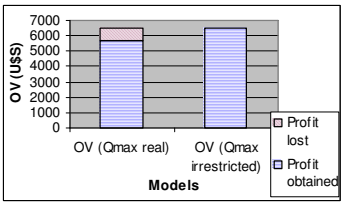


Figure 4: Objective value comparison.

Additionally, there are some other factors contributing to bigger ordered quantity: a) there is a fixed cost per contract signed; b) quantity flexibility contracts promote different discounts according to the amount purchased. In order to analyze the impact of these decisions in the objective value, financial costs of the material bought in advance should be determined. Two different strategies are presented to estimate the value of the financial costs. In the first place, it is possible to change the supplier availability considering unrestricted capacities for all periods.

Table 3. Material quantity decision for the fictitious case.

q_{jkt}	j_1k_1	j_1k_4	j_1k_8	j_2k_1
t_1	***170***	451		
t_2	700	700	***85***	
t_3	700	700	175	618
t_4	400	295	120	

Qmin active (Disjunction (7))

Table 3 shows that the solution obtained presents a few number of materials ordered before they were required. They are selected not because materials are insufficient in peak demand periods but because material costs increase in this period. The difference

between the objective value in the real case (5686.51) and the objective value in this fictitious one (6539.15) gives an approximate idea of the profit lost due to the financial costs paid caused by the supplier restricted capacity (Figure 4).

From another point of view, the quantity ordered in advance in each period according to the corresponding demand is estimated. The present discounted cost of the material acquired in periods before its actual consumption is calculated for each material family handled, representing the capital invested in stock. These two perspectives show how the solution is conditioned by supplier availability changes.

Table 4. Material purchased in advance and the corresponding actualized costs for family f_1.

Family f_1		Material k_1						Material k_2		Initial stock	Total quantity purchased	Quantity in advance	Costs in advance
		Supplier j_1		Supplier j_2		Supplier j_3		Supplier j_1					
Period	Demand	Quantity purchased	Cost	Quantity purchased	Cost	Quantity purchased	Cost	Quantity purchased	Cost				
t_1	383	500	0.5	450	0.55	0	0.6	538	0.6	350	1488	1455	803.8
t_2	580	200	0.6	200	0.7	0	0.75	300	0.72	-	700	120	128
t_3	1575	0	0.85	0	0.9	0	0.92	0	1.02	-	0	0	0
t_4	400	0	0.6	0	0.76	200	0.7	200	0.72	-	400	0	0
i = 8%											**Actualized total cost**		**854**

Table 4 shows the demand for family f_1 and the quantity bought from each supplier in the horizon planning. It also shows the cost of each material according to the supplier and period. The total quantity purchased is the summation of all materials purchased from all suppliers while the quantity in advance is calculated subtracting the demand to this quantity. The cost in advance is estimated using the materials purchased in advance and multiplying it by the corresponding material cost.

4. Conclusion

The problem tackled in this paper shows how several important decisions can be handled together so as to optimize the company global operation. In this sense, inventory optimization is considered jointly with purchase planning decisions. A new disjunctive model is proposed for solving the integration of inventory management and material purchase including also the selection of providers, specific materials and contract policies. Seasonal demand and material availability make this problem even more interesting since purchase and inventory decisions play an important role in the company profits.

With the intention of showing the problem solution, results are presented and analyzed in the previous section. This analysis shows that the quantity purchased is restricted by supplier availability in several occasions. In fact, there is a trend towards buying cheap materials in advance and stock them until they are needed. For that reason, financial costs were estimated using two different techniques showing the impact of these decisions on the company profits. The case study presented and the results obtained strengthen the relevance of the approach proposed.

References

D.J Bowersox, D.J. Closs and M.B. Cooper, 2007, Supply chain logistics management, Mc Graw-Hill, Second Edition, Singapore.

I. E. Grossmann, A. Meeraus and A. R. Vecchietti, 2005, Computational experience with LogMIP solving linear and nonlinear disjunctive programming problems, *Proceedings of the 9th Informs Computer Society Conference*, Annapolis, MA, USA.

M. Park, S. Park, F.D. Mele and I.E. Grossmann, 2006, Modeling of purchase and sales contracts in supply chain optimization, Industrial Engineering and Chemistry Research, 45, p. 5013–5026.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)

© 2009 Elsevier B.V. All rights reserved.

Adaptive Random Search: A promising method for determining the stability of mixtures

Nauro S. Junior[a], Nilo S. M. Cardozo[a] , Argimiro R. Secchi[b] and Keiko Wada[a]

[a]*DEQUI/UFRGS - Department of Chemical Engineering - Universidade Federal do Rio Grande do Sul. Rua Eng. Luis Englert, s/n., Centro. CEP 90040-040, Porto Alegre - RS, Brazil - E-mail: nauroeq@yahoo.com.br, nilo@enq.ufrgs.br*
[b]*COPPE/PEQ/LMSCP - UFRJ, Av. Horácio Macedo, 2030 - Centro de Tecnologia. CEP 21941-914, Rio de Janeiro - RJ, Brazil. E-mail: arge@peq.coppe.ufrj.br*

Abstract

Mixtures stability analysis is a calculation procedure to determine whether a mixture will exist as a homogeneous phase or as multiple phases in equilibrium, for an established condition. The mathematical problems related to this analysis generally represent a great challenge with respect to the use of optimization methods, due to the occurrence of great number of local minima. The Adaptive Random Search (ARS) method is a stochastic global optimization method that has revealed superior to other stochastic methods such as genetic algorithm and the Particle Swarm Optimization. In the present work the hybridization of the ARS method with an algorithm for the solution of nonlinear equation systems is proposed, aiming to achieve the following characteristics: i) capability of covering the whole search space without being attracted to local minima, ii) small dependence on the initial guess, and iii) low computational cost. The methodology for stability analysis using hybrid ARS was implemented for three different equations of state - Soave-Redlich-Kwong (SRK), Peng-Robinson (PR) and PC-SAFT (Perturbed Chain – Statistical Associating Fluid Theory) - and tested on some mixtures for which data are available in the literature. The results of the original and hybrid versions of the ARS method were compared to those obtained with deterministic methods: a quasi-Newton method with BFGS approximation of the Hessian and with the DIRECT, which is another deterministic global search method that has been used in stability analysis problems. The hybrid ARS method allowed determining the correct thermodynamic condition (i.e., stable or unstable) of all tested mixtures. Additionally, in the cases where multiple local minima were present, this method has provided more than one solution in a single run, which is very useful for the solution of the phase equilibrium problem. In this sense, the composition values obtained by solving the stability analysis with the hybrid ARS were used to initialize the algorithm that solves Rachford-Rice equation (flash problem). Furthermore, the hybrid ARS method was able to validate all the solutions found when solving flash problem, for both stable and unstable mixtures.

Keywords: stability analysis, global optimization, phase equilibrium.

1. Introduction

When solving a phase equilibrium problem using equations of state, the most important practical aspects is to know if the tool performs stability analysis and, if so, how are the performance and reliability of the stability analysis method.

The most used stability analysis formulations were proposed by Michelsen (1982) and are based on the Gibbs tangent plane criterion. Implementation of Michelsen's stability analysis formulations requires the use of optimization methods or methods to solve nonlinear algebraic equations systems. The main difficulty found in applying these methods to this specific problem is related to the possible existence of multiple solutions besides the trivial one or the occurrence of several local minimum. Therefore, the method used to solve stability analysis problem should be able to scan the solution space effectively.

Due to complexity of the numerical problems involved in stability analysis and although several optimization methods have been tested for this task, only a few have

presented satisfactory performance. Interval Newton method (Tan and Radoz, 2002), the tunneling global search (Nichita et al., 2008) and DIRECT (Saber and Shaw, 2008) can be mentioned as examples of optimization methods successfully applied to stability analysis. Nevertheless, the development of methods to solve stability analysis problems is still a subject of great interest, both from the scientific and industrial points of view. The method should be able to meet the following features: i) capability of covering the whole search space without being attracted to local minima, ii) small dependence on the initial guess, and iii) low computational cost.

In this context, a method that has great potential for use in stability analysis problems is the Adaptive Random Search (ARS) method of Secchi and Perlingeiro (1989). In the present work the ARS method was modified to solve mixture stability problems, taking advantage of its ability to scan the search space and determine the maximum number of possible solutions, and introducing a local search algorithm for refinement of the generated points and minimizing with this the probability of failures when the determination of unstable or stable mixtures.

2. Methodology

The methodology for stability analysis by the modified ARS method was implemented for three different equations of state - Soave-Redlich-Kwong (SRK) (Soave, 1972), Peng-Robinson (PR) (Peng and Robinson, 1976) and PC-SAFT (Perturbed Chain – Statistical Associating Fluid Theory) (Gross and Sadowski, 2001) – and tested using mixtures data found in literature. The performance of modified ARS method for the chosen mixtures stability problems was compared with the original ARS, a quasi-Newton method (FMINUNC with BFGS, Coleman *et al*., 1999), and DIRECT (Saber and Shawn, 2008), which is a global optimization method and make search inside of the specific space.

2.1 Adaptive Random Search

The ARS method is a global optimization method, based on a memorized random search, with asymmetric distribution and systematic reduction of a hyper-ellipse of samples (Secchi and Perlingeiro, 1989). In each stage of the method random points are generated inside of the hyper-ellipse using the following expression:

$$x_{i,j}^{k+1} = x\_ot_i^k + \frac{R_i}{\delta}\left[B_i\left(A_i\Psi_{i,j} - 1\right)\right]^{\delta} \qquad i = 1, NCOMP\ ;\ j = 1, NSAMPLE \qquad (1)$$

where *NCOMP* is the number of variables (in stability analysis problems is equal to the number of components), *NSAMPLE* is the number of generated random points for the search, in the step k+1, inside of hyper-ellipse, $x_{i,j}^{k+1}$ are the search points, $\underline{x}\_ot_i^k$ are the coordinates of the best point in k-th step of the algorithm (or the starting point for k = 1), $\Psi_{i,j}$ are random numbers between zero and one, R_i are the hyper-ellipse axes sizes, δ is a parameter that attributes the samples distribution form, whereas for δ = 1, the distribution is uniform. A_i is the distribution asymmetry vector (A = 2 means symmetrical distribution) and B_i is a vector that gives the direction of the asymmetry. Designating *F_ot* as the value of objective function at the point $\underline{x}\_ot_i^k$, the ARS method is executed in the following steps:

1 – Evaluate the objective function in the initial point and store this value in *F_ot* (this variable is responsible for storing, always, the best value of objective function);

2 – Generate a set of random values, inside of the hyper-ellipse space, with respective evaluation of the objective function, and store these values. This step is repeated *NSAMPLE* times, generating a vector of objective functions (F_t_j).

3 – Compare the objective function values (F_t_j) calculated in step 2 with *F_ot*. If F_t_j has at least one value better than *F_ot*, $x_{i,j}^{k+1}$ is considered promising for the global minimum search. Otherwise set $x_{i,j}^{k+1}$ is discarded and returns to step 2. If after a specified number of attempts no set of coordinates are found with the objective function value better than *F_ot* the search is finished;

4 – Since $x_{i,j}^{k+1}$ has been considered promising, *F_ot* receive the best value of F_t_j and $\underline{x}\_ot_i^k$ get their coordinates;

5 - Check whether the old value of $\underline{x}\_ot_i^k$ must be discarded or not. This is done by evaluating the distance (Equation 2) between the coordinates of the old and the new points, comparing this distance with a metric *M* given by Equation 3.

$$Dm = \|x\_ot^k - x\_ot^{k+1}\| \quad (2)$$

$$M = 0.2c[\sum_i (R)^2]^{\frac{1}{2}} \quad i = 1,2,\dots,n \quad (3)$$

where *Dm* is the distance between coordinates of iterations *k* and *k* +1 and c = 0.5 in this step. If *Dm* < *M*, x_ot^k is discarded, otherwise this point is stored too.

6 – Verify if any other point $x_{i,j}^{k+1}$matrix coordinates. is located in promising regions of the search space. However, in this step only the distances between these points are evaluated using Equation 4.

$$Dp = \left[\sum_i \left(x_{i,j}^{k+1} - x_{i,l}^{k+1}\right)^2\right]^{\frac{1}{2}}, \quad for\ j \neq l \quad (4)$$

where *Dp* is the distance between the points randomly generated. The criterion to accept a set of points is also done using the metric *M,* Eq. 3, but at this step *c* = 1.

7 - Reduce the hyper-ellipse radius by increasing the parameter δ (δ = 1, 3, 5,... when δ = 1 means uniform distribution) and return to step 2. These steps are repeated until the hyper-ellipse radius becomes less than a given tolerance, when an optimum is located. If there are still points stored, a new starting point is recovered from memory and the steps 2 to 7 are repeated, restarting with a uniform distribution.

2.2 Mixtures Stability Analysis

The procedure for a mixture stability test presented here was developed above the following consideration: if the original mixture is stable, there is no composition of a new additional phase that leads to decrease the global mixture Gibbs energy. Based on that, many formulations were developed and the most used in literature were presented in Michelsen (1982). In this work the following formulations presented by Michelsen (1982) were used:

$$lnY_i + ln\hat{\phi}_i(\boldsymbol{y}) - h_i = 0 \quad (5)$$

$$TPD(Y_i) = 1 + \sum_i Y_i\left(lnY_i + ln\hat{\phi}_i(\boldsymbol{y}) - h_i - 1\right) \quad (6)$$

where $h_i = \ln(z_i) + ln\hat{\phi}_i(\boldsymbol{z})$, $\hat{\phi}_i$ is the fugacity coefficient of component *i*, z_i are the molar fractions of the mixture for which the stability are being tested, $\boldsymbol{z}$ is the vector that contains z_i values, Y_i is an independent variable that has as unique restriction that their values must be greater than zero, $\boldsymbol{y}$ is the vector containing y_i values (molar fraction of trial phase), which are obtained by the following relationship:

$$y_i = \frac{Y_i}{\sum_i Y_i} \quad (7)$$

These equations were obtained through algebraic manipulation of the free Gibbs energy expressions, in order to obtain a final mathematical formulation which can deal more efficiently with the infinitesimal amounts of the trial phase. Complete description of the algebraic development can be found in Michelsen (1982). The values of Y_i that minimize Equation 6 are also the roots of Equation 5. In order to determine the mixture stability, it is necessary to solve just one of these equations, however, both were used in this work because the modification in ARS method was made by inclusion of an algorithm for solving nonlinear algebraic systems. The mixture instability condition is that the global minimum of Equation 6 has TPD (Y_i) < 0. The minimums of Equation 6 (or roots in Equation 5) are called stationary points and these values are, generally, found near the composition of phases in equilibrium, when a mixture is unstable.

2.3 Modifications in ARS method for determining mixtures stability

The goal of the proposed modification is to use the ability of the ARS method in generate promising values in search space together with the high convergence rate of the Newton method. This is done by using the stored coordinates as promising (determined using Equation 6), eliminating the stage of solution refinement in the ARS method (reduction of the hyper-ellipse radius by increasing the parameter δ) and using these values to solve the problem represented by Equation 5 using the Newton method.

As the only restriction for the decision variables is the positiveness, it was also proposed the following modification to generate random points avoiding negative values:

$$x_{i,j}^{k+1} = x\_ot_i^k \times exp\left\{\frac{R_i}{x_{ot_i}^k \delta}\left[B_i\left(A_i \Psi_{\mathrm{i,j}} - 1\right)\right]^{\delta}\right\} \quad (8)$$

Bringing the specific variables from stability problem to optimization problem, $x_{i,j}^{k+1}$ represents $Y_{i,j}^{k+1}$ which by Equation 7 is directly related with the molar fractions. As the molar fractions exist within a defined range it is more advantageous to use these values (y_i) instead of those (Y_i) to calculate the metric M, becoming a more robust criterion to evaluate the coordinates generated at each iteration k.

2.4 Performance indicators

When using stochastic methods, the number of objective function evaluations and the obtained solution can be different for each run the method is applied. Thus, for comparative analysis of different stochastic methods, efficiency and robustness indicators were used, as defined below, respectively.

1 – Average number of objective function evaluations and its standard deviation;

2 – Robustness (η): the robustness indicates the percentage of success to determine the global minimum of the problem.

$$\eta = 100\,\frac{n}{N^*} \quad (9)$$

where n is the number of times that the method correctly solved the problem, finding the global minimum, and N^* is the number of tests for each problem.

2.5 Solving the Flash problem

To solve multiphase and multicomponent flash problem from Rachford-Rice equation, routines have been implemented using the successive substitutions method, as shown by Leibovici and Neoschil (1995). An important point in this topic is the initial guess used in numerical procedures; as mentioned by Heidemann and Michelsen (1995), this method can be very unstable, especially for very asymmetric mixtures such as water and oil and polymer-solvent mixtures. In this work, it was used the coordinates of the minimum found in the stability problem to start the numerical methods. As the mixtures treated in this work are cases where the stability problem presents many minima, the implemented algorithm with the numerical procedures for resolution of the flash problem was started with different combinations of these coordinates.

3. Results

The optimization methods discussed previously were tested with a series of mixtures stability problems. Due to the paper length limitations, only two cases are discussed here:

Mixture 1: Methane-hydrogen sulfide at 191.15 K and 40.53 bar with 50% in molar fraction (Michelsen 1982; Nichita *et* al., 2008; among others). This problem has been formulated with PC-SAFT equations of state, with a global minimum of -0.0779 whose coordinates in molar fraction are [0.9809 0.0191] and with SRK whose minimum is -0.0766 with coordinates in molar fraction of [0.9791 0.0209]. The values of *kij* parameter were obtained, for the PC-SAFT model, in Nichita *et* al. (2008) and for SRK in Harding and Floudas (1999).

Mixture 2: Ethanol, acetic acid, ethyl acetate, water and carbon dioxide at 333.15 K and 58 bar, with composition in molar fraction of [0.2027 0.2026 0.1611 0.1617 0.2719] (Solórzano et al., 2004). This problem was formulated using PR equation of state for which the global minimum is -2.5783 with coordinates in molar fraction of [0.39110 6.7 × 10-05 7.74 × 10-07 0.60767 0.00115]. The values of *kij* parameters were obtained in Solórzano *et* al. (2004).

In both cases many local minima can be observed, this characteristic makes these problems a challenge to optimization methods. Table 1 presents the performance of the original and hybrid ARS methods for 300 tests with each problem (examples 1 and 2).

Table 1: Performance of ARS and hybrid ARS (HARS).

Case	Example 1				Example 2	
Model	SRK		PC-SAFT		PR	
Method	ARS	HARS	ARS	HARS	ARS	HARS
Robustness	52.33	99.33	32.00	95.67	97.33	97.00
Evaluation avarege number	397.88	218.14	380.88	162.94	3405.30	1599.19
Standard deviation	42.31	31.35	38.97	37.51	545.33	307.00
Minimal number of evaluations	325	151	319	99	2296	1015

The number of evaluations computed for HARS method is number of evaluations for generating promising values plus the number of iterations to solve the equations system represented by Eq. 5 with all generated promising points. As shown in Table 1, the hybridization of the ARS method decreased the required evaluations number for solve stability problem in both examples. Moreover, except for example 2, in which both methods had gotten good robustness, robustness of the HARS method for example 1was very superior to ARS for both equations of state. An important thing is that in all cases and tests presented, the initial guesses were always the trivial solution coordinates. This aspect shows that both, ARS and hybrid ARS, are not influenced by initial guesses values. Another important point is that when the mixture is stable, the trivial solution is the global minimum of the problem ($TPD = 0$).

Regarding the FMINUNC method, this solved the stability problem of the Mixture 1 with 112 objective function evaluations. However, this method has led to a local minimum with a value of -0.06598. The DIRECT method, even found the global minimum, needed 1289 evaluations to determine it. For Mixture 2, the FMINUNC method again established a local minimum, whose value is -0.04567 with 156 objective function evaluations. The DIRECT method needed 1431 objective function evaluations and found the global minimum.

Regarding the phase equilibrium problem, both mixtures are cases where exists, beyond the correct solution, another numerical solution that satisfies the equilibrium criteria (equality of the chemical potentials in all phases and mass balance restrictions), however, in terms of the Gibbs free energy, this solution does not represent a global minimum. Because of this, the stability analysis is used to validate the solution, showing the local minimum solution is unstable and the global minimum is stable. Tables 2 and 3 present these solutions for the two mixtures.

Table 2: Solutions of the phase equilibrium for Mixture 1.

Equation of State	Equilibrium type	Molar Fraction		Z = PV/RT	Condition
SRK	LL	0,91097	0,08903	0,1627	unstable
		0,12518	0,87482	0,0939	
SRK	LV	0,97838	0,02162	0,5502	stable
		0,12441	0,87559	0,0939	

Table 3: Solutions of the phase equilibrium for Mixture 2.

EOS	Equilibrium	Molar Fraction					Z = PV/RT	Condition
		Ethanol	Ac. Ac	Ethyl Ac.	Water	CO_2		
PR	LL	0,4955	0.00116	$1{,}18\times10^{-5}$	0,49465	0,00869	0,1458	unstable
		0,0631	0,29863	0,23788	0,00300	0,39735	0,0855	
PR	LLV	0,4963	0,00126	$1{,}37\times10^{-05}$	0,49442	0,00796	0.0857	stable
		0,0661	0,31538	0,25107	0,00323	0,36419	0.1482	
		0,0024	0,00141	0,00471	$4{,}29\times10^{-05}$	0,99141	0.7442	

The stability of the solutions in the flash problem was tested using hybrid ARS method, which correctly indicated stable and unstable solutions. For both mixtures, the

flash procedure was initialized with different values for each phase compositions related to some solution of the stability problem.

4. Conclusions

In this work a modification for the ARS method was proposed, named as hybrid ARS. The performance of the hybrid ARS method in stability analysis problems was compared with the conventional ARS method, a quasi-Newton, and DIRECT. Hybrid ARS method showed to be a trustworthy method with great applicability potential for this kind of problem. This method showed high robustness in all analyzed examples, which can be characterized as being of high complexity. Moreover, the hybridization of the ARS method significantly decreased the average number of objective function evaluations needed to find global minimum, being more efficient, also, than DIRECT method. Although FMINUNC establish a solution with few evaluations of objective function, this type of method is not reliable due to its dependence on initial guesses. The ARS and hybrid ARS method did not show this feature so that even the trivial solution can be used as initial guesses of the variables in the stability analysis procedure. Another important point is that these characteristics of reliability and performance delivered by the hybrid ARS method were observed for the three types of equation of state: Soave-Redlich-Kwong (SRK), Peng-Robinson (PR) e PC-SAFT (Perturbed Chain – Statistical Associating Fluid Theory).

References

Gozalpour F., Danesh A., Todd A.C. e Tohidi B., "Vapour–liquid equilibrium compositional data for a model fluid at elevated temperatures and pressures", *Fluid Phase Equilibria* **208**, 303–313, 2003.

Gross J. e Sadowski G., "Perturbed-Chain SAFT: An Equation of State Based on a Perturbation Theory for Chain Molecules", *Ind. Eng. Chem. Res.* **40**, 1244-1260, 2001.

Heidemam R.A. e Michelsen M.L., "Instability of Successive Substitution", Ind. Eng. Chem. Res. **34**, 958-966, 1995.

Leibovici C. F. e Neoschil J., "A solution of Rachford-Rice equations for multiphase systems", *Fluid Phase Equilibria*, **112,** 217-221, 1995**.**

Leibovici C. F. e Nichita D. V., "A new look at multiphase Rachford-Rice equation for negative flashes", *Fluid Phase Equilibria*, **267**, 127-132, 2008**.**

Coleman, T., Branch M.A., Grace A., "Optimization Toolbox for use with MATLAB", The Mathworks, Inc., 1999.

Michelsen M.L., "The Isothermal Flash Problem. Part I. Stability", *Fluid Phase Equilibria* **9**, 1-19, 1982.

Nichita D. V., Sánchez F. G., Gómez S., "Phase stability analysis using the PC-SAFT equation of state and the tunneling global optimization method", *Chemical Engineering Journal* **140**, 509-520, 2008.

Peng, D. Y. and Robinson, D. B. "A New Two-Constant Equation of State" Ind. Eng. Chem. Fundam.,**15,** 59-54, 1976.

Saber N. e Shaw J.M., "Rapid and robust phase behaviour stability analysis using global optimization", *Fluid Phase Equilibria* **264**, 137–146, 2008.

Secchi and Perlingeiro, "Otimização: Busca Aleatória Adaptativa". Anais do XII Congresso Nacional de Matemática Aplicada e Computacional (XII CNMAC), São José do Rio Preto, SP, 49-52, 1989.

Soave, G., ''Equilibrium Constants from a Modified Redlich-Kwong Equation of State,'' Chem. Eng. Sci., **27**, 1197, 1972.

Solórzano G. I. B., Brennecke J. F., Stadtherr M. A., "Validated computing approach for high-pressure chemical and multiphase equilibrium", *Fluid Phase Equilibria* **219**, 245–255, 2004.

Tan S.P. and Radosz M., "Gibbs Topological Analysis for Constructing Phase Diagrams of Binary and Ternary Mixtures", *Ind. Eng. Chem. Res.* **41**, 5848-5855, 2002.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Monitoring of Vinyl Chloride Suspension Polymerization Using NIRS. 1. Prediction of Morphological Properties

João Miguel de Faria Jr.,[a,b] Fabricio Machado,[b,‡] Enrique Luis Lima,[b] José Carlos Pinto[b]

[a]*Braskem S.A., Rua Hidrogênio, 3342, Pólo Petroquímico, CEP: 42810-000, Camaçari, BA, Brazil*
[b]*Programa de Engenharia Química / COPPE, Universidade Federal do Rio de Janeiro, Cidade Universitária, CP 68502, Rio de Janeiro, 21945-970, RJ, Brazil*

Abstract

The monitoring of polymer properties is frequently performed through sampling and off-line measurement of quality parameters in lab scale. Because of the long time required to perform off-line analyses, off-line data are generally inadequate for control purposes. As impurities cannot be detected in-line and can exert significant influence on the polymerization kinetics, modifications of some important quality parameters (such as the morphological characteristics of PVC particles) can only be detected at the end of the polymerization, causing major production losses. In the present work it is shown for the first time that it is possible to monitor some of the morphological properties of PVC (such as the bulk density, BD, the cold plasticizer absorption, CPA, and the average particle diameter D_p) *in situ* and in real time with the help of near infrared spectroscopy (NIRS) in suspension polymerization processes.

Keywords: Near Infrared Spectroscopy (NIRS), Suspension Polymerization, Poly(vinyl chloride) (PVC), Polymer Particle Morphology, Real Time Monitoring.

1. Introduction

Development of tailored polymer resins is required for several applications in different fields. In order to achieve the production of a polymer resin with specified end-use properties, the availability of efficient control techniques is of fundamental importance. A gap between the polymerization process and the control technique must be bridged with accurate and robust instrumentation for on-line and in-line monitoring and calculation. The lack of instruments that are able to measure and monitor the quality of the polymer resin has been recognized by many authors as the most important problem in the field of polymerization reactor control. The complex nature of typical polymerization systems is one of the main difficulties for development of adequate monitoring instruments. Despite the huge efforts invested for development of reliable and robust monitoring techniques in recent years, important polymer properties cannot be measured on-line. On the other hand, off-line measuring of polymer properties can be very laborious and take a lot of time. Besides, there is a difficult task of relating molecular properties (which can usually be evaluated remotely in the laboratory, such as

‡ Current Address: Instituto de Química, Universidade de Brasília, *Campus* Universitário Darcy Ribeiro, CP 04478, 70910-900, Brasília, DF, Brazil, E-mail: fmachado@unb.br

the average molecular weight, average particle size, and copolymer composition, among others) and final end-use properties (such as brightness, thermal and mechanical resistance, stiffness, etc.), whose quantification is much more complex and usually determined for final products only (Santos et al., 2005).

Sampling devices intended for on-line applications are often sophisticated and complicated, requiring careful design and installation procedures, which very frequently cannot be performed in real industrial sites. Additionally, several available monitoring instruments are so expensive that installation and maintenance costs can be prohibitive for most industrial facilities. According to Santos et al. (2005), it is not surprising to observe that there is an enormous gap between the control applications developed in the research laboratory and the actual control applications implemented industrially.

A major breakthrough in the area of on-line polymer reaction and quality monitoring was achieved when conventional sensing techniques were coupled to fiber optics, making *in situ* measurements and remote analysis possible for a large number of applications (Kammona et al., 1999). Among them, monitoring techniques based on the NIRS are very useful for a number of practical applications in the field of polymer reaction engineering. When quantitative analyses are necessary, NIRS monitoring technique provides process information, which is used to determine the calibration model that after validation with actual process data is employed to infer the final polymer properties. In this context, it is shown for the first time that it is possible to monitor some of the morphological properties of PVC (such as the bulk density, BD, the cold plasticizer absorption, CPA, and the average particle diameter D_p) *in situ* and in real time with the help of near infrared spectroscopy (NIRS) in suspension polymerization processes.

2. Experimental Procedure

2.1. Polymerization Reactions

The experimental apparatus consists of stainless-steel autoclaves with volumes of 0.001 and 0.1 m^3 and maximum working pressure and temperature of 20 bar and 250ºC, respectively. A thermocryostatic bath was used to control the reaction temperature. The agitation speed could be varied in the range of 0 to 2400 rpm and a Pt 100 thermocouple probe was used to monitor and control the reactor temperature.

NIR spectra were collected with the help of a NIRS-6500 equipment from NIRSystems Inc, configured to scan the spectral region ranging from 400 to 2500 nm, in reflectance and interactance mode. The instrument was connected to the bench scale reactor by using a bundle of optic fibers terminated with a 1-in diameter probe. The NSAS software (*Near-Infrared Spectral Analysis Software*) was used for data acquisition, spectral mathematical treatments, and partial least square regression. The acquisition interval of NIR spectra was adjusted to 300 seconds.

The main process variables were varied and coded based on a three-level factorial design, as illustrated in Table 1. The polymerization reactions were carried out in accordance with a factorial Taguchi planning, as shown in Table 2. A standard recipe consisting of poly(vinyl alcohol), used as a suspending agent, initiator, water, monomer and additives was charged into the reactor to perform the polymerization reactions. The detailed recipe will not be presented for confidential reasons. The agitation speed was selected as one of the design variables because it modifies the mechanism of break-up and coalescence of VCM droplets, which changes the average particle size and the PVC morphology (Jahanzad et al., 2005; Johnson, 1980; Lyoo et al., 2005 and Vivaldo-Lima et al., 1999). The temperature was selected to control the resin K-value (an indirect

measure of the average molecular weight and the processability of the PVC resins) (Burgess, 1982). The initiator amount (A) was selected as a design variable because it can be used to control the rate of reaction and the rate of heat production. The suspending agents (B and C) act on the suspension stability and on the particle size distribution (Burgess, 1982; Alexopoulos et al., 2007; Kichatov et al., 2003; Saeki and Emura 2002). At each particular experimental condition spectra were collected in the spectral regions of visible and near infrared in order to observe the effect of the process variables on the final properties of PVC resin.

Table 1 – Three-Level Variable Values

Variable	Unity	Level 1	Level 2	Level 3
X_1 - Speed ω	rpm	X	1.58X	2.16X
X_2 - Temperature	°C	Y	1.05Y	1.10Y
X_3 - VCM	g	Z	1.25Z	1.50Z
X_4 – amount of A	phm	K	1.06K	1.12K
X_5 – amount of B	phm	L	1.11L	1.22L
X_6 – amount of C	phm	M	2.00M	3.00M

Table 2 – Factorial Taguchi Planning

Experiment	Speed	Temperature	VCM	A	B	C
	X_1	X_2	X_3	X_4	X_5	X_6
1	-1	-1	+1	+1	+1	+1
2	-1	-1	+1	-1	-1	+1
3	-1	+1	+1	+1	-1	-1
4	-1	+1	+1	-1	+1	-1
5	+1	-1	-1	+1	+1	+1
6	+1	-1	-1	-1	-1	+1
7	-1	+1	-1	+1	+1	+1
8	-1	+1	-1	-1	-1	+1
9	+1	-1	+1	-1	+1	-1
10	+1	+1	-1	+1	-1	-1
11	+1	+1	-1	-1	+1	-1
12	+1	+1	+1	+1	+1	+1
13	+1	+1	+1	-1	-1	+1
14	0	0	0	0	0	0
15	-1	-1	-1	-1	+1	-1
16	-1	-1	-1	+1	-1	-1
17	+1	-1	+1	+1	-1	-1

2.2. Results and Discussion

2.2.1. Polymerization Analysis

The spectra presented in Figure 1 indicate spectral changes during the polymerization in the near infrared region between 1400 and 2100 nm. According to Santos et al. (1998, 2000), the region between 1600-1700 nm is sensitive to modification of the average size of the suspended droplets. As batches were performed at very different conditions, as shown in Table 1 and 2, this can possibly indicate that the PVC particle morphology can be changed along the reaction course; however, these changes could also be generated by others factors, instead of particle morphology. Thus, the influence of morphology on the NIR spectrum should be analyzed with the help of mathematical and statistical

techniques, such as partial least squares (PLS). The development of most NIR-based monitoring and control applications in suspension polymerization processes only becomes possible because of reflection, refraction and scattering of light on the surface of the suspended particles. The large intensity of scattering (usually considered as a disadvantage in most applications), when compared to absorption, is fundamental for determination of the morphological features of polymer beads in these systems. In the particular case of MVC suspension polymerization, it is important to notice that the average diameter of the pores of the final PVC particles has the same magnitude of the wavelengths in the NIR region (ranged from 1 to 2 μm). Thus, light scattering can be strongly influenced by the particle structure.

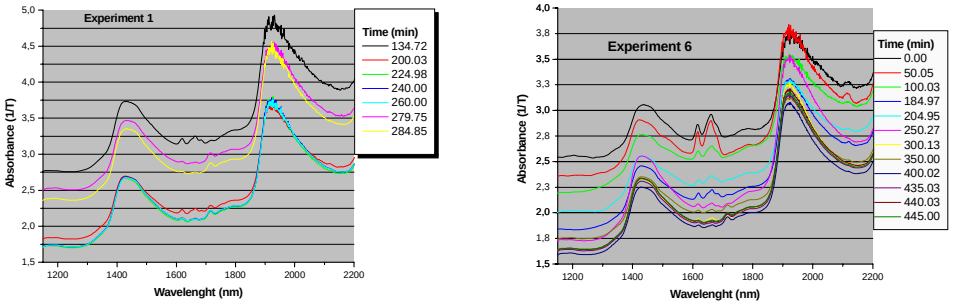


Figure 1 - Spectral Evolution along the Polymerization Time.

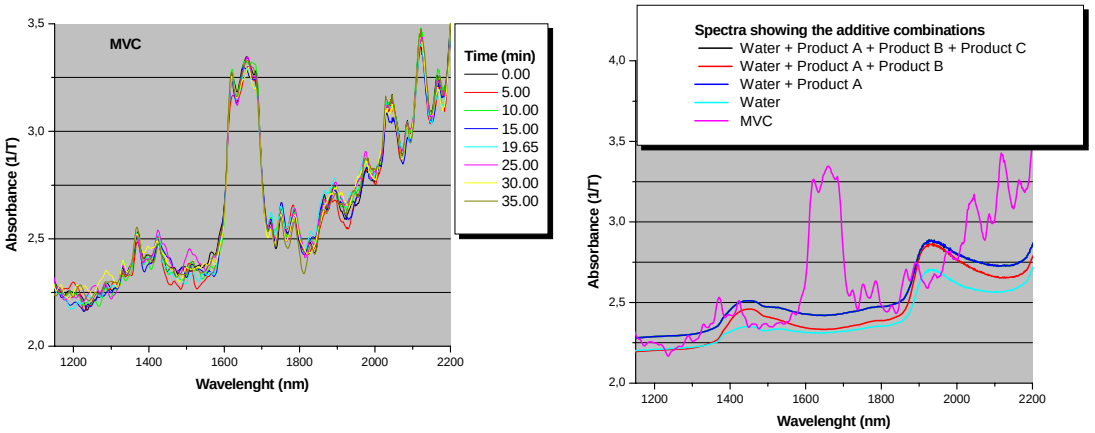


Figure 2 - Spectra of VCM and Additives used in the polymerization.

As one can see in Figure 2, the spectral absorbance presents a strong peak in the 1600 nm region, showing that this region can be useful for quantitative analyses. Based on the observations of the spectra obtained from different experimental conditions, it can be said that spectral variations are different in the whole experimental region, showing that NIRS is sensitive to the variations of the proposed experimental values. Assuming that the different experimental conditions could lead to PVC particles with different morphological characteristics, NIRS can be employed to monitor the morphological properties of PVC particles and could be used for control purposes at plant site.

2.2.1. Calibration Studies

Empirical calibration models were obtained based on the NIR spectra and PVC morphological properties obtained during actual runs. Polymer samples obtained at the end of the polymerization batches were used for determination of important resin

properties, such as BD, CPA and Dp. The standard PLS technique was used for model building. The calibration model for BD was built with eight factors, while models were built with ten factors for the others morphological parameters (CPA and Dp). Calibration was performed only with the final batch samples, because quality parameters BD, CPA and Dp can be measured only when the polymerization is finished. As shown in Figure 3, the calibration models reproduce very well the experimental data, which confirms the quality of predictions given by these models. (Models were validated with additional data, but results are not shown for lack of space.)

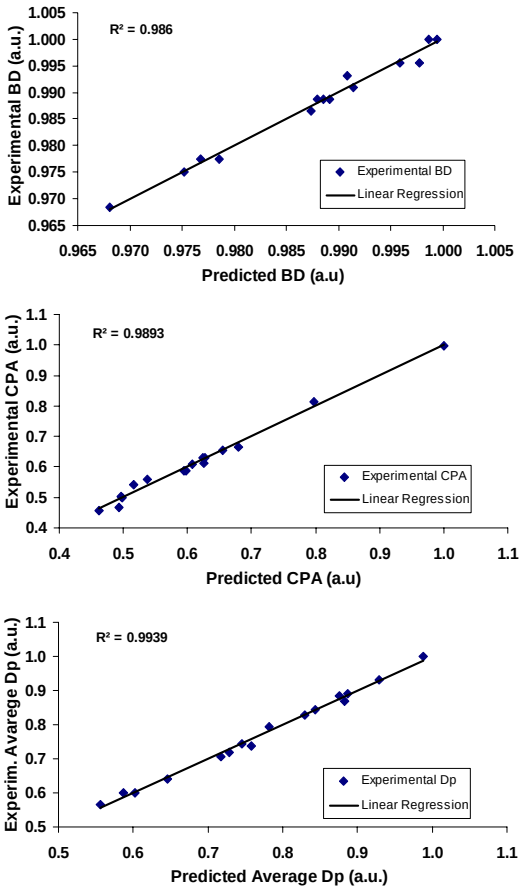


Figure 3 – Prediction of the PVC Morphological Properties. **(A)** Bulk Density; **(B)** Cold Plasticizer Absorption; **(C)** Average Particle Diameter.

The evolution of morphological properties of PVC along the batch is analyzed in the second part of this work (de Faria Jr. et al., 2009) and the reader should refer to this publication for more detailed information about dynamic trajectories. De Faria Jr. et al. (2009) showed that it is possible to follow the dynamic evolution of morphological properties (for example, bulk density and cold plasticizer absorption) in real time, in heterogeneous VCM processes, using NIR spectroscopy technique.

3. Conclusions

Several vinyl chloride polymerizations were performed in suspension mode to evaluate the NIR spectral response to different process variables. Obtained results show excellent

agreement between experimental data and predicted values given by the calibration models and indicate clearly that the NIRS technique can be used successfully to monitor properties related to PVC particle morphology. It is very important to emphasize that this is the first time in the open literature that NIR spectroscopy is used to predict morphological parameters of suspension PVC particles, such as BD, CPA and D_p. Calibration models reproduce available experimental data fairly well, showing that morphological parameters of PVC particle can be followed in-line during the polymerization process. Based on this, the in-line monitoring of these variables during the VCM polymerization with a NIR probe could be used as a fundamental piece during the implementation of advanced monitoring and control schemes, aiming at controlling the morphology of PVC resins.

Acknowledgements

The authors thank Braskem S.A. and FINEP for supplying technical resources required for this work.

References

A. F. Santos, E. L. Lima, J. C. Pinto, 1998, In-Line Evaluation of Average Particle Size in Styrene Suspension Polymerization Using Near-Infrared Spectroscopy, Journal of Applied Polymer Science 70, 1737-1745.

A. F. Santos, E. L. Lima, J. C. Pinto, 2000, Control and Design of Average Particle Size in Styrene Suspension Polymerization Using NIRS, Journal of Applied Polymer Science 77, 453-462.

A. F. Santos, M. K. Lenzi, F. Machado, J. C. Pinto, 2005, Monitoring and Control of Polymerization Reactors Using NIR Spectroscopy, Polymer-Plastics Technology and Engineering, 44, 1-61.

A. H. Alexopoulos, C. Kiparissides, 2007, On the Prediction of Internal Particle Morphology in Suspension Polymerization of Vinyl Chloride. Part I: The Effect of Primary Particle Size distribution, Chemical Engineering Science 62, 3970-3983.

B. V. Kichatov, A. M. Korshunov, P. Assorova, 2003, Particle Size Distribution of the Product of Suspension Polymerization, Theoretical Foundations of Chemical Engineering, 37, 3, 306-309.

E. Vivaldo-Lima, P. E. Wood, A. E. Hamielec, 1999, An Updated Review on Suspension Polymerization, Industrial & Engineering Chemistry Research, 36, 939-965.

F. Jahanzad, S. Sajjadi, B. W. Brooks, 2005, Comparative Study of Particle Size in Suspension Polymerization and Corresponding Monomer-Water Dispersion, Industrial & Engineering Chemistry Research, 44, 4112-4119.

G. R. Johnson, 1980, Effects of Agitation During VCM Suspension Polymerization, Journal of Vinyl Technology 2, 3, 138-140.

J. M. de Faria Jr., F. Machado, E. L. Lima, J. C. Pinto, 2009, Monitoring of Vinyl Chloride Suspension Polymerization Using NIRS. 2. Proposition of a Scheme to Control Morphological Properties of PVC, Computer-Aided Chemical Engineering (Accepted).

O. Kammona, E. G. Chatzi, C. Kiparissides, 1999, Recent Developments in Hardware Sensors for the On-Line Monitoring of Polymerization Reactions, Journal of Macromolecular Science - Reviews in Macromolecular Chemistry and and Physics, C39, 1, 57-134.

R. H. Burgess, 1982, Manufacture and Processing of PVC, London: Applied Science Publishers.

W. S. Lyoo, C. S. Park, K. H. Choi, J. W. Kwak, W. S. Yoon, S. K. Noh, 2005, Effect of Suspension Polymerization Conditions of Vinyl Pivalate on the Size and Its Distribution of Poly (Vinyl Pivalate) Microspheres, Polymer-Plastics Technology and Engineering 44, 475-487.

Y. Saeki, T. Emura, 2002, Technical progresses for PVC production, Progress in Polymer Science, 27, 2055-2131.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

How to Manage Complexity in Phase Equilibria Modeling? Application to the Bunsen Reaction

Mohamed K. Hadj-Kali,[a] Vincent Gerbaud,[a,*] Patrick Lovera,[b] Jean-Marc Borgard,[b] Pascal Floquet,[a] Xavier Joulia,[a] Philippe Carles [b]

[a] *Université de Toulouse INP- UPS- LGC, Laboratoire de Génie Chimique, 5 rue Paulin Talabot, 31106 Toulouse, France.* Vincent.Gerbaud@ensiacet.fr*
[b] *CEA, DEN, Physical Chemistry Department, F-91191 Gif sur Yvette, France.*

Abstract

The present work focuses on the thermodynamic modeling of the Bunsen section of the Iodine-Sulfur thermochemical cycle for hydrogen production on the basis of a combination of UNIQUAC activity coefficient and solvation models. The complexity of the material system, two immiscible electrolyte aqueous phases (H_2SO_4 rich and HI rich), is managed by defining an equivalent material system able to capture the main physical phenomena. Only the key species involved and their mutual interactions are taken into account, reducing the number of parameters to be estimated from a hundred to only 15. Results show a good agreement with published experimental data, with a better description of the impurities in the HI_x phase than in the sulfuric acid phase.

Keywords: H_2 production, IS thermo-cycle, Bunsen section, liquid–liquid equilibria.

1. The Sulfur–Iodine Thermochemical cycle

Hydrogen is considered to be one of the best energy carriers for the future. It can be produced from water by thermo-chemical cycles. The Iodine–Sulfur (IS) thermo-chemical water splitting cycle is one of the most promising (O'Keefe *et al.*, 1982), chosen among one hundred other possible cycles and expected to become a major source of massive hydrogen production combined with nuclear energy.
This cycle, depicted in figure 1, is divided into three sections, namely:

1. The Bunsen section (I), where sulfur dioxide reacts with excess water and iodine to produce two immiscible electrolyte aqueous phases: the upper phase containing mainly water and sulfuric acid (H_2SO_4 rich phase) and the heavy phase with water, hydroiodic acid and iodine (HI_x phase).
2. The sulfuric acid concentration and decomposition section (II),
3. The hydroiodic acid concentration and decomposition section (III).

In the two last sections, intermediate acids break down upon heating to release oxygen and hydrogen respectively while water, iodine and sulfur dioxide are recycled in the system.
After modeling the phase equilibria of the HI_x section (Hadj-Kali *et al.*, 2009), we focus in this paper on the complex liquid – liquid equilibrium occurring in the Bunsen section.

2. The Bunsen reaction

The starting point of the IS cycle is the Bunsen reaction:

$$2\ H_2O + SO_2 + I_2 \leftrightarrow H_2SO_4 + 2\ HI \qquad \text{(R1)} \qquad [1]$$

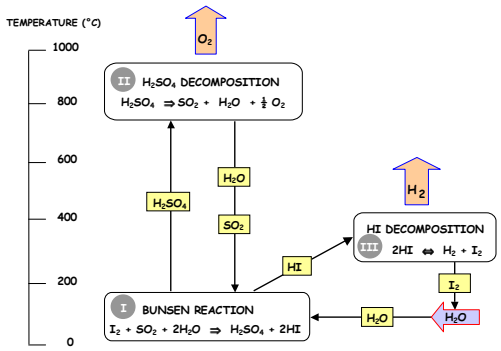


Figure 1. Sulfur – Iodine thermo-chemical cycle scheme.

This reaction is carried out at 120°C. With standard stoichiometry, its standard Gibbs free energy is positive: $\Delta G°(400\ K) = +82\ kJ.mol^{-1}$. In the late 1970s and early 1980s, General Atomics (GA, Normann *et al.*, 1981) found that the conversion is greatly improved by using excess iodine and water.

Excess iodine not only shifts equation [1] equilibrium towards the production of acids, but also causes the two acids to spontaneously separate into two aqueous solutions: hydroiodic rich (HI_x) heavy phase and sulfuric acid (SA) rich light phase. However, when this excess is too large, the compositions of the two phases stop changing because the iodine saturation point is reached.

Like iodine, the excess of water causes the equilibrium to be shifted forward but it also causes a substantial change in the reaction Gibbs free energy due to acids dilution. The sulfuric acid complexes with water and hydroiodic acid complexes with water and iodine, according to the following equations (Elder et al., 2005):

$$H_2SO_4 + 4H_2O \leftrightarrow (H_2SO_4 + 4H_2O)_{aq} \qquad \Delta G^{\circ,400\ K} = -66\ kJ.mole^{-1} \qquad (R2) \quad [2]$$

$$2HI + 8I_2 + 10H_2O \leftrightarrow (2HI + 10H_2O + 8I_2)_{aq} \quad \Delta G^{\circ,400\ K} = -104\ kJ.mole^{-1} \quad (R3) \quad [3]$$

Combining the three reactions (R1), (R2) and (R3) leads to a modified Bunsen reaction proposed by GA which has a negative Gibbs free energy: $\Delta G°(400\ K) = -88\ kJ.mol^{-1}$:

$$9\ I_2 + SO_2 + 16\ H_2O \leftrightarrow (2\ HI + 10\ H_2O + 8\ I_2)_{aq} + (H_2SO_4 + 4\ H_2O)_{aq} \qquad (R4) \quad [4]$$

3. Bunsen section thermodynamic modeling

3.1. Liquid – liquid equilibrium calculation basis

The liquid – liquid equilibrium between phases involves activities equality of each component *i* present in both phases, expressed in terms of activity coefficients γ_i by:

Equilibrium equations: $$\gamma_i(T, \mathbf{x}) \cdot x_i - \gamma_i(T, \mathbf{x}') \cdot x'_i = 0 \qquad i = 1, Nc \qquad [5]$$

At given temperature, pressure and global compositions **z**, the compositions of each phase at equilibrium (**x** and **x'**) and the split ratio τ are calculated by solving an isothermal liquid-liquid equilibrium flash model which includes equations [5-7]:

Partial mass balances: $$z_i - (1-\tau) \cdot x_i - \tau \cdot x'_i = 0 \qquad i = 1, Nc \qquad [6]$$

Summation equation: $$\sum_i x_i - \sum_i x'_i = 0 \qquad [7]$$

An excess Gibbs free energy model is then needed to calculate the activity coefficients.

3.2. The new model versus previous models

In order to compute activity coefficients, a symmetric convention is adopted that combines the UNIQUAC model (Abrams and Prausnitz, 1975) with Engel's electrolytic solvation model (Engels, 1990): the first model accounts for short range interactions by means of two terms (combinatorial and residual) and the second model accounts for complexes formation.

Until this work, all published models for this system were derived from a classical electrolyte approach. First, Davis and Conger (1980) proposed to apply the Pitzer-Debye-Hückel model (Pitzer, 1980). Later Mathias (2002) has used the electrolyte-NRTL model developed by Chen and Evans, (1986) that combines the Pitzer-Debye-Hückel model for long-range ion-ion electrostatic interactions with the NRTL theory (Renon and Prausnitz, 1968) for short-range energetic interactions. Recently, O'Connell (2008) proposed an improvement of Mathias' model.

3.3. Model description

The proposed model considers three solvation reactions:

1. Solvation of HI by H_2O:

$$m_1H_2O + HI \leftrightarrow 2C_1 \qquad \text{(RS1)} \quad [8]$$

with $\{2C_1\} \equiv \{[H_3O^+,(m_1\text{-}1)H_2O]; I^-\}$

2. Solvation of H_2SO_4 by H_2O:

$$m_2H_2O + H_2SO_4 \leftrightarrow 2C_2 \qquad \text{(RS2)} \quad [9]$$

with $\{2C_2\} \equiv \{[H_3O^+,(m_2\text{-}1)H_2O]; HSO_4^-\}$

3. Solvation of HI by I_2 in the presence of water:

$$m_1H_2O + m_3I_2 + HI \leftrightarrow 2C_3 \qquad \text{(RS3)} \quad [10]$$

with $\{2C_3\} \equiv \{[H_3O^+,(m_1\text{-}1)H_2O]; [I_3^-,(m_3\text{-}1)\,I_2]\}$

where C_i are the resulting complexes, as defined by Engels (1990), and m_i the solvation numbers. The corresponding solvation constants K_i are function of the temperature:

$$K_i = exp\left(A_{K_i} + B_{K\,i}/T\right) \qquad [11]$$

Each solvation equilibrium introduces a new species, the complex C_i, and requests 3 parameters to be estimated: m_i, A_{Ki} and B_{Ki}. Therefore, the whole system with N_C=7 species (I_2, HI, H_2O, H_2SO_4, C_1, C_2, C_3), has 93 parameters to be estimated from experimental data: 84 parameters to estimate n×(n-1) UNIQUAC binary interaction parameters with their temperature dependence and 9 parameters for the three solvation equilibria. This number is too high in view of available experimental data.

3.4. Model assumptions

On the basis of the operating conditions of Bunsen reaction (water and iodine in large excess (O'Keefe *et al.*, 1982)), the very large immiscibility gap between these two species (Kracek, 1931) and the presence of I_3^- ions in the solution (Palmer and Lietzke, 1982), a first series of hypotheses can be formulated:

a. Total dissociation of both acids (HI and H_2SO_4) in the presence of an excess of water; the second acidity of H_2SO_4 being neglected.
b. Solvation of all H^+ ions by three H_2O molecules, independently whether they come from HI or H_2SO_4 totally dissociated.

The hypotheses a and b enable to substitute the measured species HI and H_2SO_4 by two apparent species, that we write $[HI,3H_2O]$ and $[H_2SO_4,3H_2O]$. They correspond to the complexes $\{[H_3O^+,(m_1\text{-}1)H_2O]; I^-\}$ and $\{[H_3O^+,(m_2\text{-}1)H_2O]; HSO_4^-\}$ of solvation reactions (RS1) and (RS2) with $m_1 = m_2 = 3$.

c. Solvation of each hydrated complex [HI,3H_2O] by one mole of I_2 (m_3=1) following the reaction (RS3). The resulting complex is noted {[HI,3H_2O]; I_2} and corresponds to {[H_3O^+,(m_1-1)H_2O]; [I_3^-,(m_3 – 1) I_2]} with m_1=3 and m_3=1.

As a result, the number of species drops from 7 to 5 whereas the number of parameters to identify drops from 84 to 40, with only 2 solvation parameters. It is still too much regarding the number of available experimental data. Thus, a second series of hypotheses is formulated about binary interactions:

d. The [H_2SO_4,3H_2O] hydrated complex interacts like water with other species. Engels' results (1990) on H_2O–H_2SO_4 vapor-liquid equilibria modeling showed that the interaction between water and the solvation complex of H_2SO_4 by H_2O is negligible.
e. The complex {[HI,3H_2O]; I_2]} interacts like iodine with other species.

Note that the hypotheses d and e imply that the chemical theory alone allows to partially represent the non ideal behavior of the solution.

Finally, the number of parameters of the reduced model drops to 14 (including temperature dependency of UNIQUAC A_{ij} binary interaction parameters and 2 solvation parameters) and the binary interaction parameters matrix becomes:

Table 1. Binary interaction parameters matrix.

	I_2	{[HI,3H_2O]; I_2}	H_2O	[H_2SO_4,3H_2O]	[HI,3H_2O]
I_2	0		A_{12}		A_{13}
{[HI,3H_2O]; I_2}					
H_2O	A_{21}		0		A_{23}
[H_2SO_4,3H_2O]					
[HI,3H_2O]	A_{31}		A_{32}		0

The new species [HI,3H_2O], [H_2SO_4,3H_2O] and {[HI,3H_2O]; I_2} are created in the data base of Simulis® Thermodynamics properties server and their UNIQUAC parameters r_i and q_i are estimated according to Bondi's group contribution method (Bondi, 1964).

4. Available experimental data

Nowadays, available experimental data were published by Sakurai *et al.*, 1999, 2000; Giaconia *et al.*, 2007; Lee *et al.*, 2008. The influence of HI (k), I_2 (m) and H_2O (n) mole numbers for one sulfuric acid mole on liquid-liquid equilibria is studied. The data points are noted 1/k/m/n. General Atomic typical stoichiometry of reaction (R4) corresponds to a 1/2/8/14 data point. Most of those data are compiled in Lee et al. (2008). Covering the largest temperature range from 25°C to 120°C, they are used in this work.

5. Results and discussion

The reduced model reproduces demixtion for all experimental data. Table 2 reports the mean and median relative differences between experimental and calculated molar fractions at each temperature. The results highlight that the relative errors are smaller for the main species (H_2O, HI, I_2 in HI_x phase and H_2O, H_2SO_4 in SA phase) than for impurities (HI, I_2 in SA phase and H_2SO_4 in HI_x phase) in both phases. That is because impurities experimental absolute values are small and lead automatically to larger relative errors and they likely bear larger uncertainties. Otherwise, there are no significant differences between the various temperature sets, hinting at a reasonable temperature dependency of the parameters.

Table 2. Mean and median relative differences between experimental and calculated molar fractions upon the data of Lee et al. (2008)

T (°C)	All species		Impurities		Main species	
	mean	Median	mean	Median	Mean	median
25	13.5%	3.4%	30.5%	30.6%	3.3%	1.9%
40	17.1%	6.6%	32.7%	18.2%	7.7%	3.1%
60	20.0%	9.9%	41.9%	30.2%	6.9%	3.7%
80	16.0%	7.3%	31.8%	27.1%	6.5%	3.3%
100	20.5%	7.6%	45.6%	37.8%	5.5%	1.6%
120	15.9%	4.2%	33.2%	27.9%	5.6%	1.5%

Figure 2 present results by comparing the calculated and experimental molar fractions for HI_x phase at 80°C. The histograms below detail the variation of errors with variables 'm' and 'n' for a given composition ($HI/H_2SO_4/I_2/H_2O$) = (1/2/m/n).

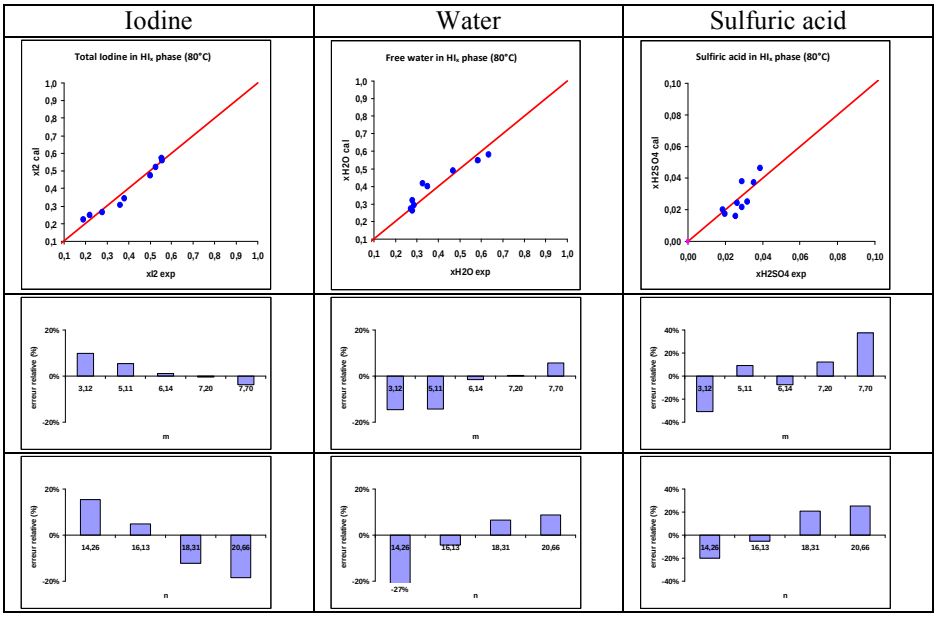


Figure 2. Calculated compositions versus experimental data for the HI_x phase at 80°C: main species are represented by water and iodine and impurities by sulfuric acid.

6. Conclusion

To our knowledge, for the first time a thermodynamic model representing the liquid-liquid equilibria of the Bunsen section of the Sulfur – Iodine thermochemical cycle over a large temperature range is proposed. This model is based on UNIQUAC's activity coefficient model combined with solvation equilibria.

Due to the high level of complexity of the system and the lack of experimental data, a fine description of all species and phenomena cannot be obtained. The formulation of

sound hypotheses, concerning the species present and their interactions, allows to reduce significantly the number of parameters and enables their estimation. The reduced model reproduces demixtion for all data points with an acceptable mean relative error: below 6% for the main species (water and sulfuric acid in the H_2SO_4 rich phase; water, iodine and hydrogen iodide in the HI_x rich phase) and below 30% for the impurities (iodine and hydrogen iodide in the H_2SO_4 rich phase; sulfuric acid in the HI rich phase).

References

Abrams D.S., Prausnitz J.M., 1975, Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems, AIChE Journal, 21(3), 116-128.

Bondi A., 1964, van der Waals Volumes and Radii, J. Phys. Chem., 68, 441-451.

Chen C.C., Evans L.B., 1986, A local composition model for the excess Gibbs energy of aqueous electrolyte systems, AIChE Journal,32:444-454.

Davis M.E., Conger W.L., 1980, An entropy production and efficiency analysis of the Bunsen reaction in the General Atomics sulfur – iodine thermochemical hydrogen production cycle, Int. J. Hydrogen Energy, 5, 475-485.

Elder R.H., Priestman G.H., Ewan B.C., Allen R.W.K., 2005, The separation of HIx in the sulphur – iodine thermochemical cycle for sustainable hydrogen production, Trans IChemE, Part B., 7, 343-350.

Engels H., 1990, Phase equilibria and phase diagrams of electrolytes, Chemistry data Series, Volume XI, Part I., Published by DECHEMA.

Giaconia A., Caputo G., Ceroli A., Diamanti M., Barbarossa V., Tarquini P., Sau S., 2007, Experimental study of two phase separation in the Bunsen section of the sulfur – iodine thermochemical cycle, Int. J. Hydrogen Energy, 32, 531-536.

Hadj-Kali, M.K., Gerbaud V., Borgard J.M., Baudouin O., Floquet P., Joulia X., Carles P., 2009, HIx System Thermodynamic model for Hydrogen Production by the Sulfur-Iodine Cycle, Int. J. Hydrogen Energy, In press.

Kracek F.C., 1931, Solubilities in the system water-iodine to 200°C, J. Phys. Chem., 35, 417-422.

Lee B.J., No H.C., Yoon H.J., Kim S.J., Kim E.S., 2008, An optimal operating window for the Bunsen process in the S-I thermochemical cycle, Int. J. Hydrogen Energy, 33, 2200-2210.

Mathias P.M., 2002, Modeling the sulfur-iodine cycle: Aspen Plus building blocks and simulation models, report GA and Sandia National Laboratories.

Normann J.H., Besenbuch G.E., Brown L.C., Thermo-chemical water-splitting cycle. Bench scale investigations and process engineering. DOE/ET/26225, 1981

O'Connell J., 2008, Oral Conference at ESAT (29/05/2008 – 01/06/2008), Cannes, France.

O'Keefe D., Allen C., Besenbruch G., Brown L., Norman J., Sharp R. and McCorkle K., 1982, Preliminary results from bench-scale testing of a sulfur-iodine thermochemical water-splitting cycle , Int. J. Hydrogen Energy,7(5), 381-392.

Palmer A.D., Lietzke M.H., 1982, The equilibria and kinetics of iodine hydrolysis, Radiochimica Acta, 31, 37-44.

Pitzer K., 1980, Electrolytes. From dilute solutions to fused salts, J. Am. Chem. Soc.,102, 2902-2906.

ProSim, 2004, Simulis® Thermodynamics User Guide, http://www.prosim.net.

Renon H., Prausnitz J.M., 1968, Local compositions in thermodynamic excess functions for liquid mixtures, AIChE Journal, 14(3), 135-144.

Sakurai M., Nakajima H., Onuki K., Ikenoya K., Shimizu S., Preliminary process analysis for the closed cycle operation of the iodine-sulfur thermochemical hydrogen production process, Int. J. Hydrogen Energy, 1999;24:603-612

Sakurai M., Nakajima H., Onuki K. And Shimizu S., 2000, Investigation of 2 liquid phase separation characteristics on the Sulfur – Iodine thermochemical hydrogen production process, Int. J. Hydrogen Energy, 25, 605-611.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Improving accuracy of refinery optimization by the on-line characterization of crude oil

Patricia A. Pantoja[a], Francisco F. Sotelo[a], Anderson C. Gatti[a], Mariana F. Ciriaco[a], Antonio C. Katata[c], Galo A.C. Le Roux[a], Claudio A.O. Nascimento[a,b]

[a]*LSCP/CESQ - Department of Chemical Engineering, Polytechnic School - University of São Paulo. Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil.*
[b]*CEPEMA-USP Center for Environmental Research and Training – Universidade de São Paulo, Rodovia Cônego Domênico Rangoni KM 270, 11573-000, Cubatão, SP, Brazil.*
[c]*Petrobras, Rua Chile, 50, Rio de Janeiro, RJ, Brazil.*

Abstract

The simulation of crude oil distillation units, necessary for the optimization and the on-line performance assessment is very sensitive to the characterization of the feed. The conventional on line characterization technique is based on considering that the crude oil feed is the combination of known crude oils present in a database combined by volume average. Alternatively, a technique based on near-infrared spectroscopy that is able to characterize the petroleum on-line was implemented in order to characterize the crude oil. Prediction models for the main physico-chemical properties have been developed using chemometrics. For Real Time Optimization (RTO) a reference model must be frequently adjusted to fresh operation data. The adaptation of the model is necessary in order to ensure that it represents the system even if there are changes in its constitutive parameters. This feature also ensures that the mathematical optimization problem to be solved in order to obtain the set points of a new operating mode for the process does represent an optimal solution, or at least is close to it. Both the characterization techniques were applied to petroleum samples available for a very long interval of operation (a year) and their result are introduced into the detailed model that is adjusted to the operating data available for this same time range. The accuracy of the predictions shows that the new on-line technique would drive the system to better operating modes.

Keywords: Petroleum; Near Infrared Spectroscopy; Real Time Optimization.

1. Introduction

Real-time characterization of crude petroleum is an important objective for the petroleum industry because it has been pointed as the main factor for the increase in productivity (Dhaval et al., 2003). The crude distillation unit (CDU) is at the heart of any petroleum refinery and is one of the most complicated operations in the field of separation processes. The products from CDU are basically mixtures of a great variety of hydrocarbons that it is almost impossible to characterize in terms of individual components. Moreover, it is sufficient to characterize these products in terms of a certain number of properties, usually measured off line, in laboratory, at intervals of 8 - 24 h. (Dhaval *et al.*, 2003).

The simulation of a CDU separation processes involves many complex nonlinear relations between independent and dependent variables of the system, although there are no chemical reactions. The independent variables are usually crude oil properties and the manipulated variables of the CDU, such as energy inputs, reflux ratios, and product flow rates; the dependent variables are the oil product qualities, the system operating performance and the plant profit (Chau-kuang *et al*. 2004).
Techniques based on near-infrared (NIR) spectroscopy have been used to characterize and classify Brazilian crude oil samples. Based on the spectroscopic data, prediction models of Simulated Distillation (SimDis), density (API degree) and viscosity were developed, using chemometrics and neural network techniques. The application of chemometrics to spectroscopic data in order to develop mathematical models allows the extraction of information in order to predict physical properties.

2. NIR spectroscopy

NIR is probably the spectroscopic technique with the most applications in industry in general and in the petroleum industry in particular (Hidajat *et al.* 2000; Hannisdal *et al.* 2005; Falla *et al.* 2006; Abbas *et al.* 2006). NIR region extends from 14000 – 4000 cm-1 (800-2500 nm). Absorption bands in the NIR region are related to the overtone and combination bands of the vibrations of functional groups (–CH, -NH and –OH). Highly overlapping absorption peaks make spectra too complex to be interpreted qualitatively and low sensitivity limits quantitative analysis. However, with advances in instrumentation and chemometrics, NIR spectroscopy has emerged as a valuable analytical tool. Mathematical algorithms extract useful chemical information from NIR spectral features. Analysis in the NIR region is particularly advantageous because it permits the usage of fiber-optic technology for remote sensing because various optical fibers are available as NIR radiation guides (Chung *et al*. 2007).

3. Crude distillation unit simulator

A good CDU simulator usually encompasses an analytical model for multi-component distillation based on equilibrium stage concept. However, perfect equilibrium is seldom achieved because of poor mixing, insufficient time of contact, etc., and hence, stage efficiencies must be incorporated in the equilibrium relations. Since there is yet no reasonable method to calculate these efficiencies accurately, these are actually treated as tuning parameters, which will allow the model predictions to match the plant performance. Efficiencies are functions of composition in both phases and other operating conditions and hence, can be expected to change with time (Dhaval *et al.*,2003).

4. Experimental

The SimDis curve (accumulated weight percent vaporized at a given temperature), the density and the viscosity were correlated to the NIR spectra. These correlations were obtained in a combination of chemometrics pretreatment techniques with artificial neural network models.

4.1. Samples

The crude oil samples came from storage tanks of a Brazilian refinery. The samples were sealed and stored below 5ºC in order to reduce the loss of light components and the natural degradation processes. API density and viscosity analysis were performed

following ASTM norms. Viscosity of the samples ranges between 5.1 cp and 36.1 cp and API from 23.3 and 33.0. More than a half of the samples have their API degree concentrated between 25 and 28. For viscosity the distribution is approximately uniform between 5 and 24 cp. There is an outlier with viscosity 36.1 cp.

4.2. NIR apparatus and software

The oil samples were analyzed in a NIR spectrometer, model FTLA2000 – 160 (ABB Bomem), 4000 – 14000 cm^{-1} region, with DTGS detector and Tg source. A type of fiber optics was used, model 041.202 – NIR fiber of length 6m (off-line tests). The transmittance NIR sensor, model 661.760-NIR, had a 1 mm path length. The fiber optics and the transmittance NIR sensor were from Hellma. NIR spectra (bi-dimensional data) were recorded over the wavelength range 5600 – 6100 cm^{-1} at 1.93 nm intervals, which corresponds to the first overtones region of C-H stretching vibrations.

4.3. Simulated Distillation

Simulated Distillation analyses (ASTM 7169) were performed on an Agilent Technologies 7890A gas chromatograph (GC), equipped with cold-on column (COC) inlet, flame ionization detector (FID) and a cryogenic column oven system. An oxygen trap (Agilent Technologies) was used in the carrier gas line (helium) to remove traces of oxygen. The detector signal was recorded as area slices (time intervals) for consecutive increasing retention times. Agilent Chemstation B.03.01 software was used for data acquisition and simulated distillation software (Wasson) was used processing data as well as boiling point distribution of the samples. Prior to injection, the samples were diluted in carbon disulfide with low benzene content (Vetec) to about 2 wt%. High purity CS_2 was used as a solvent because of its miscibility with residues, low boiling point and low response factor in the FID. The injection volumes were 0.2 µL. The distribution of boiling points for carbon number standards from C5 to C120 (Analytical Controls) were obtained using a megabore capillary column MXT-500Sim Dist 6.0 m length, 0.53 mm ID, 0.15 µm film thickness (Restek). The GC oven was programmed from -20 to 425 ºC at a rate of 15ºC /min. External standard reference gas oil (Wasson) that resembles the samples of crude oil was used, diluted in carbon disulfide to about 2 wt %. A baseline compensation analysis, or baseline blank, was performed exactly as for a conventional analysis, except that no injection was done. The blank analysis was necessary due to the usual occurrence of chromatographic baseline instability, due to the extreme temperature conditions that make the column subject to bleed. The blank analysis must be performed periodically between sample measurements because the baseline varies slightly from sample to sample.

4.4. Prediction models

Models for prediction of the Simulated Distillation (SimDis) density (API degree) and viscosity were developed, using chemometrics and neural network techniques. The methodology developed is according to “Characterization of crude petroleum by NIR” of F.S. Falla et al. 2006.

In figure 1 parity plots for the prediction of 20% and 80% are presented.

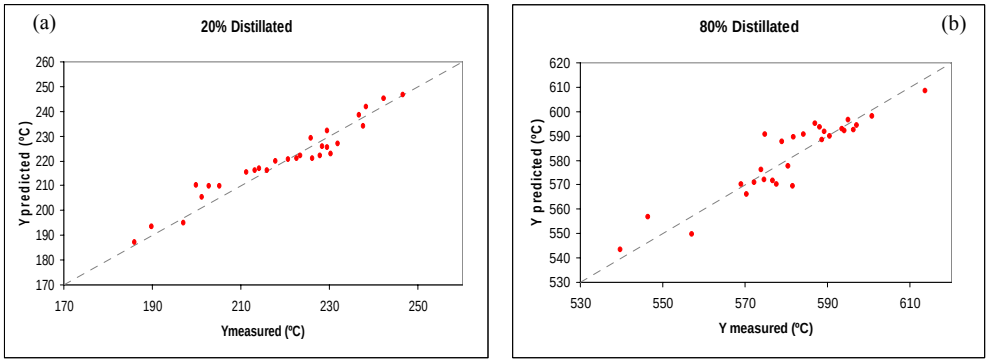


Figure 1- Parity plots for the prediction of (a) 20% and (b) 80% Distillated.

5. Methodology

A simulation model of five distillation columns from Henrique Lage (REVAP) refinery was built in Aspen HYSYS® simulation software in steady state mode. The objective is to monitor the results of the NIR analysis on the accuracy of predictions obtained through the process simulator.

In the simulation model oil characterization must be introduced. The idea is to introduce both characterization results, from the data base and obtained from NIR analysis and to assess the prediction of the models, by comparing the predictions with the data contained in the operation data base.

The procedure for the simulation is:

- The characterization of the petroleum is introduced in the simulation;
- The degrees of freedom of the system are set based on the data available on the process data base which is analyzed in order to select steady-state operating conditions (Carrillo Le Roux et al. 2008);
- As measurements are subject to errors a reconciliation step is performed.

To reconcile the variables, the optimizer tool in Aspen HYSYS® was used. Some values are corrected, interfering in other parameters, in order to minimize an error function of difference between refinery data (target) and calculated values.

The loss function for the first tower is given by:

$$L = \sum_i w_i . \left(\frac{F_i - F_i^{meas}}{F_i^{meas}} \right)^2 + \sum_j \left(\frac{T_j - T_j^{meas}}{T_j^{meas}} \right)^2$$

where F_i are the flows (Volumetric flow in m^3/h) and T_j are the temperatures (given in Kelvin). The weights w_i are related to the importance given to each measurement.

The SimDis distillation curve, the density and viscosity of some petroleum samples used at the refinery in the year 2007 were available. The first column is a preflash tower which has 15 stages. There are 14 measured variables related to this column available

from the refinery, where 5 of them are used as independent variables as this simulation unit has 5 degrees of freedom and 9 dependent variables.

The stabilizer tower is fed by naphta product from the preflash. Its model has 35 trays, a condenser, a reboiler, one inlet stream, one bottom outlet and three condenser draws. There are 15 variables measured at the refinery, where 4 of them are used as independent variables.

The bottom stream from the preflash column is fed to the third column, an atmospheric tower. Its model has 45 trays, a condenser, a reboiler, three pump-arounds, one side stripper, three inlet streams (including two vapor feeds) and seven outlets. There are 35 measured variables connected to this column available from the refinery, where 8 of them are considered independent.

The vacuum tower and the side vacuum column are not still considered in the model. Redundancy in process data is present for all the columns, because the number of variables is much bigger than the number of degrees of freedom.

6. Results

In the simulation of the preflash tower, 4 different sets of independent variables were assumed. Set 1 considers as independent the reflux flow, the gas flow (from the condenser), the naphtha flow, the subcooling temperature and the bottom reflux temperature. Set 2 considers as independent: reflux flow, gas flow, naphtha flow, bottom reflux flow and subcooling temperature. Set 3 considers: reflux flow, gas flow, bottom reflux flow, subcooling temperature and bottom reflux temperature. Set 4 includes: reflux flow, gas flow, bottom reflux flow, subcooling temperature and bottom temperature.

In order to minimize the loss function, all the flows considered in the calculation of each method were varied. Table 1 shows the value of the loss for different operation steady states of the refinery (and corresponding petroleum) and different methods.

Table 1. Value of the loss function for different input sets and operating points.

	P 1	P 2	P 3	P 4	P 5
Set 1	10,5	52,9	54,4	24,8	115,4
Set 2	4,5	24,8	16,5	11,6	23,9
Set 3	10,4	39,4	41,8	24,9	56,4
Set 4	12,3	40,5	44,1	27,7	57,4

According to table 1, at condition P1 the lowest loss is obtained while at condition P5 the biggest is reached. Set 2, that is the only one using 4 flows as input, had the best reconciliation.

A table 2 compares calculated and target parameters for the case that had minimum error that is set 2 and condition of operation P1.

Table 2. Calculated and measured variables for the case that presented minimum error.

Parameter	Calculated	Measured	Unit
Reflux volume flow	210,2	220,5	m3/h
Bottom reflux volume flow	667,1	645,8	m3/h
Bottom volume flow	1489,5	1418	m3/h
Naphtha volume flow	185,8	214,5	m3/h
Gas flow at standard conditions	1693,8	1700	Nm3/h
Top temperature	157,9	146	^{o}C
Bottom reflux temperature	433,8	375,8	^{o}C
Bottom temperature	313,2	288,4	^{o}C

7. Conclusion

The methodology is promising, because the simulation permits not only to compare the predictions but also to obtain the reconciliation of the measured variables. The real time characterization of the crude oil feed will allow the optimization problem to reach more realistic solutions.

References

Abbas º, N. Dupuy, C. Rebufa, L. Vrielynck, J. Kister, A. Permanyer (2006) "Prediction of source rock origin by chemometric analysis of Fourier Transform Infrared-Attenuated Total Reflectance Spectra of oil Petroleum: evaluation of aliphatic and aromatic fractions by Self-Modeling mixture analysis". Appl. Spectrosc. 60, 304.

G.A. Carrillo Le Roux, B.F. Santoro, F.F. Sotelo, M. Tessier, X. Joulia , (2008), "Improving Steady State Identification. In: European Symposium on Computer Aided Process Engineering", Lyon. Computer-Chemical Engineering, Amsterdam, Elsevier, 2008. v. 25. 459-464

L. L. Chau-kuang, T. Y. Chung-Kuang, T. Ming-te (2004)"Expert system of a crude oil distillation unit for process optimization using neural networks". Expert Systems with applications 26. Pg 247 – 255.

H. Chung, (2007). "Quality characterization of crude oils by particial least square calibration of NIR spectral profiles". App. Spect. Rev.,42, 251.

D.J. Dhaval, Z.D. Murtuza, V. K. S. Datyadev, S. Ganguly, D. N. Saraf (2003) "Online tuning of a steady state crude distillation unit model for real time applications". J. Process Control 13. Pg. 267 – 282.

F. S. Falla, C. Larini, G. A. C. Le Roux, F. H Quina, L. F. L. Moro, C. A. O. Nascimento (2006). "Characterization of crude petroleum by NIR". J. Petro. Sci. Eng. 51, 127.

A. Hannisdal, P. V. Hemmingsen, J. Sjöblom (2005). "Group-Type analysis of Heavy Crude oils using vibrational spectroscopy in combination with multivariate analysis". Ind. Eng. Chem. Res. 44, 1349.

K. Hidajat, S. M. J. Chong, (2000) " Quality characterization of crude oils by partial least square calibration of NIR spectral profiles". Near Infrared Spectrosc, 8, 53.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A method of representation of petroleum by a real components substitute mixture

Franklin D.R. Cuellar [a] Galo A.C.Le Roux [a],

[a]*LSCP/CESQ - Department of Chemical Engineering, Polytechnic School - University of São Paulo. Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil. E-mail address: galoroux@usp.br,*

Abstract

In the refining of petroleum it is important to predict the distribution of the products and their quality. Conventionally, petroleum is characterized by methods based on distillation analysis, like the True Boiling Point (TBP) and ASTM D86, which are the most usual analytical methods. In order to predict the distribution of the products and its relation to the operating conditions, steady-state simulation is extremely important. Conventionally, tens of pseudocomponents are generated in order to represent the petroleum. Pseudocomponents have individual properties that do not necessarily correspond to any real component and are obtained by semi-empirical methods. More recently, the approximation of the petroleum by a reduced set of real components has been proposed. In order to design the substitute mixtures, dynamic models representing the standard tests for TBP and ASTM-D86 were developed as differential algebraic equations (DAE). Results for the simulation of these properties, experimental results and properties obtained from some commercial programs are compared, showing that the models developed here are reliable.

Keywords: TBP, ASTM D86, Dynamic simulation .

1. Introduction

In the refining of petroleum it is important to predict the distribution of the products and their quality. Conventionally, petroleum is characterized by methods based on distillation analysis, like the True Boiling Point (TBP) and ASTM D86, which are the most usual analytical methods. The ASTM D86 determines the boiling range of the petroleum product by performing a simple batch distillation. The distillation is performed in a laboratory batch distillation unit at ambient pressure under conditions that are designed to provide approximately one theoretical plate fractionation. The TBP test method employs a fractionating column that has an efficiency of 14 to 18 theoretical plates operated at a reflux ratio of 5.

In order to predict the distribution of the products and its relation to the operating conditions, steady-state simulation based on first-principles, in terms of mass and energy balances and equilibrium relations, is extremely important. As the number of components is very high, the mixture composition cannot be directly used. Conventionally, tens of pseudocomponents are generated in order to represent the petroleum. Pseucomponents have individual properties that do not necessarily correspond to any real component and are obtained by semi-empirical methods. The appropriate parameters of characterization for use in thermodynamic models, such as critical temperature, critical pressure, critical volume, acentric factor and molecular weight is possible with the use of equations and empirical correlations.

More recently, the approximation of the petroleum by a reduced set of real components has been proposed (Eckert, 2005). The advantage of this approach is that the substitute mixture can be used in processes where chemical reactions take place. As the ability of modern computers to perform calculations is enormous, using a large number of real components is feasible and is no more necessary to use a reduced number of pseudocomponents. For real components, precise estimation of the critical properties is not an issue, because these are set *a priori* by the nature of the components. However the main challenge of this new method is determining the components and their quantity that are chosen to represent the original mixture. An important aspect to consider is the quality of data available, as for heavy oil very few information is available in the literature.

An important task is to fit the experimental information from the original mixture with the representation methods. Conventional analyses are standardized by the American Society for Testing and Materials (ASTM). The different standards that can be used to determine the characteristics of volatility in the mixture are the ASTM D2892, ASTM D86 and ASTM D1160.

The objective of this work is to generate the ASTM characterization curves D2892 and D86 by simulation from real components mixture. The ASTM standards were implemented in a batch distillation simulator. This step will make possible to compare the results obtained by simulation with the analysis obtained experimentally data. This will integrate a methodology for real components selection.

2. Method of real components

The method of representation of petroleum by a substitute mixture of real components avoids the use of empirical equations for determining the critical properties such as critical pressure, critical temperature, among others, essential for working with pseudocomponentes. For the ASTM D2892 (True Boiling Point curve) there are different analyses that can be applied to perform the selection of the substitute mixture:

a. Distribution of the Bubble Point (ASTM D2892).
b. Distribution of density.
c. Distribution of viscosity.
d. Distribution of Molecular Weight.

These characterization curves are dependent of the mass or volume fraction distilled. The curves (density, viscosity and molecular weight) should be related to the bubble temperature. The bubble point distribution is essential for an initial representation of the multicomponent mixture.

The temperature range of the bubble point (Initial Point-End Point), denoted as ST should be subdivide entirely in continuous intervals without overlapping,. The technical division of the intervals can be diverse. For example equidistant points can be considered, as a starting point, For the selection of the representative components the next steps are recommended:

• The curve (a) is at least necessary;
• One component is selected for each of subinterval of the curve (a).
• If some or all of the curves (b, c, d) are available, each of the properties available should be weighted to optimize the selection of representative components in accordance the following metrics:

$$\sum_{j=a}^{b,c,d} w_j \left(\frac{\zeta_{r,j,z} - \zeta_{m,j,z}}{\zeta_{m,j,z}} \right) \rightarrow \min_z \qquad (1)$$

Where j corresponds to the available properties (a, b, c e d) of components from the database respectively and the index m is attached to the measured values. The expression is calculated for the candidate component, z = 1 ,.... Zk. and the component with the lowest value of the expression above is selected to represent the range.
When the selection of real components is complete, the quantity of each component can be specified by analogy to the pseudocomponents to establish the mass or volume of fraction of distillate (Eckert et al., 2003). As defined in equation 2.

$$x_j = \left(\varphi_j^R - \varphi_j^L\right)\left(1 - \varphi_{LEj}^R\right) / \left(\sum_{i=LE+1}^{I} \left(\varphi_i^R - \varphi_i^L\right) \right) \qquad (2)$$

$$j = LE + 1, ..., I$$

where I is the total number of components, LE, is the number of light-end components, R, the right limit of the range, L, the left limit of the rang, and φ the volume (mass) fraction distilled

3. Results

This section presents the simulation results for the ASTM D86 and ASTM D2892 built in BatchSep (AspenTech 2006). The method of representation of petroleum by a real components substitute mixture was applied as the method for the approximation of the characterization. Aspen BatchSep is a tool for the simulation of distillation in batch mode that allows the generation of models with a friendly interface for the user.

3.1. Example 1

Greenfield and Lavole (1998) presented the results of the experimental curve ASTM D86 for a mixture of 7 hydrocarbons with known composition. These results are presented in figure 2. In this case, the comparison was drawn with another commercial simulator: Aspen HYSYS. The procedure used by the commercial simulator is not explicitly stated in the software manual.
Figure 1 shows that results are close but there are some differences: the D86 curves have different IP and EP for both commercial simulators. In the case of BatchSep the beginning of the curve presents a mild difference when compared to the experimental curve. This can be explained by the fact that a possible experimental loss of light components can introduce some error to the experimental results. Otherwise, the rest of the simulated curve fits well the experimental one.

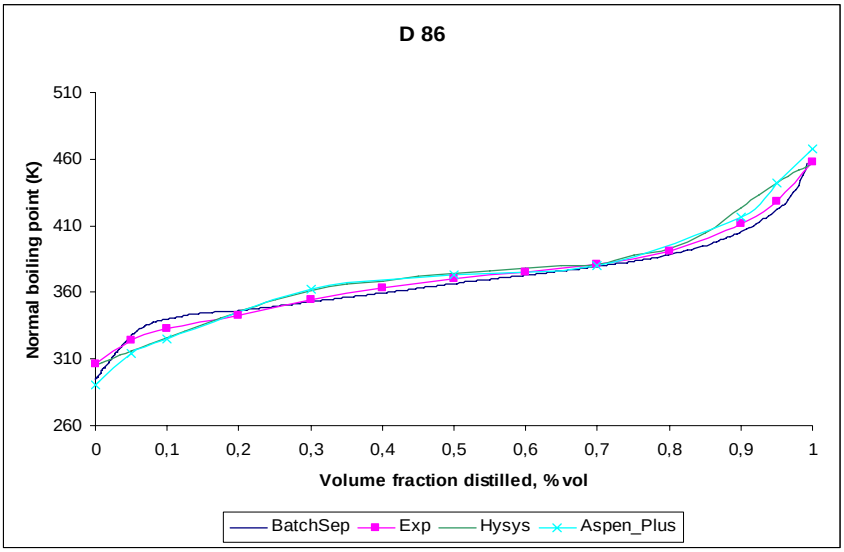


Figure 1. Prediction of the ASTM D86 with commercials simulation and the comparison with the experiment for a mixture of 7 real components.

3.2. Example 2

The experimental information presented by Miquel and Castells (1994) ASTM D 2892 data were used in order to obtain the substitute mixture of real components by the methodology proposed by Eckert (2005). The selected components and their proportion are summarized in Table 1.

Table 1. Components select into the Substitute Mixture

Name	CAS No.	Molar fraction
2,3,3,4-TETRAMETHYLPENTANE	16747-38-9	0.091711
2,4-DIMETHYLOCTANE	4032-94-4	0.114961
ISOBUTYLCYCLOHEXANE	1678-98-4	0.12756
2-ETHYL-P-XYLENE	1578-88-9	0.134563
METHYLCYCLOPENTADIENE-DIMER	26472-00-4	0.110087
1,3,5-TRIETHYLBENZENE	102-25-0	0.121606
PENTAMETHYLBENZENE	700-12-9	0.121845
N-HEPTYLBENZENE	1078-71-3	0.101989
1-ETHYLNAPHTHALENE	1127-76-0	0.075676

The curve generated with the substitute mixture in the simulator BatchSep is presented in figure 2.

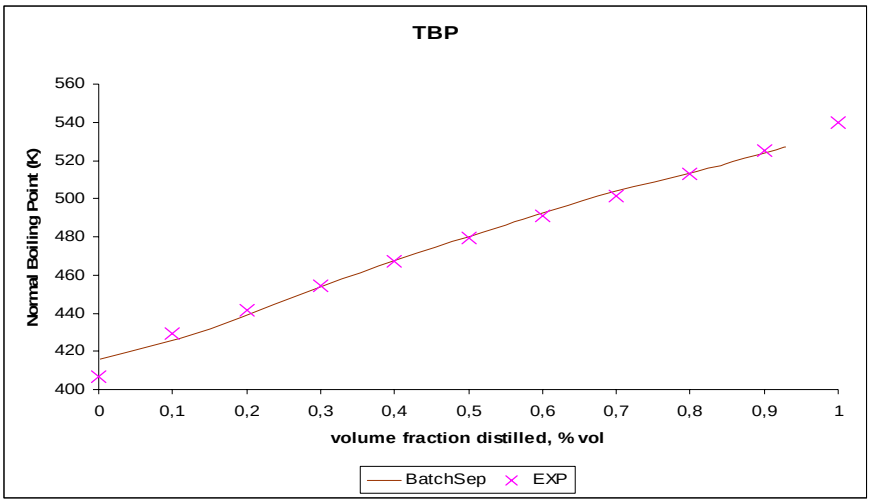


Figure 2. Prediction of the ASTM D2892 with Batchsep.

3.3. Example 3

The experimental information presented by Lojkásek and Ruzicka (1992) was used. The ASTM D 2892 and density data were used in order to obtain the real components substitute mixture by the methodology proposed by Eckert (2005). The selected components and their proportion are not summarized in this paper, but the final mixture contains 18 hydrocarbons obtained with weights $w_T = 1$ and $w_\rho = 1/100$ in equation (1).

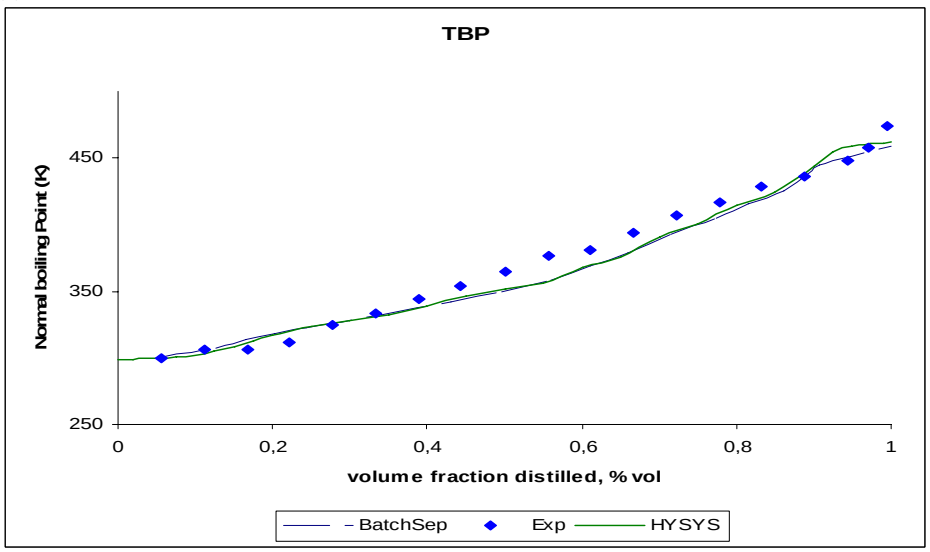


Figure 3. Prediction of the ASTM D2892 with BatchSep and HYSYS.

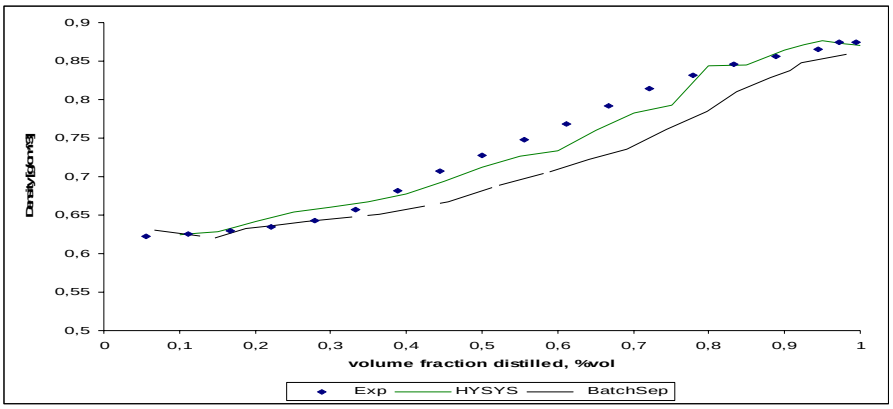


Figure 4. Prediction of the density with BatchSep and HYSYS.

The comparison of the experimental curves with the real components substitute mixture selected (BatchSep) and HYSYS (pseudocomponents) is presented in Figs 3 and 4. It can be observed that the match is fairly good, but the discrepancy is due to the availability of components in the thermodynamic data base.

4. Conclusions

The results presented in this paper show that the method of representation of petroleum by real components substitute mixture can be used in the simulation of complex mixtures and can replace the traditional method of pseudocomponentes. The algorithm suggested for the selection of the mixture needs to be replaced in order to have enough components across the range of temperature considered. This is the next step to be implemented that will generate mixed integer nonlinear problems. The real components substitute mixture must be compared also in the simulation of crude distillation units. This was already done but could not be presented here (Cuellar, 2009).

References

ASTM-METHOD D86. Cullem, 2007, Test Method for Distillation of Petroleum Products at Atmospheric Pressure, in Annual Book of ASTM Standards, American society for testing and material, vol. 05. 01.

ASTM-METHOD D2892, 2005, Test method for distillation of crude petroleum, in Annual Book of ASTM Standards, American society for testing and material, vol. 05. 01.

J. Miquel, 1994, Easy characterization of petroleum fractions (part. 2), Hidrocarbon Processing, v.73, n.12, p. 99-103.

F.D.R. Cuellar, Representação do petróleo através de componentes reais para a simulação de processos de refino, M.Sc. thesis, EPUSP, 2009

E. Eckert, 2005, New approach to the characterization of the petroleum mixtures used in the modeling of separation processes, Computer and Chemical Engineering, v.30, n.2, p. 343-356.

M. Greenfield, 1998, Macroscopic model of the D86 fuel volatility, SAE paper 982724.

M. Lojkásek and V. Ruzicka Jr., 1992, Solubility of water in blends of gasoline, methanol and a solubilizer, Fluid Phase Equilib. v 71 , p. 113–123.

Acknowledgement

Authors would like to acknowledge CAPES and FAPESP for financial support.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Sustain Activities for Real-Time Optimization Models of Ethylene Plants

Rubens Rejowski Jr.[a], Vinay Shah[b], Charles Edward Fontenot[b], Paulo de Tarso[c] and Victória Emilia Neves Santos[c]

a Aspen Technology Inc. Av. Brigadeiro Faria Lima, 3729, CEP 04538-905, São Paulo - SP, Brazil. rubens.rejowski@aspentech.com
b Aspen Technology Inc. 2500 City West Blvd. Suite 1500 Houston, Texas 77042, USA
c Braskem S.A. Rua Eteno, 1561, CEP 42810-000, Camaçari – BA, Brazil. paulo.tarso@braskem.com.br, victoria.santos@braskem.com.br

Abstract

Real-Time Optimization (RTO) applications consist of a paramount tool for the Refining and the Chemical Process Industries. It evaluates the chemical process operational conditions, calculates optimum operating conditions and sends setpoints to the plant in order to maximize the overall profit. The purpose of this paper is to report the sustain activities with respect to the RTO models for Olefins plants in a partnership formed by Aspen Technology, the world's leading supplier of process simulation software, and Braskem S.A., the Brazilian leading petrochemical company. Reliability on the RTO results leads the operators to consistently implement the suggested setpoints to the plants. In addition, operational results showed the RTO models were able to increase the ethylene production and reduce high valued chemicals losses, thus improving plant profit. The results also include drastic reductions in steam import rates.

Keywords: RTO, Olefins Plants, Sustain.

1. Introduction

Real-Time Optimization (RTO) applications are very powerful tools used to improve profitability of Refining and Chemical Process Industries. They consist mainly of a rigorous simulation model of the process coupled with managing applications, links to the plant control system and information databases. The integrated application is able to collect data from the plant and other sources, check their consistency, input them to the simulation model and determine after running an optimization cycle of parameterization and optimization cases, the optimal operating conditions that lead to maximize the plant profit. In 2005 Braskem decided to implement Advanced Process Controls (APC) and RTO in its two olefins plants located at Camacari Petrochemical Complex, Northeast of Brazil. Aspen Technology, the world's leading supplier of software for the refining and chemical industries was chosen for the project development, implementation and support.

The two Braskem RTO models started running in closed loop in 2007. Increase in high valued chemicals production along with reduction of product losses and utilities consumptions has already proven its value generated by project implementation. Due to its plant wide action and rather complex nature, APC and RTO require dedicated support by a team of engineers responsible to sustain the economical results that they return to the company. Training and commitment of operators to RTO are of

fundamental importance for the success of the project as they must be able to analyze and accept or reject the setpoints calculated by the optimizers. Care with information input to the models shall be taken on a daily basis, especially in a naphtha steam cracking unit, where feed composition changes may affect dramatically operating conditions and yields. Model setpoint calculations must also take production scheduling and planning into account. This work describes the activities required to sustain the value produced by the Braskem's Real-Time Optimizers.

2. RTO Aspen Technology Solution for Olefins Plants

AspenONE™ is the market-leading application suite that enables process manufacturers to implement best practices for optimizing engineering, manufacturing, and supply chain operations.

The RTO, Aspen Tech Solution for Olefins Plants is composed by main two modules. The first one is the plant model that is built in Aspen Plus™, which is a market-leading process modeling tool for conceptual design, optimization, and performance monitoring for the chemicals, polymer, specialty chemicals, metals and minerals, and coal power industries. The model is based on steady-state fundamental Engineering principles.

For Real-Time applications and large scope plant models, Aspen Plus™ offers the Equation Oriented mode, in which the plant steady-state models can be solved in Parameter Estimation, Data Reconciliation, Simulation and Optimization modes. In addition, Aspen Plus™ brings the DMO™ solver that solves the non linear chemical process model equations using the SQP algorithm. Moreover, Aspen Plus™ includes links to Technip's SPYRO ethylene cracker models.

The second module is Aspen OnLine™. It enables the execution of an Aspen Plus equation-oriented model for continuous online and closed-loop implementation. The tool contains a steady-state detection module and it also integrates the optimizer with Advanced Process Control (APC) tools. Figure 1 shows the online closed loop cycle for the Aspen Technology RTO application.

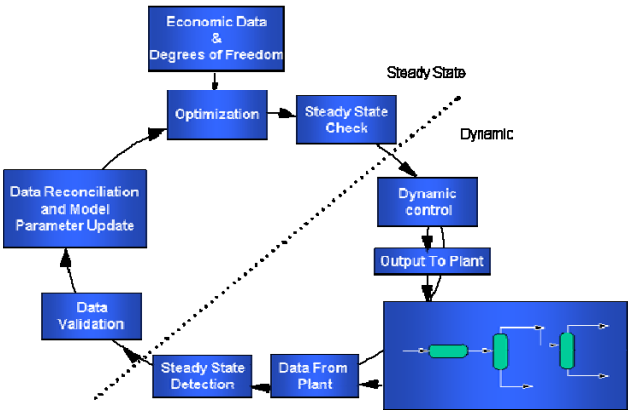


Figure 1 – Aspen Technology RTO Closed Loop Mechanism

3. Literature Review

Olefins plant remains as one of the most challenging chemical process in terms of operational and modeling perspectives. Companies seek for increasing the throughput to maximize production of ethylene and propylene and at the same time reducing the energy consumption.

The true understating of the pyrolysis reactions that take place in the furnace coils are the core component for a reliable operation and also for development of robust models. Van Goethem et al. (2008) developed new options for thermal cracking of ethane and similar light hydrocarbons. The authors validated their model in two applications that ethylene production is maximized at maximum temperature and minimal pressure reaction conditions. Guangying et al. (2008) modeled a quench oil system of an ethylene plant using real components instead of pseudo-components. Mercangöz and Doyle III (2006) developed a RTO model integrated with the control system for the pulp and mill industry. The authors report significant annual savings by the tool. Biegler and Zavala (2009) address the integration of process control models with scheduling and planning models. The authors emphasize the real-time optimization strategies that can be used to merge and replace the steady-state tasks and linear predictive control models. Gubitoso and Pinto (2007) address the operational planning of a real-world ethylene plant.

4. Sustain Activities

Sustain activities can be defined as a set of tasks that aims to maximize the benefits that a tool can deliver inside a plant or a complex. Tools related to chemical process technology, such as APC, Real-Time Optimization, Scheduling and Planning applications are heavily affected by Sustain Activities. It is crucial that companies apply a consistent set of activities that sustain these tools in order to keep the maximum possible benefits. Figure 2 shows the typical benefits from tools related to chemical process technology versus time.

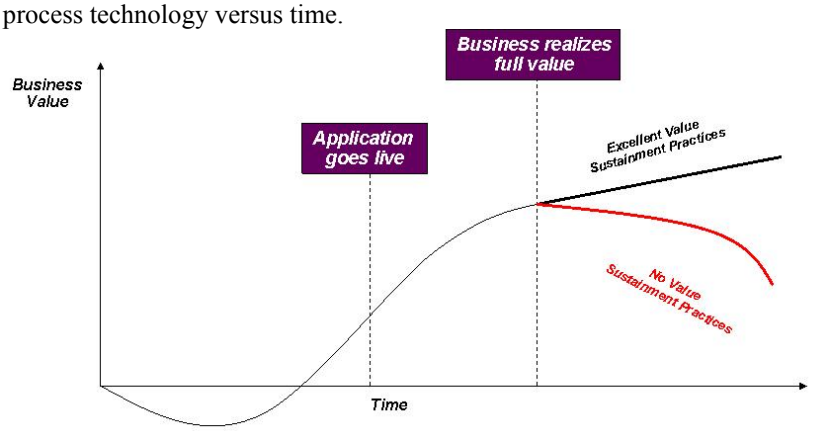


Figure 2 – Typical Benefits versus Time of Chemical Process Tools

The companies get a considerable return rate in all operation levels right after its commissioning. However, as time passes, several challenges take place, thus reducing the tools benefits. If no action is taken, the companies do not get the maximum benefits and some even abandon them. In here, the actions taken in the Aspen Tech – Braskem Sustain Program are described in order to maximize the RTO applications benefits.

4.1. Operator Training

Engineers should carry out an intensive training with operators. They must explain the RTO closed loop mechanism and emphasize the following points:

- RTO system maximizes the plant profits: The system drives the plant to its limits as the operators accept the setpoints. Any questions should be directed to the Engineers. They are able to explain why the RTO suggests the moves on the optimized variables

based on plant status, active constraints and products, raw materials and utilities prices. Also, RTO is safe. Engineers should state that the RTO respect the operational variable bounds and do not cause high disturbances to the plant;

- RTO suggested setpoints should be as clear as possible for operators to review. They should spend a minor time reviewing the setpoints. This can be critical for furnaces operators. Olefins plants typically have 10 to 18 furnaces. The RTO system regularly sends 3 setpoints for each furnace (Total HydroCarbon Feed, Dilution Steam to HydroCarbon Ratio and Severity). Therefore, the operators need to review up to 54 setpoints, which is a considerable amount of information;
- With increasing operator's confidence on the RTO results, the system can be turned to "Auto-Accept" mode, that is, if the operator does not reject the results in the review time, the setpoints are sent to the APC system. Braskem currently operates both RTO systems in Auto-Accept mode.

4.2. Furnace Feed Composition

The furnace liquid feed composition is a crucial component that has to be properly updated to enhance the model precision when calculating the furnace effluent composition. Typically, an Olefins plant can be fed by a wide range of naphthas, such as Full Range, Light, Heavy Naphthas or condensates. The naphtha composition drastically affects the furnace effluent composition. The Aspen Tech APC and RTO system relies on SPYRO^{TM} from Technip. In the beginning of the Sustain Project, the system relied on the online PIONA analyzers to characterize the naphthas that feed the furnaces. This approach presented large offsets between the model predictions and the plant production rates for several products. To eliminate these, a manual laboratory analysis was implemented, where a naphtha sample is collected from the field and a full and open component chromatographic analysis is performed. Although somewhat manual, it was able to enhance the model performance in predicting the plant yields.

4.3. RTO & Scheduling and Planning Integration

Scheduling and planning models consider plant resources, capacities and availability, task sequence and their respective duration. Main decisions for planning/scheduling models consist of discrete and continuous variables, such as the available raw materials to purchase, how much of each product should be produced in a monthly/weekly basis. These models can be applied to a chain of factories or to a single plant and are applied in a monthly or weekly basis.

The common information that Planning, Scheduling and RTO models share are key product target production rates. Therefore, the integration is performed by including the production target rates in the RTO models. The maximized profit of the original model shown in (1) is a linear equation that takes into account the difference of all product revenues by the sum of the raw material and utilities costs.

$$Profit = \sum product_i \;\; C_i - \sum feed_j \;\; C_j - \sum utility_u \;\; C_u \tag{1}$$

The new objective function is shown by equation (2).

$$Profit = \sum product_i \;\; C_i - \sum feed_j \;\; C_j - \sum utility_u \;\; C_u - C_p . \left[\overline{F_p^{TARGET}} \right] \tag{2}$$

The penalty term reduces the profit when the plant production rate is under the production target rates. This is accomplished by adding two constraints in the RTO model, shown by (3) and (4). Constraint (3) states that variable $\overline{F_p^{TARGET}}$ must be greater

than or equal to zero, whereas (4) imposes $\overline{F_p^{TARGET}}$ should be greater than or equal to the difference of the production target $\left(F_p^{TARGET}\right)$ and the plant rates for each product p $\left(F_p^{PLANT}\right)$.

$$\overline{F_p^{TARGET}} \geq 0 \qquad \forall p \qquad (3)$$

$$\overline{F_p^{TARGET}} \geq \left[F_p^{TARGET} - F_p^{PLANT}\right] \qquad \forall p \qquad (4)$$

Two possible scenarios arise for these constraints. The first is when the plant rates are already above the production target. Then, constraint (3) is activated, while (4) is relaxed. Note that in this scenario, the penalty is equal to zero. The second scenario is characterized when the plant rate is under the production target rates. Then, constraint (4) is activated, whereas (3) is relaxed. Therefore, the overall plant profit is penalized.

5. Results

The Sustain Project performed at the Olefins plants at Braskem showed significant results. Due to a confidentiality policy, only relative process results are shown by the following pictures. Figure 3 shows a drastic reduction in the High Pressure steam make-up to total furnace feed and to the ethylene production ratios when the operators accepted the suggested RTO setpoints. In this particular case, Braskem was able to save US$100,000 in High-Pressure steam import in eight days. In addition, operators started accepting RTO setpoint more frequently after the intensive training, Engineers support and after getting acquainted with the setpoint review procedure.

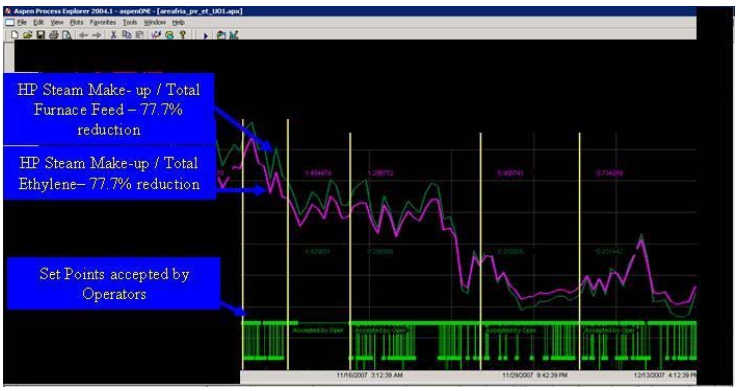


Figure 3 – Reduction in High Pressure Steam Make-up to Olefins Plant

Figure 4(a) shows the ethylene production at the cold side at one of the plants and the ethylene flow rate predicted by the model at the furnaces header. Before the laboratory implementation, the model presented a large offset since there are ethylene losses along the plant. The laboratory analysis procedure implemented eliminated the large offsets. Figure 4(b) shows the RTO system was able to maintain the ethylene production rate at the production target and to increase the pyrolysis gasoline production flow rate.

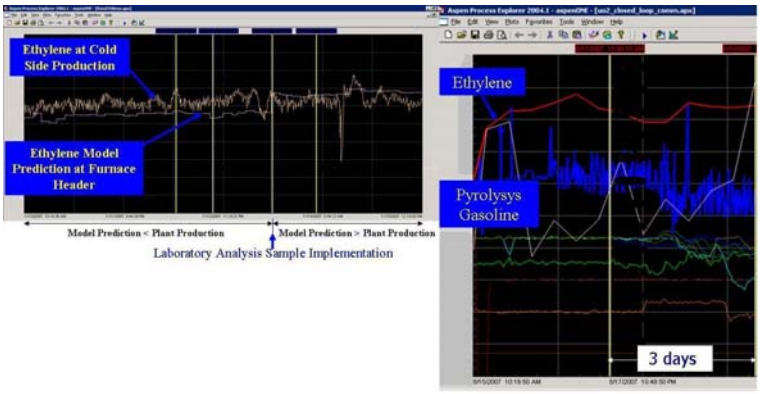


Figure 4 – Laboratory analysis scheme(a) and ethylene production target rates(b)

Figure 5 shows reduction of ethylene and methane losses in the demethanizer column. In this particular case, the RTO and the APC systems were able to reduce the methane losses by 43% and ethylene losses by 59%.

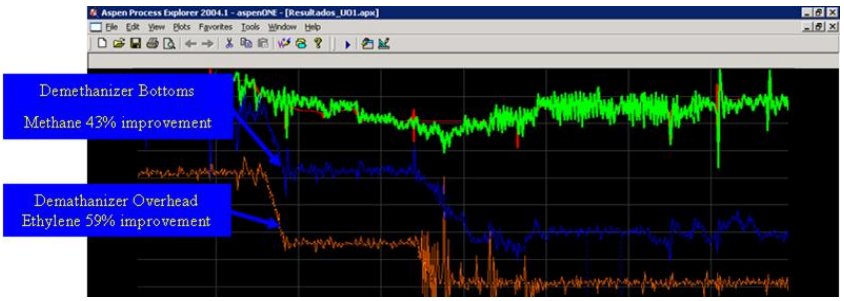


Figure 5 – Demethanizer losses improvements

6. Conclusions

Results proves that the Sustain Activities applied to the RTO systems at Braskem Olefins plants brought several operational improvements in the plants and also helped in sustaining the value generated by the optimizer.

References

L.T. Biegler, V.M. Zavala, 2009, Large-Scale nonlinear programming using IPOT: An integrating framework for enterprise-wide dynamic optimization, Computers&Chemical Engineering, 33, 3, 575-582.

G. Guangying, Q. Xianliang, J. Bin, Li Xingang, 2008, Simulation of quench system of ethylene plant using real components, Chemical Engineering Research and Design, 86, 9, 1049-1052.

F. Gubitoso, J.M. Pinto, 2007, A planning model for the optimal production of a real-world ethylene plant, Chemical Engineering and Processing: Process Intesification, 46, 11, 1141-1150.

M. Mercangöz, F.J. Doyle III, 2008, Real-time optimization of the pulp mill benchmark problem, Computers&Chemical Engineering, 32, 4-5, 789-804.

M. W. M. Van Goethem, S. Barendregt, J. Grievink, J. A. Moulijn, P. J. T. Verheijen, 2008, Towards synthesis of an optimal thermal cracking reactor, Chemical Engineering Research and Design, 86, 7, 703-712.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Stochastic dynamic predictions using kriging for nanoparticle synthesis

Andres F. Hernandez,[a] Martha A. Grover[a]

[a] School of Chemical & Biomolecular Engineering, Georgia Institute of Technology. Atlanta, GA. 30332-0100. E-mail: martha.grover@chbe.gatech.edu

Abstract

Kriging is an empirical modeling approach that has been widely applied in engineering for the approximation of deterministic functions, due its flexibility and ability to interpolate observed data. Despite its statistical properties, kriging has not been developed to approximate stochastic functions or to describe the dynamics of systems with multiple outputs. Our paper proposes a methodology to construct approximate models for multivariate stochastic dynamic simulations using kriging, by combining ideas from design of experiments and dynamic systems modeling. We then apply the methodology in the prediction of a dynamic size distribution during the synthesis of nanoparticles.

Keywords: Kriging, Dynamic Systems Modeling, Stochastic, Nanoparticle

1. Introduction

One of the most challenging problems in nanoparticle synthesis is the controlled generation of a monodisperse size distribution while sustaining a high yield. Because of the small size scale, molecular simulations are needed for modeling the manufacturing process. Unfortunately, the computational time of these simulations is usually too high for tasks like process control, requiring the creation of approximated models.

Kriging (also known as Gaussian process regression model) has recently been applied in the field of dynamic systems modeling [1, 2]. It is possible to approximate the state variable dynamics by implementing a recursive state-space formulation of a kriging model. The construction of this recursive function requires enough information to map from one discrete time to the next discrete time. For complex dynamic systems, obtaining the information to describe this dynamic function usually is time consuming, which requires minimizing the number of simulations while the system information is maximized over the state space. This problem can be viewed as a sequential selection of simulated data [3].

2. Methodology

2.1. Kriging

Kriging is a modeling approach for generalized linear regression models which accounts for the correlation in the residuals between the regression model and the observations [4, 5]. Kriging formulates a local correlation between the residuals of the linear regression model as a function of the location of the model inputs. Assume there is a set S of input/output pairs $\left\{\left(\mathbf{x}^i, \mathrm{y}^i\right)\right\}$, where $\mathbf{x}^i \in \mathbb{R}^d$, $y^i \in \mathbb{R}$, i = 1… n. A usual

covariance function to model the correlation at short distances between samples in the set is [1]:

$$\mathrm{V}_{ij}\left(\mathbf{x}_i,\mathbf{x}_j\right)=\sigma_c^2\cdot\exp\left[\sum_{k=1}^{d}\frac{\left(\left|\mathrm{x}_{i,k}\text{ - }\mathrm{x}_{j,k}\right|\right)^2}{\theta_k^2}\right]+\sigma_u^2\cdot\delta_{ij} \tag{1}$$

where δ_{ij} is the Kronecker delta and σ_c^2, $\theta_1^2\ldots\theta_d^2$, σ_u^2 are the kriging parameters that control the features of the correlation. In particular, σ_u^2 is included to model the random noise component of the function to be fitted. By using this covariance function, a kriging prediction for a new input $\mathbf{x}_o$ is [6]:

$$\hat{\mathrm{y}}\left(\mathbf{x}_o,S\right)=\mathrm{h}\left(\mathbf{x}_o\right)^T\hat{\boldsymbol{\beta}}+\mathrm{V}^T\left(\mathbf{x}_o,S\right)\cdot\mathrm{V}^{-1}\left(S,S\right)\cdot\left(\mathbf{y}\text{ - }\mathrm{H}\left(S\right)\hat{\boldsymbol{\beta}}\right) \tag{2}$$

where h, H represent a set of *p* regression functions evaluated at the unknown input and the inputs in *S*, $\hat{\boldsymbol{\beta}}$ is the generalized least-squares estimator of the regression coefficients and $\mathbf{y}=\left[\mathrm{y}^1,\ldots,\mathrm{y}^n\right]^T$. The corresponding variance of the prediction is:

$$\sigma_{\hat{\mathrm{y}}}^2\left(\mathbf{x}_o\right)=\mathrm{V}\left(\mathbf{x}_o,\mathbf{x}_o\right)\text{ - }\left[\mathrm{h}\left(\mathbf{x}_o\right)^T\ \mathrm{V}^T\left(\mathbf{x}_o,S\right)\right]\begin{bmatrix}0 & \mathrm{H}^T\left(S\right)\\ \mathrm{H}\left(S\right) & \mathrm{V}\left(S,S\right)\end{bmatrix}^{-1}\begin{bmatrix}\mathrm{h}\left(\mathbf{x}_o\right)\\ \mathrm{V}\left(\mathbf{x}_o,S\right)\end{bmatrix} \tag{3}$$

Frequently, the kriging parameters are obtained by maximizing the likelihood prediction over the multivariate normal distribution of the outputs in the set *S*.

2.2. Application of kriging in dynamic systems modeling

An approximated model for dynamic systems can be written based on the kriging structure presented in the previous section. A kriging model can be used as an autoregressive model where inputs and outputs are fed back and employed as regressors to predict a current output [7]. A recursive state-space formulation for the state variables $\mathbf{w}\in\mathbb{R}^m$ and control inputs **u** can be written as [2]:

$$\begin{gathered}\mathbf{w}\left(k+1\right)=\hat{\mathrm{f}}\left[\mathbf{x}_o\left(k\right),\ S\right]\qquad \mathrm{k}=0,\ 1,\ 2,\ \ldots\\ \mathbf{x}_o\left(k\right)=\left[\mathbf{w}\left(k\right),\ \mathbf{u}\left(k\right)\right]\\ t=k\cdot\Delta t\\ S=\left\{\left(\mathbf{x}^i,\ \mathrm{y}^i\right)\right\},\ \mathrm{y}^i=\mathbf{w}^i\left(k+1\right)=\mathrm{f}\left[\mathbf{x}^i\left(k\right)\right]\end{gathered} \tag{4}$$

where $\hat{\mathrm{f}}$ is represented by the kriging model in Eq. (2). Our approximated model employs the concept of storage and retrieval of information [8], in which pre-computed function evaluations are stored and then used to drive an equation-free prediction of the system dynamics [9]. From the stored data, a subset of information is extracted in the set *S* to approximate the dynamics at a new value of **w** and **u** in a one-step-ahead

prediction. Notice that $\mathbf{x}$ belongs to a higher dimension than the state-space $\mathbf{w}$, ($\mathbf{x} \in \mathbb{R}^d$, $\mathbf{w} \in \mathbb{R}^m$ and $d \geq m$) because the variables in the kriging model should also contain control inputs. By taking the approximated value from the previous step, our model uses the set S to approximate the dynamic behavior. Due to the recursive nature of Eq. (4), error will propagate from one step to the next. However, it is possible to formulate a correction in the prediction using a truncated Taylor-Series expansion [10] that combines error estimation in the kriging parameters using the Fisher Information Matrix of the likelihood function, along with error estimates from the previous kriging prediction via Eq. (3).

2.3. Sequential Design and Analysis of Computer Experiments (Sequential DACE)

A sequential DACE design was used to improve the construction of the reference vectors in our methodology. This sequential DACE was implemented based on the variance of the kriging prediction, Eq. (3), as an indicator of regions where our approximated model is more uncertain [11].

To create an approximated model with multiple inputs and outputs, one kriging model is built for the prediction of each output variable using maximum likelihood estimation to define kriging parameters. Since the multivariate model is built with several kriging models, the uncertainty in all models should be considered to define new simulations. It is proposed as a measure of uncertainty in the model, the sum of the kriging variance over the normalized state variables. The algorithm continues until a satisfactory level of prediction variance is reached. Figure 1 shows a schematic of our algorithm.

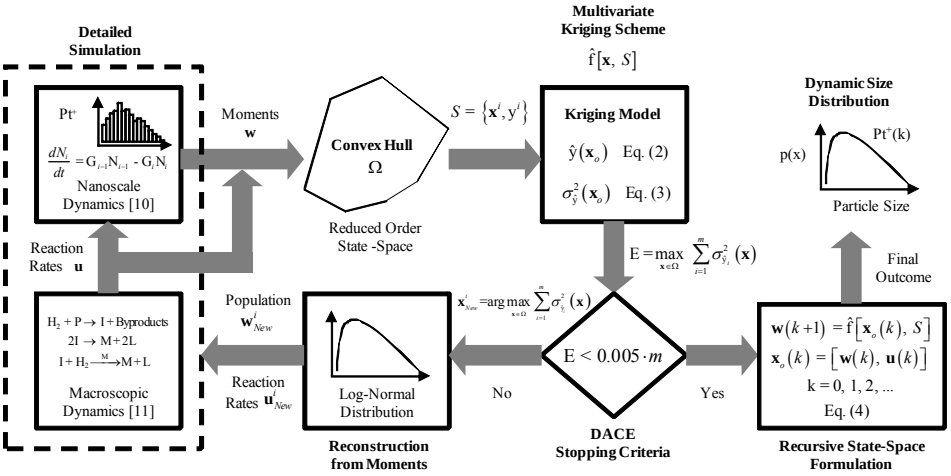


Figure 1. Sequential DACE design to approximate nanoparticle size distribution using a multivariate kriging scheme.

In this paper, we describe our detailed simulations as an ODE system that describes single adatom processes [12] for platinum deposition on porous alumina under supercritical CO_2 conditions [13]. The ODE system models the macroscopic dynamics of the hydrogen and precursor concentrations as well as the nanoscale dynamics of the nanoparticle distribution, which has a white noise term to generate stochastic behaviour. Our methodology creates a reduced state using the first three moments of the nanoparticle distribution, along with the nucleation and growth rates of the macroscopic

model to make the kriging model practical. Detailed simulations are collected for several macroscopic process settings of initial concentrations over a predetermined time period. The convex hull of these simulations defines the reduced state space.

From the convex hull, a set S of sampled information is selected to build the multivariate kriging scheme, and the uncertainty prediction is evaluated. Then, the sequential design of the stochastic simulations is applied over the convex hull closing the loop with the detailed simulations, assuming that the nanoparticle distribution is approximated by a log-normal distribution with corresponding raw moments calculated from the approximated model. The new sample point is collected at the point $\mathbf{x}$ of maximum uncertainty E. Our multivariate approximated model is rebuilt incorporating the new information in the set S, guiding the sequential design procedure. After the set S has been built, dynamic size distributions are obtained using the final approximated model for different macroscopic process settings.

3. Results and Discussion

Figure 2 presents the reconstructed nanoparticle size distributions from the prediction of the first three moments assuming that the nanoparticle distribution is log-normal. The sequential DACE procedure improves the prediction of the nanoparticle distribution by selecting additional sample points usually located close to the boundaries of the sampling space. The final kriging model uses a total of 23 sample points that creates a good approximation of the location of the peak and the width of the distribution. Our improved kriging model also shows a satisfactory result in the prediction of the long-term behavior of the nanoparticle distribution.

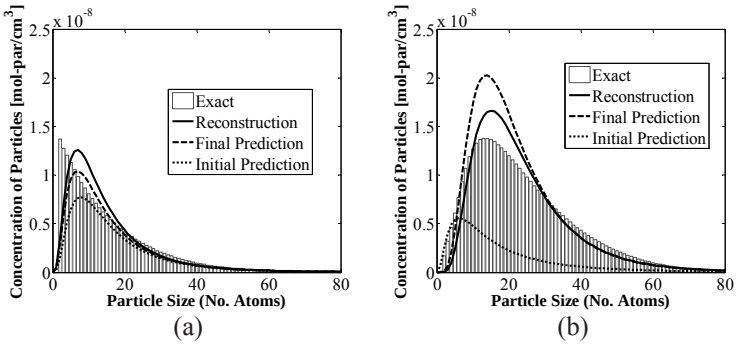


Figure 2. Reconstructed nanoparticle size distributions using the first three moments predicted by the kriging model. (a) Distributions at $t = 3$ *sec*. (b) Distributions at $t = 60$ *sec*. Initial Concentrations: $H_2 = 5 \times 10^{-5}$ *mol/cm*^{3}, *Precursor* $= 1 \times 10^{-5}$ *mol/cm*^{3}.

Comparing the nanoparticle distribution prediction made at $t = 3$ *sec* and $t = 60$ *sec*, the final kriging model exhibits a more accurate prediction over time than the initial kriging model. The effect of the error propagation and the uncertainty in the prediction can be seen in Figure 3, where the dynamics of the first three moments are shown. The addition of new sample points improves the prediction of all moments in the entire time frame of interest, showing the robustness of the approximated model. Some ideas to improve the performance of our model include the selection of different covariance and regression functions in the kriging expression or a better selection of the reduced state variables of

the system (increasing the number of moments extracted from the distribution or adding some macroscopic variables like concentration profiles).

The implemented sequential DACE design uses the prediction uncertainty of the model as a measurement in the selection of new sample points. Since the kriging model predicts exactly the information of the actual sample points, the prediction variance is equal to zero at those points. This situation creates several local maxima, most of them located at the boundaries of the reduced state space. We improve the chance to find the global maximum by randomizing the initial value of the optimization problem. 25 random points in the reduced state space are used to evaluate the objective function prior the optimization and the initial value is selected by the maximum value of the function. The results of the optimization procedure of the sequential DACE design are shown in Figures 4 and 5.

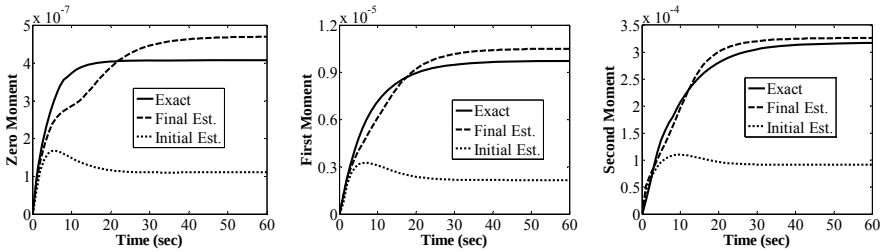


Figure 3. Dynamic approximation of the first three moments of the nanoparticle size distribution before and after use sequential DACE design on the approximated model. Initial Concentrations: $H_2 = 5 \times 10^{-5}$ *mol/cm*^{3}, *Precursor* $= 1 \times 10^{-5}$ *mol/cm*^{3}.

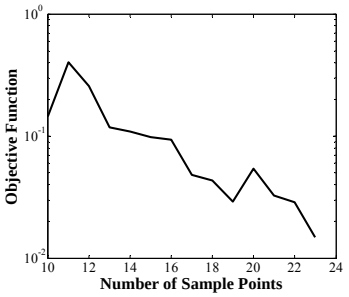


Figure 4. Improvement using sequential DACE design, as a function of the number of sample points used in the kriging models.

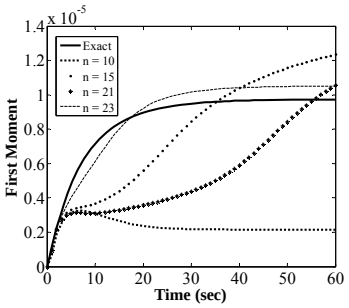


Figure 5. Improvement in the dynamic approximation of the first moment by using the sequential DACE design. Initial Concentrations: $H_2 = 5 \times 10^{-5}$ *mol/cm*^{3}, *Precursor* $= 1 \times 10^{-5}$ *mol/cm*^{3}.

The addition of the new 13 sample points decreases the value of the normalized kriging variance as well as the dynamic prediction of the first moment. The relevance of this fact is that the model can select by automatically where new and relevant information is needed from a full simulation in order to improve the approximation.

A global dynamic analysis was perform from 40 different macroscopic process settings, using 30 different multivariate kriging schemes and using the Euclidean distance

between the exact and the predicted trajectories at each time as an error measurement. After our sequential DACE design was implemented, the overall dynamic error in the trajectories decreases by 37%. A completely different sampling strategy must be developed when we are interested in a single nominal trajectory.

4. Conclusions and Future Work

This paper demonstrates the application of sequential DACE techniques to improve the accuracy of an approximate model from a stochastic simulation. By using a detailed simulation for nanoparticle synthesis, our model develops approximates of the nanoparticle size distribution obtained from the full simulations. The sequential DACE design based on the prediction uncertainty of the model shows a good performance in the selection of additional sample points to improve its accuracy.

One aspect that has not been investigated yet is the approximation of stiff systems. In our methodology, one of the keys to an accurate prediction is an appropriate selection of a time step that captures the dynamics of the system. A preliminary idea of how to approximate these systems will involve mixing information generated at different time steps that can be correlated at smaller time steps.

Since our approximate model requires recorded information to define the boundaries of the reduced state space for future sampling, the creation and management of databases is relevant. Ideas coming from data mining can be implemented in the construction of the preliminary reference vectors for our kriging model. By including these ideas, we expect to improve the construction of the convex hull.

References

[1]. Rasmussen, C.E. and C.K.I. Williams, *Gaussian Processes for Machine Learning*, ed. T.M. Press. 2006.

[2]. Hernandez, A.F. and M.G. Gallivan, *An exploratory study of discrete-time state-space models using kriging*, in *IEEE American Control Conference*. 2008: Seattle. p. 3993-3998.

[3]. Sacks, J., et al., 1989, "Design and analysis of computer experiments". *Statistical Science*. vol. 4, no. 4: pp. 409-435.

[4]. Cressie, N., *Statistics for Spatial Data*. 1993: Wiley Interscience.

[5]. Koehler, J.R. and A.B. Owen, *Computer Experiments*, in *Handbook of Statistics*, S. Ghosh and C.R. Rao, Editors. 1996, Elservier Science: New York. p. 261-308.

[6]. Martin, J.D. and T.W. Simpson, 2005, "Use of Kriging Models to Approximate Deterministic Computer Models". *AIAA Journal*. vol. 43, no. 4: pp. 853-863.

[7]. Azman, K. and J. Kocijan, 2007, "Application of gaussian processes for black-box modelling of biosystems". *ISA Transactions*. vol. 46: pp. 443-457.

[8]. Pope, S.B., 1997, "Computationally efficient implementation of combustion chemestry using in-situ adaptive tabulation". *Combustion Theory and Modelling*. vol. 1, no. 1: pp. 41-63.

[9]. Kevrekidis, I.G., C.W. Gear, and G. Hummer, 2004, "Equation-free: The computer-aided analysis of comptex multiscale systems ". *AIChE Journal*. vol. 50, no. 7: pp. 1346-1355.

[10]. Todini, E. and M. Ferraresi, 1996, "Influence of parameter estimation uncertainty in Kriging". *Journal of Hydrology*. vol. 175: pp. 555-566.

[11]. Kleijnen, J.P.C. and W.C.M.v. Beers, 2004, "Application-driven sequential designs for simulation experiments: Kriging metamodelling". *Journal of the Operational Research Society*. vol. 55: pp. 876-883.

[12]. Ratsch, C. and J.A. Venables, 2003, "Nucleation theory and the early stages of thin film growth". *Journal of Vaccum Science & Technology. A*. vol. 21, no. 5: pp. S96 - S109.

[13]. Gummalla, M., et al., 2004, "Multiscale Hybrid Modeling of Film Deposition Within Porous Substrates". *AIChE Journal*. vol. 50, no. 3: pp. 684-695.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Estimating the Normal Boiling Point of Organic Compounds Based on Elements and Chemical Bonds

Xia Li, Xiang Shu-Guang and Jia Xiao-Ping

The Hi-Tech Institute for Petroleum and Chemical Industry，Qingdao University of Science and Technology，Qingdao，266042，Shandong，China

Abstract

Based on elements and chemical bonds, a new method for estimating the normal boiling point of pure organic compounds from chemical structure was proposed. This method considers the contributions of interactions between elements and chemical bonds in the molecule. On the basis of 4,060 kinds of credible experimental data, 7 correlation equations are obtained by regression analysis. A mean absolute average deviation of 15.16K, less much than mostly used group contribution methods for alkyne, chloride derivative, iodine derivative, especially for alkane, bromide derivative, aromatic hydrocarbon, alicyclic hydrocarbon and sulphide organic compounds. Compared with the other group contributions methods, the proposed method in this paper showed significant improvements in accuracy and the ability to distinguish among isomers.

Keywords: the normal boiling point, elements, chemical bonds, group-contribution techniques

1. Introduction

The boiling point of organic compounds is of importance in the design of new organic compounds and in some industrial processes since they are the factors that primarily control solubility and vapour pressure [1-3].

Despite the enormous amount of available boiling point data, there are very few useful general means for quantitatively relating the boiling points of a compound to its chemical structure [4]. Most of the work in the prediction of the phase transition temperature of compounds has been focused mainly on boiling point estimation. Recently simple group contribution methods were proposed to predict the boiling points. It was found that the boiling temperature can be estimated simply by using molecular fragment values. Most of the methods were developed mainly from small data sets and can be used to predict only certain classes of organic compounds [5].

From the experience gained with the development of the previous group contribution methods, based on elements and chemical bonds, a new method for estimating the normal boiling point of pure organic compounds from chemical structure was proposed.

2. Physical Property Data Set

A data base containing boiling points, molecular weight, molecular descriptors, etc. was developed in Yaws' Handbook of Thermodynamic and Physical Properties of Chemical Compounds [6]. The data set for the normal boiling points (measured at 1 atm) consists of 4,060 compounds. The normal boiling points were credible experimental data. The compounds considered in this study were substituted aromatic including heterocyclic

compounds. The substituents used were the non-hydrogen bonding and single hydrogen bonding groups. These were methyl(CH_3), methylene (CH_2), cyano (CN), oxygen (O), hydroxyl(OH), halogens, nitro(NO_2), aldehyde (CHO),made(CONH2), keto (CO), carboxyl(COOH), Sulfur(S), sulfoxide(SO), mercapto (SH), thiocyanate (SCN), and isothiocyanate (NCS) groups.

3. Development of the new method

3.1 Molecular Descriptors

The elements and chemical bonds were regarded as basic contribution units in this method. Elements were divided into C, H, O, F, Cl, Br, I, N and S. Chemical bonds were divided into cyclic and non-cyclic. According to the different atoms, the chemical bonds were divided into C-C, C-H, C-O, cyclical C-C, cyclical C-O, and so on. Because the chemical bond in benzene was different from C-C and C=C, it was regarded as a special one. The list of structural groups for the new method (9 elements, 29 chemical bonds and 32 adjacent chemical bonds) were given in table 2.

3.2 Statistical Analysis

The Statistical Analysis of the data was performed using Matlab 7.0. The correlation coefficient and standard error were used as a measure of correlation for the equations developed. However, in selecting the best model emphasis was placed upon reducing the standard error.

3.3 Regression algorithm

From the experience gained with the development of the previous method, six linear regression of equation were employed.

$$T_b^k = m + \sum_j n_j \Delta T_{bj} \tag{1}$$

$$T_b = m + \frac{\sum_j n_j \Delta T_{bj}}{M^k} \tag{2}$$

$$T_b = m + \frac{\sum_j n_j \Delta T_{bj}}{M^k + l} \tag{3}$$

$$T_b = m + \frac{l + \sum_j n_j \Delta T_{bj}}{M^k} \tag{4}$$

$$T_b = m + k * (\sum_j n_j \Delta T_{bj})^l \tag{5}$$

$$T_b^l = m + \frac{\sum_j n_j \Delta T_{bj}}{M^k} \tag{6}$$

3.4 Correlation coefficient and group contributions

The correlation coefficients were found in Table 1. The contributions of elements, chemical bonds and the adjacent chemical bonds interaction were showed in Table 2, Table 3 and Table 4.

Table 1 The correlation coefficient for six linear regression of equations

correlations	m	k	l
Correlation 1	100.638	2.425	-
Correlation 2	10.892	0.604	-
Correlation 3	71.215	0.616	3.569
Correlation 4	37.069	0.619	489.408
Correlation 5	-25.668	27.548	0.397
Correlation 6	9.556	0.690	0.739

Table 2 The elements contribution for six linear regression of equations

elements	correlation 1	correlation 2	correlation 3	correlation 4	correlation 5	correlation 6
C	240320	782.910	839.850	841.670	114.640	222.400
H	28212	53.126	56.257	53.593	13.620	13.896
O	237280	756.010	823.740	817.720	115.180	216.640
F	95650	208.790	280.010	210.410	45.101	54.531
Cl	-155650	-109.020	-146.970	-126.550	-76.313	-16.131
Br	422880	1970.100	2032.800	2092.900	202.960	591.470
I	676750	3128.200	3223.100	3316.900	324.490	938.540
N	780110	1456.100	1619.800	1535.200	371.080	373.950
S	72825	351.300	379.940	402.480	36.894	103.750

Table 3 The contribution of chemical bonds for six linear regression of equation

chemical bonds	correlation 1	correlation 2	correlation 3	correlation 4	correlation 5	correlation 6
C-C	18652	48.224	52.788	52.958	8.469	12.609
C-O	-20133	46.858	33.214	45.012	-9.801	16.598
O-O	2458	47.524	24.630	38.489	0.039	12.299
O-H	718760	863.370	1069.400	933.160	343.120	197.610
C-H	-24453	-28.265	-31.964	-28.521	-11.603	-6.073
C-F	-134440	215.030	58.425	231.510	-60.153	97.711
C-Cl	615500	1952.800	2062.100	2085.700	298.710	559.620
C-Br	422880	1970.100	2032.800	2092.900	202.960	591.470
C-I	676750	3128.200	3223.100	3316.900	324.490	938.540
C-S	374130	934.160	1022.900	997.440	179.510	256.850
S-S	739810	1876.200	1998.800	1903.300	345.920	521.080
S-H	485890	992.620	1092.000	1019.700	232.760	267.250
C-N	-232150	-310.240	-364.760	-332.700	-111.490	-67.245
N-H	198200	272.680	326.360	303.670	96.577	61.182
N-N	-330090	-361.230	-400.540	-348.010	-156.670	-72.967
cyclical C-C	52679	82.322	92.760	84.927	25.001	20.488
cyclical C-O	90786	185.950	211.660	207.020	43.115	48.322
cyclical C-S	389500	1004.500	1101.000	1063.100	185.530	276.910
cyclical C-N	-36751	-39.512	-0.233	-8.676	-15.713	-4.859
C=C	60230	-3.186	4.808	7.272	32.892	-6.126
C=O	457480	614.510	729.390	663.240	218.650	145.840
C=S	186130	1085.300	1070.300	1198.800	97.746	327.880
C=N	-201780	-273.430	-253.300	-253.890	-90.481	-60.664
N=O	55093	187.760	197.250	197.880	26.707	53.176
cyclical C=C	-20851	-51.780	-49.945	-51.298	-8.349	-13.837
cyclical C=N	106730	134.760	176.370	143.740	49.635	31.476
C ≡ C	149130	25.339	46.749	17.972	75.249	-3.479
C ≡ N	305480	150.140	310.530	208.650	151.050	24.377
benzene	591560	684.240	824.250	703.750	282.020	154.100

Table 4 The contributions of the adjacent chemical bonds interaction for six linear regression of equation

the adjacent chemical bonds interaction	correlation 1	correlation 2	correlation 3	correlation 4	correlation 5	correlation 6
CH_3	-104880	-158.130	-189.040	-170.840	-46.805	-36.446
CH_2	11360	31.982	35.531	36.574	5.416	8.847
CC-CC_2	31033	26.431	30.312	15.806	14.014	6.647
CC-CC_3	80859	110.570	126.710	104.160	36.607	27.832
C_2C-CC_2	127440	193.840	224.360	196.320	57.775	48.234
C_2C-CC_3	217360	315.360	372.390	324.710	98.957	76.099
C_3C-CC_3	336220	459.980	552.920	477.440	153.490	107.650
C_2C=CC	30005	63.936	66.662	73.000	12.549	16.169
C_2C=CC_2	224030	283.310	355.600	310.510	103.170	64.649
H_2C=CC	-101830	-151.200	-177.810	-140.520	-45.310	-36.194
H_2C=CC_2	-46540	-87.804	-102.000	-75.157	-20.609	-21.663
CRC（ortho）	31212	46.820	61.002	49.555	13.634	10.994
CRC（para）	20961	38.531	46.535	40.560	9.224	9.819
CRO（ortho）	-126000	-98.122	-125.810	-110.520	-61.269	-19.466
CRO（para）	132790	180.530	222.410	191.660	62.335	41.568
ORO（ortho）	19101	19.490	16.704	14.063	7.779	5.701
ORO（para）	221950	137.670	195.000	164.990	105.510	26.854
CRN（ortho）	30320	221.410	239.060	230.780	13.143	62.553
CRN（para）	64651	237.440	286.990	258.320	28.161	60.186
CRS（ortho）	-166220	-126.090	-129.820	-107.540	-79.427	-24.883
CRS（para）	-129230	-96.662	-92.418	-77.653	-61.246	-18.962
NRN（ortho）	777570	886.490	1094.600	964.550	367.910	198.390
NRN（para）	1017700	1042.800	1253.700	1137.600	483.930	225.640
CHO	6964	-19.275	-3.241	-15.799	6.983	-4.160
COO	-20083	-6.329	-41.545	-35.664	-11.736	3.467
NH_2	-13300	-105.410	-64.020	-86.257	-3.547	-23.710
NO_2	27546	93.881	98.624	98.941	13.353	26.588
triatomic ring	-167720	-353.690	-390.680	-274.480	-67.937	-94.181
four-membered ring	-131290	-338.550	-362.300	-276.710	-52.184	-91.562
five-membered ring	3812	-126.410	-129.950	-124.380	6.280	-37.263
six-membered ring	57043	19.234	37.358	21.224	29.240	1.195
sever-membered ring	30570	-91.893	-64.870	-79.766	20.197	-28.648

Note: R: ring structure

3.5 Results and discussion

We will compare the above results with those from Joback and Reid (JR) [7], Devotta and Pendyala (DP) [8], Constantinuou and Gani (CG) [9], Marrero-Moreion and Pardillo (MP) [10]. All methods have in common the fact that they only require the knowledge about the molecular structure and therefore are comparable. The probability of a prediction failure (extreme deviation between experimental and estimated value) was chosen as the most important criterion for the reliability of a model.

The first correlation proposed here gives a mean absolute average deviation of 15.88K for 4060 compounds (second correlation: 15.89K (4060), third correlation: 15.16K (4060), fourth correlation: 15.34K (4060), fifth correlation: 15.78K (4060), sixth correlation: 16.47K (4060), JR: 33.45K (4060), DP: 33.37K (4060), CG: 16.43K (3893), MP: 17.02K (3549)). This means that it combines the lowest deviation with the broadest range of applicability.

In order to test the predictive capability of the method, experimental normal boiling temperatures for 10 components, not in the database used for regression, were compared with the predicted values. The results showed that the method proposed in this paper had significant improvements in accuracy.

To test the predictive capability of the method to differentiate the same kind of isomers, a detailed procedure for the estimation of T_b is given in Table 6 for heptane

isomers. The results showed that the method proposed in this paper had the ability to distinguish among isomers.

Table 5 Comparison of the method proposed for compounds, not used for regression

methods	N	AAE
Correlation 1	10	2.78
Correlation 2	10	1.91
Correlation 3	10	1.74
Correlation 4	10	1.84
Correlation 5	10	2.74
Correlation 6	10	1.77
JR	10	5.83
DP	10	5.90
CG	10	3.96
MP	10	3.17

Table 6 Comparison of the method proposed for heptane isomers

Compound name	$T_b^{\exp}$	T_b^{cal}	AAE	AAPE
heptane	371.58	371.01	-0.57	0.15
3-methyl-hexane	365	365.05	0.05	0.01
2-methyl-hexane	363.2	362.18	-1.02	0.28
3,3 - dimethyl - pentane	359.21	363.52	4.31	1.20
2,2 - dimethyl - pentane	352.34	355.92	3.58	1.02
2,4 - dimethyl - pentane	353.64	356	2.36	0.67
2,3 - dimethyl - pentane	362.93	362.16	-0.77	0.21
3-ethyl-pentane	366.62	367.89	1.27	0.35
2,2,3 -trimethyl- butane	354.03	356.77	2.74	0.77

3.6 Examples

To illustrate the application of the proposed method (using correlation 2), the results for the estimation of T_b is given in Table 7 for 1, 3-dihydroxy-4-methylbenzene.

Table 7 The estimation result of 1, 3-dihydroxy-4-methylbenzene

Molecular Descriptors	n_i	ΔT_{bi}	M	$\sum_i n_i \Delta T_{bi}$	T_b^{cal}	$T_b^{\exp}$	AAE	AAPE
C	7	782.910						
H	8	53.126						
O	2	756.010						
C-C	1	48.224						
C-H	6	-28.265						
C-O	2	46.858	124.139	9682.146	537.21	543.15	5.94	1.09
O-H	2	863.370						
Benzene	1	684.240						
CH_3	1	158.130						
CRO(ortho)	1	-98.122						
ORO(para)	1	137.670						

4. Conclusions

Based on elements and chemical bonds, a new method for estimating the normal boiling point of pure organic compounds was developed. The predictions are based exclusively on the molecular structure of the compound. The first correlation proposed in paper gives a mean absolute average deviation of 2.78K (second correlation: 1.91K, third correlation: 1.74K, fourth correlation: 1.84K, fifth correlation: 2.74K, sixth correlation: 1.77K, JR: 5.83K, DP: 5.90K, CG: 3.96K, MP: 3.17K) for 10 components, not in the database used for regression. The results of the new method were in most cases far more accurate than previous methods and had the ability to distinguish among isomers.

List of symbols

AAE	a mean absolute average deviation (K)
AAPE	a mean relative average deviation (%)
m , k, l	correlation coefficient
M	molecular weight
N	number of compounds
n_i	number of group i in compounds
n_j	number of group j in compounds
T_b	the normal boiling point of compounds (K)
$T_b^{\exp}$	the experimental normal boiling temperature (K)
T_b^{cal}	the predicted normal boiling temperature (K)
ΔT_b	the contribution of elements or chemical bonds (K)
ΔT_{bj}	he contribution of group j in compounds (K)

References

Pardillo-fontdevila E., R. Gonzalez-Rubio. 1997, A Group-interaction contribution approach for the estimation of chemical-physico properties of branched isomers, *Chem. Eng. Commun*, 163:245.

Reid R.C， Prausnitz J.M., Poling, B.E., 1987, The Properties of Gases and liquids, 4th ed., McGraw-Hill, NewYork.

Lohninger H., 1993, Evaluation of Neural Networks based on radial basis functions and their application to the prediction of boiling points from structural parameters. *J. Chem. Inf. Comput. Sci.* 33:736.

Cordes W., Rarey J, 2002, A New Method for the Estimation of the Normal Boiling Point of Non-electrolyte organic compounds. *Fluid Phase Equilib*, 201, 409.

Nannoolal Y., Rarey J., Ramjugernath D., Cordes W., 2004, Estimation of Pure Component Properties, Part 1: Estimation of the Normal Boiling Point of Non-Electrolyte Organic Compounds via Group Contributions and Group Interactions . *Fluid Phase Equilib*, 226, 45.

Yaws Carl L. 2003, Yaws' Handbook of Thermodynamic and Physical Properties of Chemical Compounds. Knovel.

Joback KG., Reid R., 1987, Estimation of pure component properties from Goup-contributions, *Chem. Eng. Commun.* 57:233.

Devotta S., Pendyala VR, 1992， Modified Joback group contribution method for normal boiling point of aliphatic Halogenated compounds. *Ind Eng Chem Res*, 31:2042.

Constantinous L., Gani R., 1994, New group contribution method for estimating properties of pure compounds, *AlChE J.* 40:1697.

Marrero-Morejon J., Pardillo-Fontdevila E., 1999, Estimation of pure compound properties using group-interaction contributions, *AlChE J.* 145:615.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Sequential design of dynamic experiments in modeling for optimization of biological processes

Mariano Cristaldi[a], Ricardo Grau[a], Ernesto Martínez[b]

[a]*INTEC (CONICET-UNL), Guemes 3450, Santa Fe 3000, Argentina*
[b]*INGAR (CONICET-UTN), Avellaneda 3657, Santa Fe S3002 GJC, Argentina*

Abstract

Finding optimal operating conditions fast with a scarce budget of experimental runs is a key problem to speed up the development and scaling up of innovative bioprocesses. A methodology for model-based design of dynamic experiments in modeling for optimization is proposed and successfully applied to the optimization of a fed-batch bioreactor related to the production of r-interleukin-11 whose DNA has been cloned in an *E. coli* strain. A library of tendency models is used to increasingly bias bioreactor operating conditions towards an optimum. Parametric uncertainty of tendency models is iteratively reduced using Global Sensitivity Analysis (GSA). At each iteration, the 'most informative' tendency model is used for designining the next dynamic experiment. Model selection is based on minimizing an error measure which separates parametric uncertainty from structural errors to trade-off exploration with exploitation.

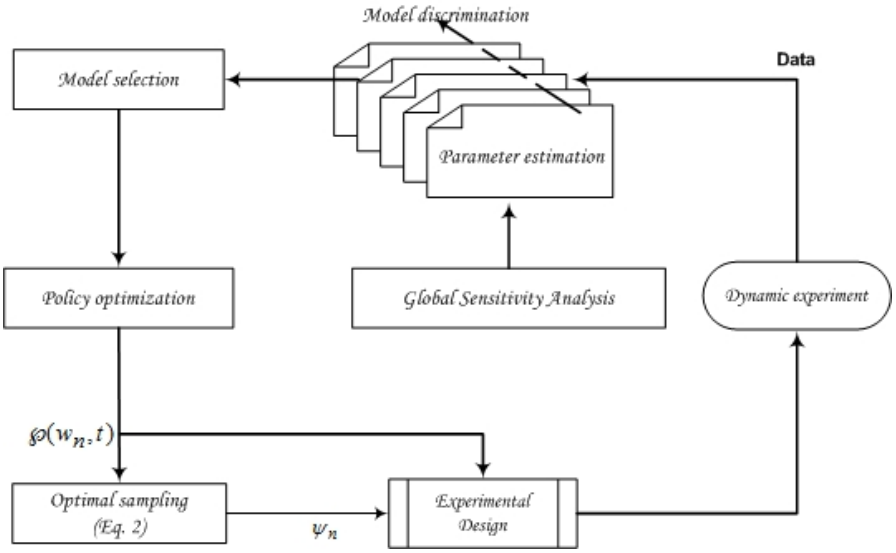
Keywords: Modeling, optimization, experimental design, dynamic experiments.

1. Introduction

Findings in the 1950s that DNA is the molecule that encodes proteins, which in turn controls all the cellular processes including metabolic pathways, have provided the impetus for the biotechnology era (Walsh, 2007) in the development of high valued-added pharmaceuticals such as insulin, humanized antibodies, interferons or interleukins along with biofuels (Fischer *et al*., 2008) and new materials (e.g. biodegradable polymers). The recombinant microorganism is typically grown in a fed-batch bioreactor to high cell concentration and then expression of the heterologous protein is triggered so as to obtain considerable quantities of the target product (Cooney, 1983). In a fed-batch culture, the feed rate of the carbon source, usually glucose, must be manipulated in order to restrict overflow metabolism and glucose repression. To this aim, model-based optimization of a bioreactor operating condition seems to be the safe and economic approach to resort with. Application of Pontryagin's maximum principle for fed-batch bioreactor optimization has been studied by several researchers (Mahadevan and Doyle III, 2003). However, most of these optimization methodologies for bioprocess scaling up and productivity inprovement have not been widely adopted for industrial use since the perfect model assumption is far from realistic and bioreactor behavior is quite often deviant from model predictions. Also, relevant measurement are sparse and delayed. Considering the large uncertainty and poor reproducibility in novel bioprocesses along with metabolic regulation, the development of an accurate mathematical model of bioreactor dynamics is a costly and very difficult undertaking. A better approach in innovative bioprocesses is to improve the operating policy by resorting to tendency models (Visser *et al.*, 2000) for designing optimally informative experiments which iteratively reduce the performance prediction uncertainty (Martínez and Wilson, 2003).

2. Modeling for optimization

At each iteration, the core idea of "modeling for optimization" is to select from a library of tendency models the one which allows computing inputs that increasingly improve the operating policy and bias data gathering accordingly. To this aim, a dynamic experiment is designed around the current policy, and optimal sampling times are calculated so as to maximize information content regarding performance improvement. The experiment is carried out and new data are collected. Based on incoming data the sub-set of model parameters for each tendency model are re-estimated which selectively reduces its parametric uncertainty. Based on total modeling errors a tendency model is selected for policy re-optimization. With the new input policy, a new iteration begins. The identification-optimization cycle is continued until no performance improvement is obtained and the input policy converges. Fig. 1 provides an overall picture of the proposed methodology.



Fig, 1. Modeling for optimization using designed dynamic experiments

2.1. Problem statement

In what follows let's assume that the dynamic behavior of the bioreactor under study may be modeled alternatively by a library of ℓ tendency models

$$\frac{dx_i}{dt} = f(x_i(t), \wp(w,t), \theta, t) \;\; 0 \le t \le t_f, \quad x_i(0): \text{given}, i = 1,2,...,\ell \qquad (1)$$

and the optimization objective to be maximized iteratively via experimental runs is

$$J(w) = h(x(t_f)) + \int_0^{t_f} g(x(t), \wp(w,t), \theta, t)dt \qquad (2)$$

where $x(t)$ is an n_s-dimensional vector of time dependent state variables, w is an m-dimensional vector of parameters for the input policy $\wp$, θ is a p-dimensional vector of model parameters and t_f is the final time of a batch run. The function g is the instantaneous reward function along the state trajectory $x(t)$ defined by a given policy parameterization whereas the function h is the specific reward for the final state of the batch run when using the input policy $\wp(w,t)$.

2.2. Model selection

Model selection is based on distinguishing between parametric uncertainty and structural errors in performance prediction using tendency models (see Fig. 2 for details). For a given tendency model and a plausible realization of its parameters, the corresponding simulated trajectory of the process performance index $\tilde{J}_i, i = 1,2,..., n_{sp}$. At the *ith* sampling point, a sample average $\bar{J}_i$ of n different model parameterizations can be used to characterize the *parametric uncertainty* for the tendency model as follows (Asprey and Machietto, 2002; Chen and Asprey, 2003):

$$\varepsilon_{J_i \to \bar{J}_i} == \frac{1}{n_{sp}\, n} tr\left[\left(\tilde{J}_{ij} - \bar{J}_i\right) W_{ij} \left(\tilde{J}_{ij} - \bar{J}_i\right)^T \right]; i = 1,2,..., n_{sp};\ j = 1,2,..., n \qquad (3)$$

As a measure of *structural errors* inherent to a given tendency model, the average performance trajectory $\bar{J}_i$ is compared to the actual (observed) trajectory J_i

$$\varepsilon_{\bar{J} \to J} = \frac{1}{n_{sp}} tr\left[\left(\bar{J}_i - J_i\right)^T W_{ii} \left(\bar{J}_i - J_i\right) \right]; i = 1,2,..., n_{sp}; W_{ii} :\ \textit{weighting matrix} \qquad (4)$$

where n_{sp} is the number of sampling points. The total error of a model is expressed as the sum of parametric uncertainty in Eq. (3) plus its structural error defined as it is shown in Eq. (4). For policy optimization, model selection in each iteration is based on choosing the tendency model from the model library whose total error is the lowest. More elaborated strategies for model selection can also be developed. For example, initially model selection may emphasize reducing parametric uncertainty and as more data are gathered model selection is more based on structural errors.

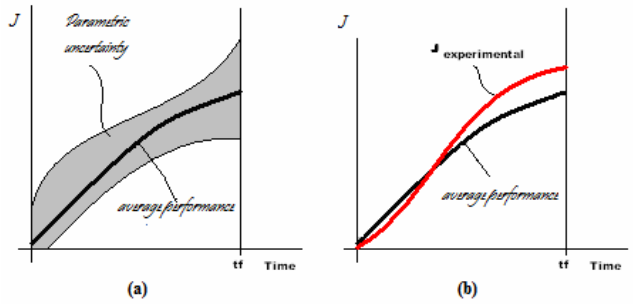


Fig. 2. Model selection based on total error: (a) Parametric uncertainty; (b) structural errors.

2.3. Optimal sampling

For a given policy in the current policy iteration of Fig. 1 optimal sampling times ψ^{opt} along a batch run must be calculated so as to bring new information to selectively reduce parametric uncertainty which affect the most the value estimation of the performance index J trajectory. Assuming model parameters are set to $\hat{\theta}$ and the

current version of the optimal policy is $\wp(w,t)$, the issue of optimal sampling is related to calculating at which sampling times $\psi^{opt} \in \Psi$ in a dynamic experiment the values of measured process variables are most informative in modeling for optimization assuming that the policy evaluation step should narrow down the uncertainty about the optimal input. To this end, the following optimization problem is solved:

$$\psi^{opt} = \max_{\psi \in \Psi} \det\left|M(\hat{\theta}, \wp(w_n,t), \psi)\right| , \; M = Q^T Q , \; Q = \begin{pmatrix} Si_{11} & \cdots & Si_{1n} \\ \vdots & \ddots & \vdots \\ Si_{m1} & \cdots & Si_{mn} \end{pmatrix} \qquad (5)$$

where each entry of the matrix Q , Si_{ij}, measures the sensitivity of the performance index $J(w)$ at the *i-th* sampling time with respect to *j-th* parameter of the operating policy. To calculate Si_{ij}, Global Sensitivity Analysis (GSA) is proposed (Saltelli et al., 2006). GSA is a variance-based technique that decomposes model outputs variability as a combination of uncertainty intervals for each independent input factor and its interactions with other factors using Monte Carlo sampling techniques.

3. Case study

To illustrate the proposed methodology results obtained in the optimization of fed-batch fermentation process for the recombinant protein **rIL-11** using a genetically modified *E. coli* strain are presented. Production of recombinant proteins in *E. coli* has been widely applied in both laboratory research and bioproduct manufacturing since this microorganism is considered a reliable source of proteins. This method may achieve profitable mass productivity due to high density cell growth and fast product formation. A structured kinetic model proposed by Tang *et al.* (2007) which describes state variables trajectories such as: biomass (X), substrate (S), intracellular recombinant protein concentration (P) will be used as an *in silico* bioreactor to generate the required data in the modeling for optimization approach. Four unstructured (tendency) models which differ in their biomass growth kinetics are used as guidelines for policy optimization so that the mismatch between the "real" bioprocess and alternative models of the fed-batch bioreactor is accounted for by increasingly biasing data gathering. Also, the operation policy has been defined based on the substrate feeding rate and induction time t_{ind} as the main components subject to optimization, including the initial culture condition. The performance index $J(t)$ is related to the amount of recombinant protein obtained at the final time of production runs. Tendency model equations and their alternative biomass growth kinetics are:

$$\frac{dX}{dt} = \mu X; \qquad \frac{dS}{dt} = -\frac{\mu}{Y_{xs}} X - f(X,t) \; ; \; \frac{dP}{dt} = r_P - \mu P \qquad (6)$$

$$r_P = \begin{cases} 0, & t < t_{ind} \\ K_P^{\max}\left(\frac{S}{K_S + S}\right)\left[\frac{1}{1+\left(\frac{P}{KI_P}\right)^5}\right], & t \geq t_{ind} \end{cases}$$

$$\begin{aligned} &\textit{First order} : \mu = \mu_{\max} S, \; f(X,t) = 0 \\ &\textit{Monod} : \mu = \mu_{\max} \frac{S}{K_s + S}, \; f(X,t) = 0 \\ &\textit{Contois} : \mu = \mu_{\max} \frac{S}{K_X X + S}, \; f(X,t) = 0 \\ &\textit{Monod} : \mu = \mu_{\max} \frac{S}{K_s + S}, \; f(X,t) = m.X \end{aligned}$$

Based on experimental data provided by Tang et al. (2007), a rather rough estimation of each tendency model parameters was made and referred to as "initial values" in Table 1.

Due to the significant level of parametric uncertainty a ±50% confidence interval around these initial values for each parameter is assumed in the first policy optimization iteration. A uniform distribution over its confidence interval is assumed initially for each model parameter. Experimental data have been conveniently pre-treated to reduce significantly the signal-to-noise ratio and outliers are not present.

Table 1. Initial parameterizations of tendency models based on experimental data

Parameter	Unit	Model			
		1st Order	Monod	Contois	Maintenance
$\mu_{\max}$	h^{-1}	0.2000	0.6301	0.5607	0.5261
K_s	$g\ L^{-1}$	2.0184	1.4956	-	0.7190
Y_{xs}	$g_{biomass}\ g_{substrate}^{-1}$	0.3982	0.4506	0.4826	0.4464
$K_P^{\max}$	$g\ L^{-1}$	0.0759	0.0629	0.0557	0.0536
KI_p	$g\ (L\ h)^{-1}$	0.0877	0.0609	0.0627	0.0600
K_x	$g_{substrate}\ g_{biomass}^{-1}$	-	-	1.7291	-
m	h^{-1}	-	-	-	0.0100

At any time *t*, the input policy is defined by a vector *w* of parameters corresponding to two different degrees of freedom for process optimization. A subset of the policy parameters corresponds to inputs that can be modified only from run-to-run but are time-invariant in a given run such as the substrate feeding concentration, run duration or induction time. The remaining entries are parameters which are used here for describing the profile of time-varying controls such as the feeding rate. In the latter case, a key issue is the mathematical description to be used so as to provide ample room for different variability patterns within economic and safety constraints with a minimum number of independent parameters. Even though other profile functions (high-order polymonials, Gaussian Processes, etc.) can be used for shape flexibility with a small number of parameters, the following quadratic inverse polynomial is used hereafter:

Table 2. Sequential optimization of a E. coli culture for rIL-11

Parameter	Units	Initial Condition	1st iter	2nd iter	3rd iter	4th iter	5th iter	6th iter
Z	$L\ h^{-1}$	1	-	-	-	-	-	-
A	$L\ h^{-2}$	0	0.0544	0.0121	0.0911	0.1431	0.2385	0.2389
B	h^{-1}	-	$3\ 10^{-4}$	$3\ 10^{-4}$	$3\ 10^{-4}$	$3\ 10^{-4}$	$3\ 10^{-4}$	$3\ 10^{-4}$
C	h^{-2}	-	$3\ 10^{-4}$	$4\ 10^{-4}$	$8\ 10^{-4}$	0.0156	0.0222	0.0223
S_f	$g\ L^{-1}$	10	30	30	30	30	30	30
t_{feed}	h	6	0	0	3.19	4.05	5	5
t_{ind}	h	4	4	4	4	4	4	4
t_f	h	12	16	16	16	16	16	16
V_0	L	6	5.31	10	5	6.61	5	5
X_0	$g\ L^{-1}$	0.05	0.1	0.1	0.1	0.1	0.1	0.1
S_0	$g\ L^{-1}$	6	3.93	7	7	7	7	7
$J.V_f$	g	**1.60**	**6.06**	**3.75**	**7.15**	**6.40**	**7.22**	**7.22**

$$F_{in} = \begin{Bmatrix} 0 & t < t_0 \\ \dfrac{At}{1 + Bt + Ct^2} & t \geq t_0 \end{Bmatrix} \quad (7)$$

As it is shown in Table 2, despite the rough approximation of tendency models, they provide a very valuable guideline for fast optimization with a handfull of experiments. Based on their total errors, *Monod* and *Contois* kinetics are the most infomative for re-calculating the input policy for dynamic experiments (see Table 3 and Table 4 below).

Table 3. Total error in 1st iter.

Model	$E_{\hat{j}_i \to \langle j \rangle}$	$E_{\langle j \rangle \to J}$	E_{total}
1st Order	0.0044	0.0209	0.0254
Monod	$7.13\ 10^{-12}$	$6.23\ 10^{-5}$	**$6.23\ 10^{-5}$**
Contois	$1\ 10^{-4}$	$5.78\ 10^{-5}$	$1.58\ 10^{-4}$
Manteinance	1.1403	$6.35\ 10^{-5}$	1.1404

Table 4.Total model errors in 4th iter

$E_{\hat{j}_i \to \langle j \rangle}$	$E_{\langle j \rangle \to J}$	E_{total}
$6.41\ 10^{-7}$	0.0279	0.0279
$6.92\ 10^{-6}$	0.0011	0.0011
$1.65\ 10^{-5}$	$9.85\ 10^{-4}$	**0.0010**
$2.01\ 10^{-4}$	0.0045	0.0047

4. Final remarks

A systematic procedure is proposed for sequential design of dynamic experiments in modeling for optimization using a library of tendency models for safe exploration of alternative parameterizations of the input policy to improve operating conditions. At each iteration, model selection is based on the total model error which accounts separately for parametric uncertainty and structural errors. Since tendency models are initially plagued with uncertainty model selection using poorly estimated total errors makes possible to trade off exploitation with exploration which is instrumental for model-based optimization with imperfect models. Global sensitivity analysis has been used to formulate the optimal sampling in each experiment as an optimization problem.

References

S. Asprey, S. Machietto, 2002, Designing robust optimal dynamic experiments, *J. of Process Control,* 12, 545-556.

B. Chen, S. Asprey 2003, On the design of optimally informative dynamic experiments for model discrimination in multiresponse nonlinear situations, *Ind. Eng. Chem. Res*., 42, 1379-1390.

C. Cooney, 1983, Bioreactors: design and operation, *Science* 19, 728-740.

C. Fischer, D. Klein-Marcuschamer, G., Stephanopoulos, 2008, Selection and optimization of microbial hosts for biofuels production. *Metabolic Engineering*, (**in press**).

R. Mahadevan, F. Doyle III, 2003, On-Line optimization of recombinant product in a fed-batch bioreactor, *Biotechnol. Prog*., 19, 639-646.

E. Martínez, J. Wilson, 2003, Evolutionary optimization of batch process systems using imperfect models, *Proc. Indian Natn. Sci. Acad*., 403-428.

A. Saltelli., M. Ratto, S. Tarantola, F. Campolongo, 2006, Sensitivity Analysis Practices. Strategies for Model-Based Inference, *Reliability Engineering & System Safety*, 1109-1125.

S. Tang, J. Chen, Z. Zhang, 2007, Structured models for recombinant human interleukin-11 fermentation, *Biochemical Engineering J.*, 35, 210-217.

D. Visser, et al., (2000). Tendency Modeling:A New Approach to Obtain Simplified Kinetic Models of Metabolism, *Metabolic Engineering*, 252-275.

G. Walsh, 2007, *Pharmaceutical biotechnology: concepts and applications*, John Wiley & Sons Ltd, Chichester, England.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Technology Pipelines for Learning in Energy System Models

Bri-Mathias S. Hodge[a], Joseph F. Pekny[a,b], Gintaras V. Reklaitis[a]

[a]*School of Chemical Engineering, Purdue University, Forney Hall of Chemical Engineering, 480 Stadium Mall Drive, West Lafayette, IN 47907, USA*
[b]*e-Enterprise Center, Discovery Park, Purdue University, Gerald D. and Edna E. Mann Hall, 203 S. Martin Jischke Drive, West Lafayette, IN 47907, USA*

Abstract

Technological learning has been treated as an exogenous variable in most energy system models despite the possibility of conditions produced by model variables having significant effects on the development of technologies. When technological learning has been endogenously treated within energy system models, it has often been accomplished through the use of learning curves. By assigning all technological change to a single variables such as cumulative sales, or even including research spending in a two-factor learning curve, the ability to examine the effects of directed policy are severely limited.

Technology pipelines represent the potentials and limitations of each successive generation of a technology individually while using the aggregated improvements learning curves represent to progress through each development stage of the particular technology generation. While multiple generations of a technology may be under concurrent development, by treating each generation individually the true capabilities and limitations may be exposed. With the possibilities for a technology known, development stage dependent policies may be implemented for each generation, instead of a blanket program for the entire technology. The use of the technology pipelines can help to bring energy system models more accurate means of simulating the policies necessary for selecting and nurturing energy technologies until they reach commercial availability. A case study on solar photovoltaic technologies is provided as an example of the technology pipeline framework.

Keywords: Technological learning, Energy systems modeling, Learning curves

1. Introduction

Technological change will play a large role in the trajectory of future energy systems. Current technologies will evolve at different rates and new technologies will gain a foothold. These innovations will help determine the structure of future energy systems and, because they are path dependent, it is critical that they are represented within energy system models. This endogenous technological learning is the key to producing not only realistic future scenarios but also the path by which they may take to their future state.

Technological change, that is an increase in economic output without a corresponding increase in the economic inputs, has historically been treated as a constant or near constant progression in combined energy-economy models. This exogenous

representation is able to show the effects of change on the system, but does not account for how the change occurs. Implicit in this representation is the assumption that economic policies have no effect on research and development or on the adoption of new technologies. While this assumption may be quickly dismissed as simplistic, the economic description of technological change is much more difficult to formulate. The first step in this process is to realize that research and development must be treated as a fixed cost instead of as a raw material input into a product (Romer, 1990).

The basic idea of learning curves, that the cost of production for one unit of a product decreases as the total number of units produced increases, was first formulated more than 70 years ago (Wright, 1936). Though this idea of learning-by-doing has been successfully applied to many disparate industries (Yelle, 1979), it neglects to consider the effects of research and development explicitly, instead effectively attributing all improvements to economies of scale. Learning rates for a number of energy technologies have been calculated (McDonald and Schrattenholzer, 2001). The expansion of the learning curve idea to include a learning-by-searching variable (Kouvaritakis *et al.*, 2000) created the two factor learning curve and allows an explicit representation of the role research and development plays in technological change.

Integrating endogenous technical learning into energy system models has only become a subject of research attention comparatively recently. One of the first attempts at incorporating technological learning into an energy model was made by using investments made by the model as the only source of technical advancement (Messner, 1997) within the mixed integer programming MESSAGE III model. Concurrently the linear programming GENIE model was adapted to incorporate single factor learning curves as a measure for technological change for the study of electrical technology adoption (Mattsson and Wene, 1997). Similarly, single factor learning curves were used to endogenize technological learning in versions of the MARKAL (Loulou *et al.*, 2004) and ERIS (EC-TEEM, 2000) models that were used to study the effects of implementation of the Kyoto protocols on the electricity industry (Barreto, 2001). The variable elasticities used in two-factor learning curves have been formulated for a number of technologies and applied within the POLES (Criqui and Kouvaritakis, 2000) energy model. Additional two factor progress ratios for mature, evolving and emerging electricity technologies have also been calculated (Jamasb, 2007).

2. Research Pipelines

In order to more accurately represent the way new energy technologies are nurtured and developed a new formalism for their representation has been developed. Since one of the goals of long-term energy policy must be the selection and maturation of new technologies the problem can be likened to the management of a portfolio of new products candidates in the pharmaceutical industry (Blau *et al.*, 2004), a task that has already received significant research effort (Ding and Eliashberg, 2002), (Kavadias and Loch, 2004), (Rogers *et al.*, 2002).

As is common in the evolution of products in the pharmaceutical industry we will classify the progress of technologies into three stages. During the research stage a new means for accomplishing the goals specified for the technology are discovered and refined. It is during this first stage where the initial cost of the unit is defined and the initial capabilities of the technology are established. Production of the technology

during the research stage is limited to the laboratory scale. During the development stage work continues on advancing the capabilities of the technology but work also begins on improving the process by which the technology may be built. The economic realities of the technology often become more apparent during this second stage and technologies that have severe limitations that may hinder their ability to become economically competitive with existing technologies often fail at this stage. Production of the technology progresses during the development phase from the laboratory scale to the pilot plant scale. The final stage in the progress of the technology is the manufacturing stage. Here the production processes that were previously developed are refined further and new fabrication procedures may also be implemented.

In the widely-used punctuated equilibrium model of technological change learning occurs in two stages: short bursts of innovation followed by incremental progress. These revolutionary changes beget new technologies that create new markets and render previous technologies obsolete. While viewing technological change from this high level perspective is good for providing histories of technologies, it misses the significant changes that may occur within a technology by only representing the change that occurs within established technologies incrementally. The use of technological generations is necessary to differentiate between particular instances within a technology that have the same general capabilities but possess fundamental underlying differences in structure, material or design. Technological generations are meant to represent significant advances in performance within a technology regime (Lawless and Anderson, 1996). Therefore in the technology pipeline model of learning multiple generations of a technology may exist concurrently, each following the three evolutionary stages previously mentioned.

3. Example: Photovoltaic Solar Cell Technologies

Solar photovoltaics are an ideal choice to illustrate the concept of technology pipelines due to their well documented history and status as an emerging player in main stream electricity applications. Additionally, the specific generations of solar photovoltaics are much more well-defined than for many other energy technologies.

First generation photovoltaic technology is based on single junction cells made from crystalline silicon wafers. While modules based on these cells currently make up the bulk of the commercial market, their relatively low theoretical efficiency limit and the high cost of the composite materials and costly processing steps have combined to spur research on subsequent generations of photovoltaic technology that may have a lower floor cost. Thin film, or second generation, photovoltaic cells were conceived as a means of limiting the use of expensive raw materials in cell fabrication. A variety of exotic, and costly, materials are used in their production but the small amounts of material needed, as well as more economical processing methods, help to limit costs. Unfortunately the cheaper processing methods used for cadmium telluride, copper indium gallium selenide and amorphous silicon cells also produce more defects that limit cell efficiency, which ultimately raises the cost. Second generation technologies have already advanced to the production stages though they currently account for only one-tenth of solar cell production. Third generation cells attempt to retain the low cost approach of second generation cells but also increase efficiency through the use of tandem cells, multiple electron-hole pairs, hot carriers, multiband cells and thermophotonic devices (Green, 2002). Third generation technologies have only been

produced in limited quantities for specialty applications and are thus classified as at the beginning of the development phase within the technology pipeline framework.

In order to demonstrate the potential benefits of using the technology pipeline approach to examine policy level decisions a pair of scenarios for the worldwide solar photovoltaic industry was developed. Each generation of solar photovoltaic technology is assigned a two-factor learning curve that is used to calculate the cost of this particular instance of the technology. The detailed price data need for determining the individual learning by doing and learning by searching rates for second and third generation technologies is not yet public because they are still in early stages of commercialization. Therefore uniform learning rates were assumed for all three generations of the technology based on the longer history of the first generation technology. While this may over or under estimate the potential of the second and third generation technologies it allows the effects of the allowing generation specific policies to be examined in a more general manner. The learning by doing rate used is 18.4% and the learning by searching rate is 14.3% (Kobos *et al.*, 2006). Because the exact learning rates can have a major impact on determining the potential cost reductions for a technology it is important that the most accurate known values are used; those chosen are a good compromise of the varying values found in the literature. It is also important to note that the costs for the modules are based upon the reported market price, not the production costs.

In both the base case scenario and the generational policy scenario the photovoltaic industry's production capacity is assumed to grow at a rate of 15% per year for a period of ten years. While this may seem optimistic it is actually less than the actual rate of growth of 18% from 1993 to 1997 (Shah *et al.*, 1999) and that of 50% production growth rate of the industry in the USA from 2005 to 2007 (EIA, 2008). In both scenarios expansions in production capacity are split between the three generations based upon the current cost of the technology, with the lowest priced technology generation receiving a significantly larger share. Worldwide research and development spending is kept constant throughout the ten year period at $500 million per year, a figure that includes both industrial and governmental expenditures. The allocation of this research budget between the three generations of photovoltaic electricity generation technologies is what differentiates the two scenarios. In the base case scenario all funding is tied to the expansions in capacity; that is it follows the lead of the industry. In contrast, the generational policy scenario allocates the governmental portion of the funding based on the highest cost technology while the industrial portion of the funding follows the same rules as expansions in production capacity. This highest cost approach is predicated upon the assumption that those generations with the highest cost also have the most room for improvement.

The results of a discrete event simulation that models each of the two scenarios indicate that the implementation of generational specific policies can play a large role in the reducing the industry wide cost of a technology. The model is implemented as three states representing the stages of the technology generation. Any additions to research spending or production trigger a recalculation of the generational cost and a check on whether the generation should proceed to the next stage in development or remain in its current state. As may be seen in Figure 1, the allocation of additional research funding to the third generation produces a significant decrease in the module cost of the technology when compared to the base case. This result can also be inferred from the

structure of the learning curve formulation. Since the initial amount of research spending on the third generation technology is much smaller a doubling of the cumulative research spending, and hence significant cost reduction, requires less spending than an equivalent reduction for either of the more mature technology generations. This model assumes that the policy shift in research funding occurred before the start of the simulation time frame, as research spending does not translate into immediate discoveries. A typical time lag between research spending and its effect on the technology's learning curves is 2-3 years.

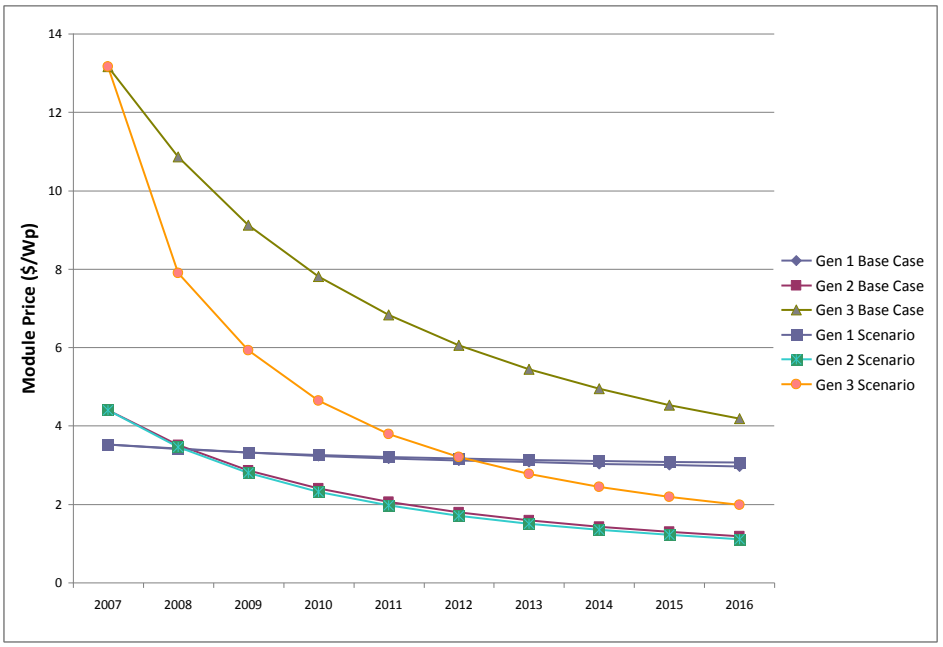


Figure 1: Solar photovoltaic module prices for the three technology generations in the base case and policy specific scenarios

Through the use of the research pipeline concept we have been able to examine how governmental and industrial policies that are focused on a particular segment of a technology may elucidate more information on the potentials of individual technology generations. Following the common single technology single learning curve model any changes in policy could only reproduce the potentials and limitations of the first generation photovoltaics that still dominate the industry. As second generation technologies have started to reach and surpass the cost levels of the first generation, a closer examination of the succeeding generations is necessary to illuminate the future path of the photovoltaic industry.

4. Conclusions

Energy system models have often treated technological learning as an exogenous variable. Government and industrial policies may play a large role in technological learning and thus models that aim to analysis these policies should treat technological learning endogenously. When endogenous technological learning has been considered, it has been in the form of industry-wide learning curves that prohibit the investigation of policy choices that focus on particular breeds of a technology. The technology pipeline

framework presented represents a decomposition of the learning curve concept often used to represent endogenous technological chance. An example from the field of solar photovoltaic electricity production has been used to illustrate how this decomposition may lead to the simulation of more detailed technology fostering policies.

References

L. Barreto, 2001, Technological Learning in Energy Optimisation Models and Deployment of Emerging Technologies, Swiss Federal Institute of Technology Zurich.

G. Blau, J. Pekny, V. Varma, P. Bunch, 2004, Managing a Portfolio of Interdependent New Product Candidates in the Pharmaceutical Industry, The Journal of Product Innovation Management, 21, 227.

P. Criqui, N. Kouvaritakis, 2000, World Energy Projections to 2030, International Journal of GLobal Energy Issues, 14, 1-4, 116.

M. Ding, J. Eliashberg, 2002, Structuring the New Product Development Pipeline, Management Science, 48, 3, 343.

EC-TEEM, 2000, Energy technology dynamics and advanced system model, International Journal of GLobal Energy Issues, 14, 1.

EIA, 2008, Solar Photovoltaic Cell/Module Manufacturing Activities 2007 (Department of Energy, Energy Information Administration).

M. Green, 2002, Third generation photovoltaics: solar cells for 2020 and beyond, Physica E, 14, 65.

T. Jamasb, 2007, Technical Change Theory and Learning Curves: Patterns of Progress in Electricity Generation Technologies, The Energy Journal, 28, 3.

S. Kavadias, C. Loch, 2004, *Project Selection Under Uncertainty: Dynamically Allocating Resources to Maximize Value*. F. Hillier, Ed., International Series in Operations Research & Management Science (Kluwer), pp.

P. Kobos, J. Erickson, T. Drennen, 2006, Technological learning and renewable energy costs: implications for US renewable energy policy, Energy Policy, 34, 1645.

N. Kouvaritakis, A. Soria, S. Isoard, 2000, Modelling energy technology dynamics: methodology for adaptive expectations models with learning by doing and learning by searching, International Journal of GLobal Energy Issues, 14, 1-4, 104.

M. Lawless, P. Anderson, 1996, Generational Technological Change: Effects of Innovation and Local Rivalry on Performance, Academy of Management Journal, 39, 5, 1185.

R. Loulou, G. Goldstein, K. Noble, 2004, Documentation for the MARKAL Family of Models (Energy Systems Technology Analysis Programme).

N. Mattsson, C.-O. Wene, 1997, Assessing New Energy Technologies Using an Energy System Model with Endogenized Experience Curves, International Journal of Energy Research, 21, 385.

A. McDonald, L. Schrattenholzer, 2001, Learning rates for energy technologies, Energy Policy, 29, 4, 255.

S. Messner, 1997, Endogenous Technological learning in an energy systems model, Journal of Evolutionary Economics, 7, 291.

M. Rogers, A. Gupta, C. Maranas, 2002, Real Options Based Analysis of Optimal Pharmaceutical Research and Development Portfolios, Industrial & Engineering Chemistry Research, 41, 6607.

P. Romer, 1990, Endogenous Technological Change, Journal of Political Economy, 98, 5.

A. Shah, P. Torres, R. Tscharner, N. Wyrsch, H. Keppner, 1999, Photovoltaic Technology: The Case for Thin-Film Solar Cells, Science, 285, 5428, 692.

T. Wright, 1936, Factors Affecting the Cost of Airplanes, Journal of Aeronautical Sciences, 3, 4, 122.

L. Yelle, 1979, The Learning Curve: Historical Review and Comprehensive Survey, Decision Sciences, 10, 2, 302.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Discrete Element Modeling for Multiscale Simulation of Aggregation Processes

Alexander Reinhold and Heiko Briesen

Max Planck Institute for Dynamics ofComplex Technical Systems, Sandtorstraße 1, 39106 Magdeburg, Germany.
Present Adress: Chair for Process Systems Engineering, Center of Life and Food Sciences Weihenstephan, TU München, 85350 Freising

Abstract

First steps of combining the favorable predictive capability of the discrete element method and the computational efficiency of the population balance approach are presented. Aggregation due to gravitational settling and shear induced aggregation are implemented and discussed as case studies. Results of the multiscale simulation show excellent results for 99 % of the total particle number.

Keywords: multiscale simulation, discrete element method, shear induced aggregation

1. Introduction

Population balance modeling is a commonly used tool to describe aggregation processes on a macroscopic level. Analytical aggregation rate kernels and suitable numerical discretization have to be used to obtain meaningful solutions of actual population balance models. However, a predictive simulation of aggregation is still not possible today due to the lack of predictive models for the aggregation rates.

The physics of aggregation are governed by many effects: particle mass fraction, particle shape, particle size distribution, spatial distribution of particles and the leading aggregation mechanism. For the derivation of well known aggregation rate kernels (e.g. the classical shear induced or Brownian aggregation rate kernel by Smouluchowski (1917)) very coarse assumptions have to be made in order to obtain an analytical solution: low particle mass fraction, ideal spherical particles, no dependency on particle size distribution, well mixed compartment and a single type of mechanism. With discrete element simulations, the Brownian aggregation rate kernel is reproduced for mass fractions lower than 0.1% (Trzeciak (2006), Heine and Pratsinis (2007)). However, it was also shown that Brownian aggregation rate kernels depend on the particle mass fraction and even depend on the specific particle size distribution.

Discrete element modeling of aggregation is capable of avoiding many modeling assumptions and provides more accurate and predictive models. However, discrete approaches are computationally expensive and therefore not immediately suited for a large population of particles. Equation-free multiscale modeling (Kevrekidis (2004)) provides a framework to combine the favorable predictive capability of discrete element simulation and the computational efficiency of population balances.

This paper discusses basic principles of the multiscale simulation of aggregation and focuses on two specific case studies: Shear induced aggregation and aggregation due to gravitational settling. The mandatory modeling assumptions for these aggregation rate

kernels are reconstructed in the discrete model to validate their accuracy. Additionally, the first results of multiscale simulations are presented.

2. Method

2.1. Population Balance

The particle population comprises of ideal spherical particles in a well-mixed compartment. The volume v (or radius $\lambda(v)$) is used as the only internal coordinate. Pure aggregation phenomena can be modeled by the population balance equation:

$$\frac{\partial \tilde{n}(t,v)}{\partial t} = \int_0^v r(\tilde{n}, v', v-v')dv' - \int_0^\infty r(\tilde{n}, v, v')dv' \quad , \tag{1}$$

where $\tilde{n}$ is the number-based particle size distribution and r are the aggregation rates. The aggregation rate measures the frequency of effective collisions per time and volume between particles of the sizes v and v'. With the assumption of binary aggregation, the aggregation rate kernel $\beta(v,v')$ is usually separated:

$$r(\tilde{n}, v, v') = \beta(v,v')\tilde{n}(v)\tilde{n}(v') \quad . \tag{2}$$

Available theoretically motivated aggregation rate kernels require the assumption of ideal spherical and homogenously distributed particles (Ramkrishna (2000)). Hence, the assumptions for the population balance in Eq. (1) are mandatory if analytical aggregation rate kernels are used. In this paper two case studies on aggregation are analyzed. Each one depends on a different set of physical parameters beside the dependency on particle radii. The shear induced aggregation rate kernel $\tilde{\beta}_{shear}$ depends on the shear rate γ. The aggregation rate kernel for aggregation due to gravitational settling $\tilde{\beta}_{grav}$ depends on the gravitational acceleration g, dynamic viscosity η, mass density of particles ρ_{part} and mass density of fluid ρ_{fluid}. The aggregation rate kernels for these cases can e.g. be found in Hounslow et al. (2001) and Ramkrishna (2000), respectively:

$$\tilde{\beta}_{shear}(v,v') = \frac{4\gamma}{3}\left(\lambda(v)+\lambda(v')\right)^3 \quad , \tag{3}$$

$$\tilde{\beta}_{grav}(v,v') = \frac{2\pi}{9}\frac{g\left|\rho_{part}-\rho_{fluid}\right|}{\eta}\left(\lambda(v)+\lambda(v')\right)^2\left|\lambda(v)^2-\lambda(v')^2\right| \quad . \tag{4}$$

2.2. Multiscale Framework

What happens, if the mandatory assumptions of spherical and homogeneously distributed particles are insufficient? Or assuming a dense particle population, the assumption of binary collisions becomes invalid. Analytical aggregation rate kernels and consequently analytical aggregation rates are no longer available.

Adopting the equation free multiscale framework (Kevrekidis et al. (2004)), this problem can be avoided. Instead of a closed analytical formulation of the problem, missing analytical equations are replaced by discrete element simulations. As discrete element simulations are very time consuming, the multiscale framework can profit from the time efficiency of the population balance equation plus the detail of modeling assumptions of the discrete element model.

The framework is based on the numerical solution of the population balance equation. Generally, any discretization method can be used to transform Eq. (1) into a set of ordinary differential equations. For this case study the fixed pivot technique presented by Kumar and Ramkrishna (1996) is used. The continuous particle size distribution $\tilde{n}$ is discretized in intervals i with a number of particles n_i in each interval. Each interval has a representative particle volume v_i. The resulting ordinary differential equations are integrated stepwise. For each step of the time integrator, aggregation rate kernels $\beta(v_i, v_j)$ are required for all pairs of representative particle sizes.[1]

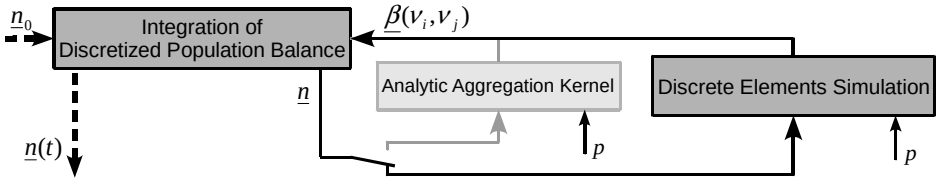


Figure 1: Scheme of multiscale framework

The basic scheme of the equation free multiscale framework adopted to aggregation processes is shown in Figure 1. Given the current state vector $\underline{n}$ and the required physical parameters p, the aggregation rate kernel can be calculated in two ways. If modeling assumptions allow usage of a given analytical aggregation rate kernel, the framework reduces to a closed analytic formulation on the macroscale level. If analytical expressions are leaking or invalid, statistical evaluation of discrete element simulations may be used instead.

2.3. Discrete Element Model

Using a discrete element model, the assumptions of ideal spherical particles and homogeneous spatial distribution are in principal not necessary. However, for the work presented here, these assumptions are exactly reconstructed to validate the framework with analytical aggregation rate kernels. The discrete element simulation is divided into three sections: initialization, simulation, statistical analysis.

2.3.1. Initialization

The discrete element simulation requires a number N of particles initialized with radii and position in the simulation volume of the size $V_{sim} = N / \sum_i n_i$. The position $\vec{x}_Q$ of each particle Q is equally distributed in the simulation volume. Trzeciak et. al (2006) and Heine and Pratsinis (2007) have shown that the aggregation rate kernel depends on the particle size distribution for Brownian aggregation. Hence, to construct a most flexible framework, the size distribution of the discrete particles must approximate the discretized particle distribution $\underline{n}$. The radius λ_Q for each particle Q is sampled from the probability distribution

$$P(\lambda_Q = \lambda(v_i)) = \frac{n_i}{\sum_i n_i} \tag{5}$$

to achieve this requirement.

[1] Note, that the actual choice of these representative particle sizes depend on the discretization of the inner coordinate.

2.3.2. Simulation

The derivation of the analytical aggregation rate kernels assumes constant particle velocities for shear and gravitational settling, respectively:

$$\vec{u}_{shear,Q} = \begin{pmatrix} 0 & \gamma & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{x}_Q \quad , \tag{6}$$

$$\vec{u}_{grav,Q} = \frac{2}{9} \frac{\lambda_Q{}^2 (\rho_{fluid} - \rho_{part})}{\eta} (0,1,0)^T \quad . \tag{7}$$

These velocities are integrated iteratively for small time steps τ. With each step, collisions between particles are checked. Each collision between two selected particles is counted once. If two particles collide again later, this collision is ignored. Collisions between two particles do not lead to an aggregation. Thus, the particle size distribution is conserved during the simulation.

2.3.3. Statistical Analysis

Given the number $C(\nu_i,\nu_j)$ of counted collisions between particle size classes i and j, the aggregation rate kernel can be calculated by :

$$\hat{\beta}(\nu_i,\nu_j) = C(\nu_i,\nu_j) \frac{V_{sim}{}^2}{T_{sim} V_{obsv}} \frac{1+\delta_{i,j}}{N_i (N_j - \delta_{i,j})} \frac{1}{1 - p_{init,(i,j)}} \quad , \tag{8}$$

where $\delta_{i,j}$ denotes the Kronecker delta and T_{sim} denotes the time span of the simulation. The observed volume V_{obsv} is slightly smaller than the simulation volume. The term $p_{init,(i,j)}$ is the probability of two particles fulfilling the collision condition just after initialization. A detailed description of these terms and the derivation of Eq. (8) is given by Reinhold (2008).

As the simulation is initialized randomly, the achieved aggregation rate kernel has a certain coefficient of variation. This coefficient of variation is approximated in Eq. (9) by assuming the number of counted collisions being normally distributed:

$$\theta_{\hat{\beta}(\nu_i,\nu_j)} = \frac{1}{\sqrt{\hat{\beta}(\nu_i,\nu_j)}} \sqrt{\frac{1+\delta_{i,j}}{T_{sim} N} \frac{\sum_k n_k}{n_i n_j}} \quad . \tag{9}$$

Eq. (9) can be used as a measure of quality for the estimates of the aggregation rate kernels. An increase of simulation time has the same effect on the coefficient of variation as the increase of the number of particles for the simulation. The accuracy of calculating small aggregation rate kernels becomes worse. Aggregation rate kernels for particles of the same size class have a coefficient of variation $\sqrt{2}$ times larger, than aggregation rate kernels for particles of different size classes. Additionally, the aggregation rate kernel for denser particle populations can be predicted more accurately.

3. Results

3.1. Discrete Element Model

To validate the implemented discrete element model for each case, a set of 10^5 stochastically independent simulations of $N = 6000$ particles and a time span of $T_{sim} = 0.5\,s$ were performed.

Particle sizes of $\lambda_1 = 4\,\mu m$ and $\lambda_2 = 8\,\mu m$ were initialized with particle densities of $n_1 = n_2 = 10^{13}\,m^{-3}$ for shear induced aggregation. The shear rate was $\gamma = 5\,s^{-1}$ and the time steps were limited to $\tau_{\max} = 2.5\,ms$. The mean aggregation rate kernel for two particles of size λ_1 was calculated and resulted in a relative error of 0.041% compared to the analytical aggregation rate kernel Eq. (3). The relative error of the aggregation rate kernels for the other aggregation pairs was significantly smaller.

Particle sizes of $\lambda_1 = 6\,\mu m$, $\lambda_2 = 8\,\mu m$ and $\lambda_3 = 10\,\mu m$ were used for aggregation due to gravitational settling. The particle densities were $n_1 = n_2 = n_3 = 10^{13}\,m^{-3}$. The physical parameters were set to the following values: $g = 9.81\,m\,s^{-2}$, $\eta = 10^{-3}\,kg\,m^{-1}\,s^{-1}$, $\rho_{part} = 900\,kg\,m^{-3}$ and $\rho_{fluid} = 10^3\,kg\,m^{-3}$. Time steps were limited to $\tau_{\max} = 21.5\,ms$. The mean aggregation rate kernel for particles of sizes λ_1 and λ_2 has a relative error of 0.046% compared to the analytical aggregation rate kernel Eq. (4). Again, the relative error of the aggregation rate kernels for the other aggregation pairs was significantly smaller.

3.2. Multiscale Simulations

For the case studies the following initial condition was used:

$$\tilde{n}(v, t_0) = \frac{n_0}{v_0} e^{-v/v_0} \quad , \tag{10}$$

with $n_0 = 10^{15}\,m^{-3}$ and $\lambda(v_0) = 1\,\mu m$. For the methodological comparison, a simple forward Euler algorithm was used to integrate the resulting ordinary differential equations with constant time steps of $\tau_{pop} = 0.4\,s$. The discrete element simulation was executed each time step simulating $N = 32000$ particles with a time span of $T_{sim} = 0.2\,s$. The results are shown in Fig. (2). Shear induced aggregation was simulated using a shear rate of $\gamma = 3.5\,s$. For aggregation due to gravitational settling the following parameters were used: $g = 9.81\,m\,s^{-2}$, $\eta = 10^{-3}\,kg\,m^{-1}\,s^{-1}$, $\rho_{part} = 1110\,kg\,m^{-3}$ and $\rho_{fluid} = 1000\,kg\,m^{-3}$.

The evolution of shear induced aggregation shows a more broadened particle size spectrum, whereas the aggregation due to gravitational settling is much steeper. The plots show an excellent agreement for particle size distribution values above $\tilde{n}(v) = 10^{25}\,m^{-4}$. The error in the remaining region represents not more than 1% of the total particle number. The reason for this deviation is connected to the initialization of the discrete element simulations. Just a discrete number of each particle size fraction can be initialized. If the expectation value of this number is of the order of 1 or below, no particles of this size category might be initialized. Hence, no aggregation rate kernel

can be calculated. These aggregation rate kernels are set to 0. The particle size distribution cannot propagate properly in regions, where aggregation partners would be necessary, which are unlikely to be initialized.

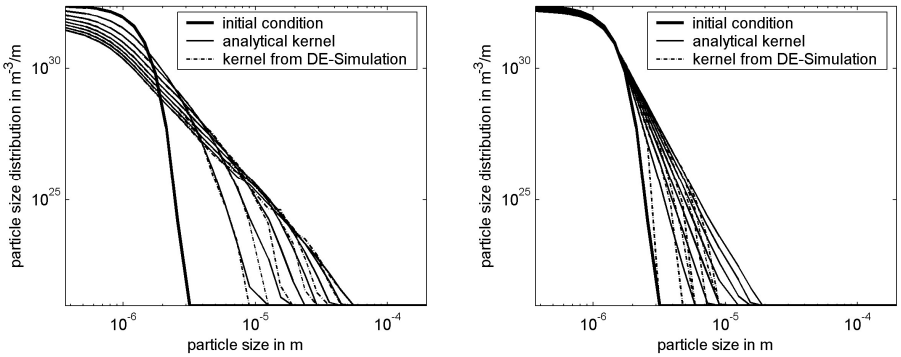


Figure 2: Evolution of particle size distribution for shear induced aggregation (left) and aggregation due to gravitational settling (right). The numerical solutions are depicted in 8 time steps of 42 seconds from 0 to 336 seconds.

3.3. Conclusions

Discrete element simulations are capable of covering a broad spectrum of modeling assumptions. Furthermore, if implemented carefully, they lead to results with no systematic error, as presented in section 3.1. The statistical accuracy of these simulations is solely limited by spent computational efforts. The first results of the multiscale framework already show an excellent agreement for the majority of the particle population. As this is only a validation study using known analytical kernels, it can only show the potential of this approach. Nevertheless, when reaching for predictive aggregation models, we conclude that the presented multiscale framework is capable of combining a high detail in modeling with efficient population balance models.

References

M. Heine, S. Pratsinis, 2007, Brownian coagulation at high concentration, Langmuir, 23, 9882-9890

M. J. Hounslow, H. S. Mumtaz, A. P. Collier, J. P. Barrick, A. S. Bramley, A micro-mechanical model for the rate of aggregation during precipitation from solution, Chemical Engineering Science, 56, 2543-2552

I. Kevrekidis, C. Gear, G. Hummer, 2004, Equation-free: The computer-aided analysis of comptex multiscale systems, AICHE Journal, 50, 1346-1355

S. Kumar, D. Ramkrishna, 1996, On the solution of population balance equations by discretization - I. A fixed pivot technique, Chemical Engineering Science, 51, 1311-1332

D. Ramkrishna, 2000, Population Balances -Theory and Applications to Particulate Systems in Engineering, Academic Press

A. Reinhold, 2008, Multiskalenmodellierung von Aggregationsprozessen, diploma thesis, Otto-von-Guericke University, Magdeburg

T. Trzeciak, A. Podgórski, J. Marijnissen, 2006, Stochastic calculation of collision kernels: Brownian coagulation in concentrated systems, WCPT5, World Congress on Particle Technology

M. Smoluchowski, 1917, Versuch einer mathematischen Theorie der Koagulationskinetik kolloider Lösungen, Zeitschrift für Physikalische Chemie, 92, 129-168

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal synthesis of heat exchanger networks using enthalpy-temperature functions to describe streams

Diego G. Oliva, Javier A. Francesconi, Miguel C. Mussati and Pio A. Aguirre

INGAR Instituto de Desarrollo y Diseño (CONICET-UTN), Avellaneda 3657, Santa Fe (3000), Argentina

Abstract

This work presents a mixed integer non linear programming (MINLP) model for heat exchanger network synthesis including a detailed description of the process streams resorting to a mathematical function that matches temperature with enthalpy. This allows the model to account for non linear behavior of streams by taking into consideration their composition. To achieve the objective, the energy balance equation proposed in the "synheat" model by Yee and Grossmann (1990) is modified but equations involving temperature are kept the same. A new equation is added to that model to describe the relationship between temperature and enthalpy by correlation. As the original model, this modified MINLP model provides the network structure that minimizes the total annual cost. The mathematical program has non-convex equations, and only locally optimal solutions can be guaranteed. The approach between cold and hot streams has more detailed information because describes the real behavior of streams, and avoids matches that a linear functionality or a classical model may allow. Correlations between temperature and enthalpy are a function of the components present in a particular stream. The required information can be collected from bibliography or public databases as those provided by NIST (National Institute of Standards and Technology). The proposed methodology is a useful tool when the minimum temperature approach between streams is small. Examples are presented and discussed to compare results when streams are described by linear and non linear functionality between temperature and enthalpy.

Keywords: Heat exchange network HEN, synthesis, MINLP models, synheat, temperature–enthalpy functionality.

1. Introduction

Two major methodologies have been proposed for the synthesis of heat exchanger network problems: sequential and simultaneous approaches. The pinch design method (Linnhoff and Hindmarsh, 1983) is one of the most-widely known, and is based in sequential approaches, in which targets are obtained sequentially using heuristic rules. On the other hand, methodologies based on mathematical programming techniques were developed. A complete review and classification of different methods were performed by Furman and Sahinidis (2002).

Yee and Grossmann (1990) developed a basic framework for heat exchanger network synthesis using a staged-superstructure formulated as a MINLP (Mixed Integer Non Linear Programming) model aiming at minimizing simultaneously the utility demands and the capital cost of the network. Several extensions of that framework were proposed addressing flexibility, incorporating detailed design aspects of the exchanger units, and focusing on global optimality issues.

Some modifications to the staged-superstructure models are here proposed to address process synthesis. When a process is to be synthesized is important to consider the composition of the streams involved because reactors behavior and structure of the entire process are function of these variables. Then, a modification to the classical stage-superstructure HEN models is made to take into account energy content and temperature approach between streams. These considerations provide a better description of streams in heat exchangers, and avoid infeasible streams matches.

2. Mathematical model

This derived model is inspired by the pioneering work of Yee and Grossmann (1990). Consequently, the superstructure proposed and the nomenclature used has some similarities with that work. Figure 1 shows the structure and variables involved. At the stage outlet, temperature and enthalpy are problem variables. It is possible to modify the original energy balances equations replacing temperature by enthalpy, maintaining streams feasibility equations with temperature, and adding a new equation to the original model that relates temperature and enthalpy. Non linear correlations for stream enthalpy as function of the stream molar composition and temperature are adopted.

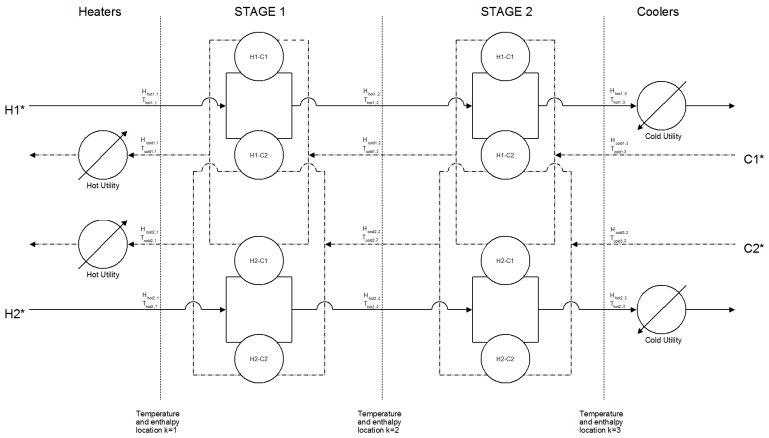


Figure 1. Superstructure with the new variables incorporated to the original model.

The new MINLP model can be written as follows (nomenclature is the same as in Yee and Grossmann, 1990):

2.1. Overall energy balance for each stream

The total heat balance for streams is given by:

$$(H_{IN_i} - H_{OUT_i}) = \sum_{k \in ST} \sum_{j \in CS} q_{i,j,k} + q_{CU,i} \quad \text{i} \in \text{HS} \qquad (1)$$

$$(H_{OUT_j} - H_{IN_j}) = \sum_{k \in ST} \sum_{i \in HS} q_{i,j,k} + q_{HU,j} \quad \text{j} \in \text{CS} \qquad (2)$$

2.2. Energy balance for each stage

The energy balance for each stage is a modified version of the Yee and Grossmann (1990) model.

$$(H_{i,k} - H_{i,k+1}) = \sum_{j \in CS} q_{i,j,k} \quad \text{i} \in \text{HS}, \text{k} \in \text{ST} \qquad (3)$$

$$(H_{j,k} - H_{j,k+1}) = \sum_{i \in CS} q_{i,j,k} \quad \text{j} \in C\text{S, k} \in \text{ST} \tag{4}$$

2.3. Inlet enthalpy assignment

These equations assign the inlet enthalpy to each stream.

$$H_{IN_j} = H_{j,k+1} \quad \text{j} \in \text{CS} \tag{5}$$

$$H_{IN_i} = H_{i,1} \quad \text{i} \in \text{HS} \tag{6}$$

2.4. Enthalpy feasibility

Here, the modified model replaces temperature by enthalpy to keep the order in each stage, and indirectly reaches temperature feasibility.

$$H_{\text{i,k}} \geq H_{i,k+1} \quad \text{i} \in \text{HS, k} \in \text{ST} \tag{7}$$

$$H_{\text{j,k}} \geq H_{j,k+1} \quad \text{j} \in \text{CS, k} \in \text{ST} \tag{8}$$

$$H_{OUT_i} \leq H_{i,k+1} \quad \text{i} \in \text{HS} \tag{9}$$

$$H_{OUT_j} \geq H_{j,1} \quad \text{j} \in \text{HS} \tag{10}$$

2.5. Heating and cooling duties

$$(H_{i,k+1} - H_{OUT_i}) = q_{CU,i} \quad \text{i} \in \text{HS} \tag{11}$$

$$(H_{OUT_j} - H_{j,1}) = q_{HU,j} \quad \text{j} \in \text{CS} \tag{12}$$

Equations (13) to (25) and (28) are the same as those proposed by Yee and Grossmann (1990). Equations (26) and (27) are added to correlate temperature and enthalpy. This allows the model to provide a better description of the streams.

2.6. Upper bound constraints

$$q_{i,j,k} - Q_{\max} y_{i,j,k} \leq 0 \quad \text{i} \in \text{HS, j} \in \text{CS, k} \in \text{ST} \tag{13}$$

$$q_{CU,i} - Q_{\max} y_{CU,i} \leq 0 \quad \text{i} \in \text{HS} \tag{14}$$

$$q_{HU,j} - Q_{\max} y_{HU,j} \leq 0 \quad \text{j} \in \text{CS} \tag{15}$$

2.7. Temperature differences

$$\Delta T_{i,j,k} \leq T_{i,k} - T_{j,k} + \Delta T^{\max}(1 - y_{i,j,k}) \tag{16}$$

$$\Delta T_{i,j,k+1} \leq T_{i,k+1} - T_{j,k+1} + \Delta T^{\max}(1 - y_{i,j,k}) \tag{17}$$

$$\Delta T_{CU,i} \leq T_{i,k+1} - T_{OUT,CU} + \Delta T^{\max}(1 - y_{CU,i}) \tag{18}$$

$$\Delta T_{HU,i} \le T_{OUT,HU} - T_{j,1} + \Delta T^{\max}(1 - y_{HU,j}) \tag{19}$$

2.8. Logarithmic mean temperature differences

$$LMTD_{i,j,k} - \left[\frac{1}{6}(\Delta T_{i,j,k} + \Delta T_{i,j,k+1}) + \frac{2}{3}\sqrt{\Delta T_{i,j,k}\Delta T_{i,j,k+1}}\right] \le 0 \tag{20}$$

$$LMTD_{CU,i} - \left[\frac{1}{6}(\Delta T_{CU,i} + T_{OUT_i} - T_{IN_{CU}}) + \frac{2}{3}\sqrt{\Delta T_{CU,i}(T_{OUT_i} - T_{IN_{CU}})}\right] \le 0 \tag{21}$$

$$LMTD_{HU,j} - \left[\frac{1}{6}(\Delta T_{HU,j} + T_{IN_{HU}} - T_{OUT_j}) + \frac{2}{3}\sqrt{\Delta T_{HU,j}(T_{IN_{HU}} - T_{OUT_j})}\right] \le 0 \tag{22}$$

2.9. Heat exchanger area requirement

$$area_{i,j,k} - \frac{q_{i,j,k}}{U_{i,j}\, LMTD_{i,j,k}} = 0 \tag{23}$$

$$area_{CU,i} - \frac{q_{CU,i}}{U_{CU,i}\, LMTD_{CU,i}} = 0 \tag{24}$$

$$area_{HU,j} - \frac{q_{HU,j}}{U_{HU,j}\, LMTD_{HU,j}} = 0 \tag{25}$$

2.10. Enthalpy-temperature function

By incorporating Equation (26) and (27) to the original model of Yee and Grossmann (1990), a better description of streams in a heat exchanger is possible. It should be noted that in equation (26) $T_{i,k}$ is a function of $H_{i,k}$, and that function is expressed in terms of constant values, i.e. A_i, B_i, C_i, D_i, E_i, F_i and H_i; and analogously for equation (27), except for the suffix j, where values of these constants are taken from the NIST database depending on the stream composition.

$$H_{i,k} = \frac{1000.F}{MW}\left[\begin{array}{l} A_i.\left(\frac{T_{i,k}}{1000}\right) + B_i.\frac{\left(\frac{T_{i,k}}{1000}\right)^2}{2} + C_i.\frac{\left(\frac{T_{i,k}}{1000}\right)^3}{3} + \\ D_i.\frac{\left(\frac{T_{i,k}}{1000}\right)^4}{4} - \frac{E_i}{\left(\frac{T_{i,k}}{1000}\right)} + F_i - H_i \end{array}\right] \qquad i \in HS \tag{26}$$

$$H_{j,k} = \frac{1000.F}{MW}\left[A_j.\left(\frac{T_{j,k}}{1000}\right) + B_j.\frac{\left(\frac{T_{j,k}}{1000}\right)^2}{2} + C_j.\frac{\left(\frac{T_{j,k}}{1000}\right)^3}{3} + D_j.\frac{\left(\frac{T_{j,k}}{1000}\right)^4}{4} - \frac{E_j}{\left(\frac{T_{j,k}}{1000}\right)} + F_j - H_j \right] \qquad j \in CS \qquad (27)$$

2.11. Total annual cost

The objective function to be minimized is the Total Annual Cost (TAC)

$$TAC = \sum_{i\in HS} CCU\, q_{CU,i} + \sum_{j\in CS} CHU\, q_{HU,j} + \\ + \sum_{i\in HS}\sum_{j\in CS}\sum_{k\in ST} CF_{i,j} y_{i,j,k} + \sum_{i\in HS} CF_{i,CU}\, y_{CU,i} + \sum_{j\in CS} CF_{j,HU}\, y_{HU,j} + \\ + \sum_{i\in HS}\sum_{j\in CS}\sum_{k\in ST} C_{i,j}\left(area_{i,j,k}\right)^{\beta} + \sum_{i\in HS} C_{i,CU}\left(area_{i,CU}\right)^{\beta_{CU}} + \sum_{j\in CS} C_{j,HU}\left(area_{j,HU}\right)^{\beta_{HU}} \qquad (28)$$

3. Example

Water streams with different temperature levels are used. Table 1 shows the streams specifications.

Table 1. Streams Data.

Stream	TIN (K)	TOUT (K)	F(kg seg-1)	Cp (kJ kg-1 K-1)	F.Cp (kW K-1)	h (kW m-2 K-1)
H1	370	290	10.00	4.19	41.87	1.80
H2	480	380	7.00	2.02	14.14	2.00
HU	627	627	-	-	-	2.50
C1	280	350	8.00	4.19	33.52	1.80
C2	400	600	5.00	2.02	10.10	2.00
CU	280	285	-	-	-	1.00

Note that original model uses Cp values from Table 1 while proposed model uses parameters depicted in Table 2.

Table 2. Water coefficients used in temperature-enthalpy correlation.

	Water	
Coefficient	Condensed phase	Gas phase
A	-203.606	30.092
B	1523.290	6.833
C	-3196.413	6.793
D	2474.455	-2.534
E	3.855	0.082
F	-256.548	-250.881
G	-488.716	223.397
H	-285.830	-241.826

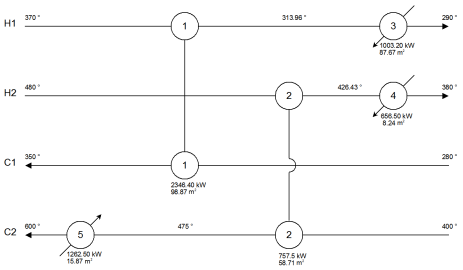


Figure 2. Network using the traditional model.

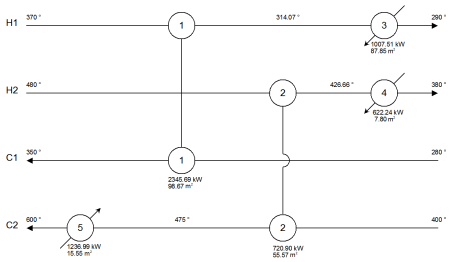


Figure 3. Network using the new model.

Table 3. Differences in the computed energy content according to the used model.

Stream	Energy Content Computed [kW]		Difference Between Models	%
	Traditional Model	New Proposed Model		
Hot 1	3349.60	3353.20	3.60	0.11
Hot 2	1413.50	1343.14	-70.36	-4.98
Cold 1	2346.40	2345.69	-0.71	-0.03
Cold 2	2020.00	1957.89	-62.11	-3.07

Differences between models are the energy computed in each stream (Table 3) and the temperature approach between streams. For example, the traditional model computes an energy content of 2020.00 kW for cold stream 2 while the proposed model with the temperature-enthalpy correlation computes 1957.89 kW. Figure 2 and Figure 3 show that the arrangements obtained with the traditional model and with the proposed model are similar. However, the proposed model provides a better solution than the original model regarding the TAC value, i.e. the original model computes $174,744.80 while the proposed model computes $171,622.91. Both models were implemented in GAMS (Brooke et al., 1997) and solved with the solver DICOPT solver (Duran and Grossmann, 1986).

4. Conclusion

A more rigorous description of process streams was proposed using a correlation. The temperature approach and the streams energy content are better described and computed using the temperature-enthalpy correlation as proposed in this model than the linear correlation frequently used. The proposed methodology is a useful tool when the minimum temperature approach between streams is small.

References

B. Linnhoff, E. Hindmarsh, Chemical Engineering Science 38 (1983) 745-763.

K.C. Furman, N.V. Sahinidis, Indust. and Engineering Chemistry Research 41 (2002) 2335-2370.

T.F. Yee, I.E. Grossmann, Computers & Chemical Engineering 14 (1990) 1165-1184.

Secretary of Commerce on behalf of the United States of America, NIST Chemistry WebBook http://webbook.nist.gov/chemistry/, 2008.

M. A. Duran and I. E. Grossmann, An Outer-Approximation Algorithm for a Class of Mixed-Integer Nonlinear Programs, Mathematical Programming, 36 (1986), pp. 307–339.

A. Brooke, D. Kendrick, and Meeraus, A., GAMS Language Guide Release 2.25, 92 ed., GAMS Development Corporation, 1997.

Acknowledgements. The financial support from CONICET is acknowledged.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Computer Simulation for Oil Wells with Artificial Lift Method by Electrical Submersible Pump

Evellyne S. Batista,[a] Rutácio O. Costa,[b] André L. Maitelli,[a] Tiago S. Barbosa,[a] Andrés Ortiz Salazar [a]

[a]*LAUT/DCA – Laboratory of Oil Automation/ Department of Engineering Computer and Automation - University of Rio Grande do Norte, Campus Universitário Lagoa Nova, 59072-970 Natal – RN, Brazil*
[b]*PETROBRAS UN-RNCE*
Av Euzébio Rocha, 1000 Cidade da Esperança, 59164-100 Natal - RN, Brazil

Abstract

The main goal this paper is related with the development of computational simulator for the artificial lift method based on Electrical Submersible Pump (ESP). This simulator will be able to represent the dynamic behavior, considering the source's model and electric energy transmission for the motor, the ESP model's electric motor (including the thermal calculation), flow simulations and three-dimensional animations for each ESP subsystem (motor, pump, seal, and so on). The simulator proposes an improvement in oil wells monitoring through a better knowledge of physical phenomena more relevant of the process. These information can prevent failures, facilitating decision-making in a smaller portion of time. Similarly, it allowing to control the increased oil production through adjustments of operating conditions (frequency, head well pressure, operations points, etc.).
Keywords: Electrical submersible pump, artificial lifting, modeling and optimization.

1. Introduction

For the oil production is necessary, first of all, discover the potential field producer through seismic and geological studies. Later, a well crossing one or more hydrocarbon bearing zones should be drilled. During the drilling is possible to confirm the presence of oil. In this case, the well should be coated and completed with a system of natural or artificial elevation that allows the lifting of fluids to the surface. The ESP is a method of artificial lift in which the power is transmitted by a cable to the electric motor that connects the centrifugal pump as illustrated in figure 1. The flow produced depends of reservoir characteristics, fluids properties, internal geometry of pump, rotation motor and suction and discharge pressures. It can change the flow produced, within certain limits, controlling the opening of control's valve (beam) or the rotation of centrifugal pump by adapting the motor's frequency. The cost is high in case of failure, because it needs of the drilling rig. There are cases in which all equipment must be replaced, further increasing costs. The time waiting for rig may be high, especially in offshore applications, generating losses due to delay in oil production.

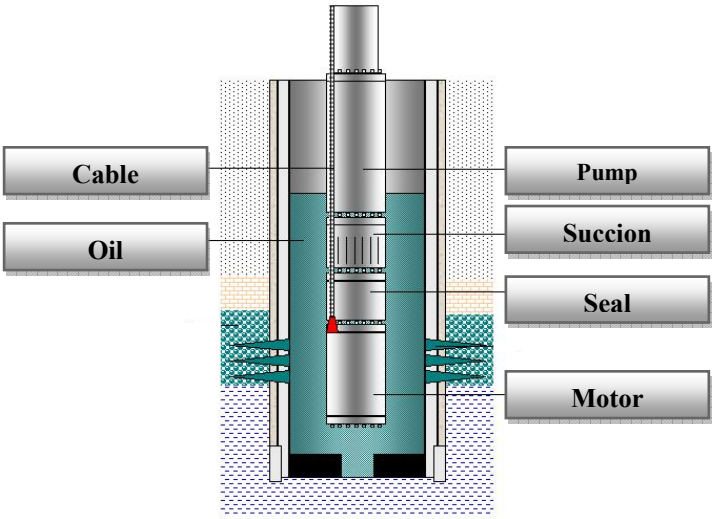


Fig.1. ESP system

1.1. Motivation

In the technical-economic view, it is expected that the flow will be as large as possible. This context leads to adjustment in beam on surface and inverter's frequency on electric motor. Another basic requirement is related with the failures time. It is expected that the failures average time will be as long as possible to decrease loss of profit and maintenance costs. Sometimes is necessary to stop motor because operation schedules or failures on electric system. In this case, it is strongly recommend hold the stop status for a time even motor's temperature decrease for a safe work level. Currently, this time is defined based on operator experience and reading of diverse operational variables. A simulator can be used to train the operators about the real conditions work through features fluids, reservoir scheme, well design and pump models. Currently, there is not a commercial simulator that provides the previous functionalities.

2. Description of Model

In the technique of computer simulation is established a model able to describe or represent the real problem to be submitted an "experimental" manipulation in a computer. Through simulations, it is possible to visualize the main stages of process. In this context, this paper proposes a simulator for the computational method of artificial lift, considering the characteristics of the reservoir, model of fluid flow inside a pump ESP, model of electric motor considering the heating effect and integration of Variable Speed Drives (VSD).

2.1. Motor

The electric motors used to boost the ESPs are induction phase, tubular geometry usually two poles which rotate around 3500 rpm for a frequency of 60Hz. The motors are designed to work in difficult conditions: immersed in fluids that being produced, high pressures and high temperatures. To support these conditions, the motors have a mineral oil to ensure electrical isolation, lubrication of bearings and cooling of motor.

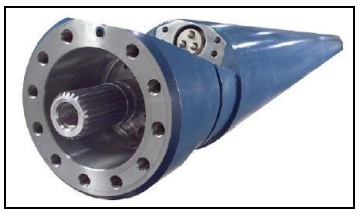


Fig.2. Induction motor

The model of dynamic behavior of electric motor was based in [1]. This model is based on linear processing called dq0 that allows the calculation of torque linking voltages, currents, flow dispersion and angular speed instantly. From these relations is estimated power factor, active power, efficiency and the heat generated in motor windings.

2.2. Pump

The pump used is a centrifugal with multiple stages, each stage consisting of an impelidor and diffuser. The impelidor creates a centrifugal field which is responsible for increasing the speed and pressure of the fluid. The flow through the diffuser the fluid will find a progressive increase in the flow that cause drop in speed and increase in pressure (Bernoulli's theorem). Each stage provides an increase of the pressure fluid. The pump's stages are based on requirements fluids productions. In figure 3 is represented a stage of the pump and the principle of operation for several stages.

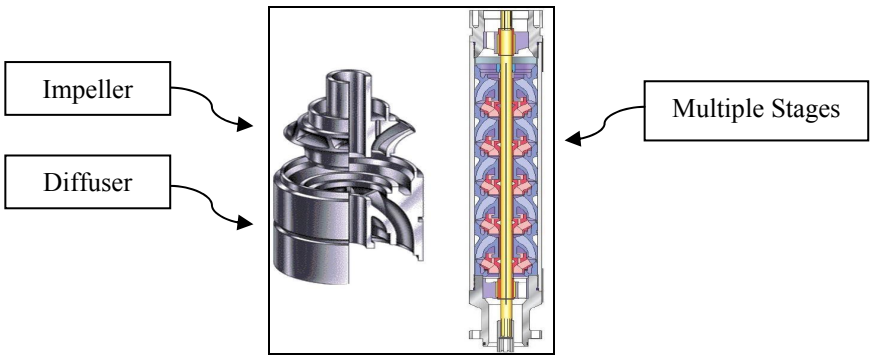


Fig.3. Representation of centrifugal pump

The shape and size of impelidor and diffuser determine the flow to be pumped, stages number and their ability to lift or head. The manufacturers publish a curve of each model of pump. These curves are given considering a single stage draining water should be corrected when applied to other values of fluid viscosity. If present biphasic flow inside the pump the effect of gas must be considered. The simulator will correct the curve of performance the pump when pumping viscous oils. There are several procedures in the literature to correct pump performance for viscous applications. The simulator is based on the Hydraulic Institute. In figure 4, there is the behavior of the curve of the pump for water and a viscosity of 40cp.

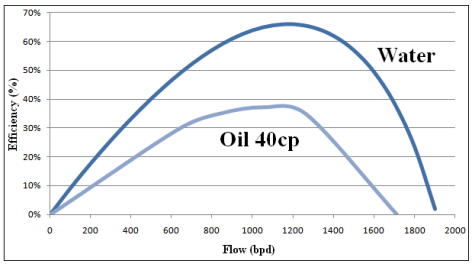


Fig.4. Effect of viscosity

The simulator will use motors with variable speed (VFD), so it is necessary to calculate the new performance curve for different frequencies of electric motor.
According to [2] changes following relations:

1. The flow (Q) changes directly proportional to rotation (N):

$$\frac{Q_2}{Q_1} = \frac{N_2}{N_1}$$

2. The head changes directly proportional with the square of the change in rotation:

$$\frac{Q_2}{Q_1} = \left(\frac{N_2}{N_1}\right)^2$$

3. The power changes directly proportional with the cube of the change in rotation:

$$\frac{Pot_2}{Pot_1} = \left(\frac{N_2}{N_1}\right)^3 \quad ie \quad \frac{Q_2}{Q_1} = \frac{N_2}{N_1} = \sqrt{\frac{H_2}{H_1}} = \sqrt[3]{\frac{Pot_2}{Pot_1}}$$

In figure 5, is an example of application of affinity laws illustrating the performance curve of pump for various frequencies.

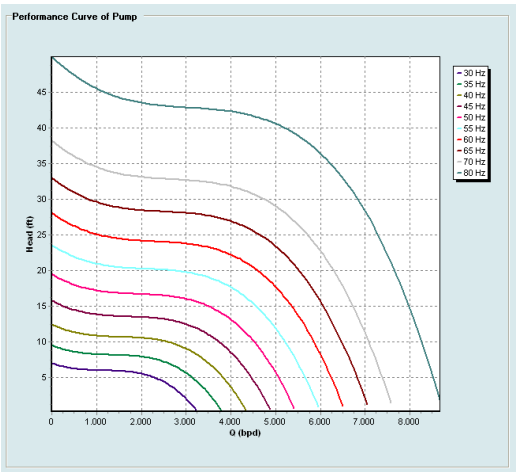


Fig.5. Application of affinity laws

2.3. Reservoir

For development a simulator of artificial lift is required detailed knowledge of the factors that determine the flow through the porous medium to the well. The pressure's curves available in a reservoir, called IPR curves (inflow Performance Relationship) [3]. They represent the relationship between the pressure flow at the bottom of the well and flow.

- IPR Vogel

$$\frac{q_o}{q_{\max}} = 1 - 0{,}2\frac{p_{wf}}{p_e} - 0{,}8\left(\frac{p_{wf}}{p_e}\right)^2$$

Such that

p_e :static pressure	$q_{\max}$:maximum flow	q_0 : flow	p_{wf} : pressure

2.4. Flow of fluids

The fluids that flow through an oil well are complex mixtures of water and hydrocarbons. This work is considering the flow of oil, water and gas.

This simulator used the calculation of multiphase flow in two sections of pipe: casing, between the perforation and pump, in the production tubing between the pump and the surface. The correlations adopted are widely used in the petroleum industry, as described by [4].

3. Results

3.1. Graphical User Interface.

The main screen of the simulator is shown in Figure 6, so that the three-dimensional animation shows the dynamic behavior of the whole process. Also will be able to make the animation of each component of the BCS system (pump, motor, separator, protector). On the right side of the screen, the simulator will show the curve of the pump, dynamic level, the engine torque curve, rotation of the engine, the current curves.

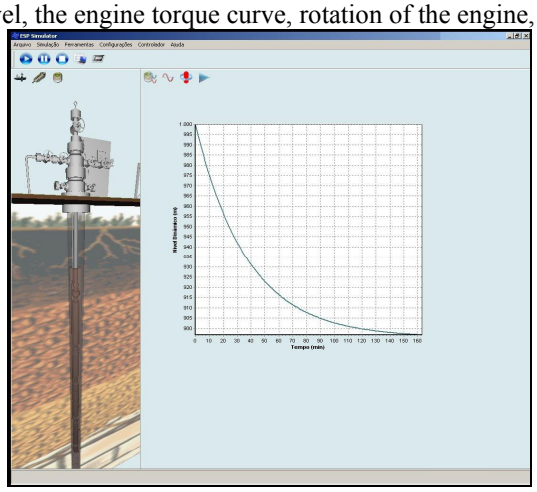


Fig.6. The main screen of the simulator

The setup screen of the simulator has been subdivided into properties of fluids, mechanical scheme, correlations of multiphase flow, reservoir and setting the pump.

3.2. Simulation of a case

In implementation of a case picked up the model of the pump, number of stages and a frequency of 60 Hz, as shown in Figure 7.

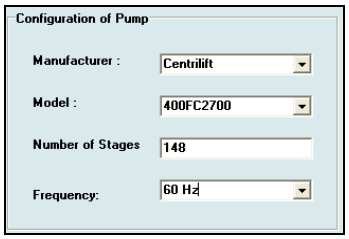


Fig. 7. Setting the pump.

The figure 8 illustrates the input data for the properties of fluids and reservoir.

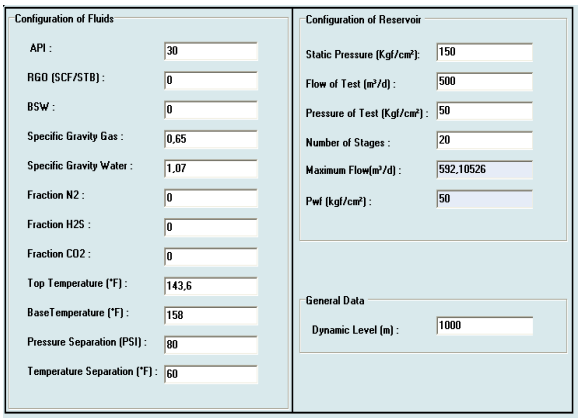


Fig.8. Setup screen of the properties of reservoir and fluids

The graph representing the dynamic level of the well (m), depending on the time (min) to the input data described, is shown in Figure 6. Until the system reaches the equilibrium (steady), recording the final dynamic level in the 905m with steady flow in the 491, 4 m³/d. The simulator allows the adjustment of the feeding frequency in the 70Hz which will be a flow of 564,2 m³/d and dynamic level 1240 m.

4. Conclusions

It is viable to develop a computer simulator for the transient BCS system, from the integration of models available in the literature for the behavior of the reservoir, the pump, the flow of fluids and the electric motor. The computer simulation presented in this paper allows the prediction of operational parameters of the BCS system. The operation of the system in a more rational may contribute to increased production of oil and the useful life of equipment, reducing costs of intervention and intermissive profit. The simulator will also be useful for training of operators, helping to increase the efficiency of the industry of oil production.

5. References

1. Ong, Chee-Mun (1998), “Dynamic Simulation of Eletric Machinery”, 1º edition, Prentice
2. Hall.Mattos, Edson Ezequiel. Falco. R. “Bombas Industriais”/, 2º edition - Rio de Janeiro: Interciência, 1998
3. Lee, H.K. “Computer Modeling and Optimization for Submersible Pump Lifted Wells”, paper SPE 17586 presented at the 1988, China (November 1-4)
4. Brill, J.P., Mukherjee, H., “Multiphase Flow in Wells”, Monograph Volume 17 SPE Henry L. Doherty Series, SPE, Richardson-Texas, 1999.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Computation of Crossover Pressure for Synthesis of Supercritical Fluid Separation Systems

Silvio A.B. Vieira de Melo,[a*] Gloria M.N. Costa,[a] Ana C.C. Viana,[b]

Fernando L.P. Pessoa[c]

[a] *Programa de Engenharia Industrial - Escola Politécnica – Universidade Federal da Bahia, Rua Prof. Aristides Novis, 2, Federação, 40210-630, Salvador, Bahia, Brazil*
[b] *Escola Politécnica – Universidade Federal da Bahia, Rua Prof. Aristides Novis, 2, Federação, 40210-630, Salvador, Bahia, Brazil*
[c] *Departamento de Engenharia Química, Escola de Química, Universidade Federal do Rio de Janeiro - Centro de Tecnologia, bloco E, S. 209, Ilha do Fundão, 21949-900, Rio de Janeiro-RJ, Brazil*
** Corresponding author: sabvm@ufba.br*

Abstract

The crossover region is a phenomenological observation which may be used to synthesize supercritical fluid extraction process (SCFEP) and is supported by a number of experimental studies. This phenomenon means that there may be regions of temperature and pressure in the near critical fluid state where the solute solubility decreases with an isobaric increase in temperature. Accurate calculation of solid solubility in supercritical fluid does not imply that crossover pressure calculation has also good accuracy. In this paper, a new method to predict crossover pressure based on cubic equations of state is developed and applied to systems containing supercritical CO_2 and bixin. Peng-Robinson-LCVM-UNIFAC equation of state is used in a predictive way and crossover pressure is determined for several conditions. The method described here can in principle be applied to any solid-fluid system used in supercritical fluid extraction process synthesis.

Keywords: Supercritical fluid, crossover pressure, process synthesis

1. Introduction

The crossover region is a phenomenological observation supported by a relevant number of experimental studies. This phenomenon means that there may be regions of temperature and pressure in the near critical fluid state where the solute solubility decreases with an isobaric increase in temperature (Kelley and Chimowitz, 1989). Despite many database available for the solubility of solids in supercritical fluids, most of models fits experimental data to binary parameters only for solubility calculation and much less attention is given to accurate determination of crossover pressure.

The aim of this work is to present and discuss the principal mathematical aspects and equation constraints for computation of crossover pressure. A modified Peng-Robinson equation of state is used to exemplify crossover pressure calculation for the binary system containing supercritical carbon dioxide with bixin at several conditions.

2. Crossover Pressure Approach Modeling

The crossover pressure is defined as the point where the slope of the solubility versus temperature curve changes sign and the opposite effects of solute vapor pressure and solvent density on solubility compensate each other. Mathematically, it is expressed as

$$\left(\frac{\partial \ln y_2}{\partial T}\right)_P = 0 \tag{1}$$

where y_2, T and P denote for solute solubility, temperature and pressure, respectively.

It is well known that when plotting solubility versus pressure, isothermal curves cross each other twice and these intersections are called lower crossover pressure (LCP) and upper crossover pressure (UCP) points. At pressures between LCP and UCP, the solubility decreases increasing temperature because the solvent density effect overcomes the vapor pressure effect. At pressures below LCP or above UCP, the solubility increases with temperature as a consequence of the rapid increase in vapor pressure compared to a less pronounced solvent density effect (Foster et al., 1991).

Considering in the calculation of solid-SCF equilibrium that: (i) SCF does not dissolve in the solid phase; (ii) the molar volume of the solid is independent of pressure; (iii) the fugacity coefficient of pure vapor at temperature T and vapor pressure is unity, then, the solubility of solute can be expressed as

$$\ln y_2 = \ln P_2^s + \frac{v_2^s\left(P - P_2^s\right)}{RT} - \ln \hat{\phi}_2 - \ln P \tag{2}$$

where P_2^s and v_2^s are vapor pressure and molar volume of solid solute, respectively, and $\hat{\phi}_2$ is the fugacity coefficient of solid solute in the fluid phase (Foster et al., 1991).

Calculation of crossover pressure could be done in principle using Eq. (2) because at this point solubility values are equal. However, it is mathematically very hard to find the root P of this equation at the crossover point (Vieira de Melo *et al.*, 2009). To avoid this kind of problem, it is recommended to combine Eqs. (1) and (2), and the following equation is obtained for calculating the crossover pressure:

$$\frac{d \ln P_2^s}{dT}\left(1 - \frac{v_2^s P_2^s}{RT}\right) - \frac{v_2^s}{RT^2}\left(P - P_2^s\right) - \left(\frac{\partial \ln \hat{\phi}_2}{\partial T}\right)_P = 0 \tag{3}$$

Solute vapor pressure is calculated by

$$\ln P_2^s = A - \frac{B}{T} \tag{4}$$

The Clapeyron equation gives the relation between solid vapor pressure and enthalpy of sublimation, as follows

$$\frac{dP_2^s}{dT} = \frac{\Delta H^{sub}}{T.\left(v_2^g - v_2^s\right)} \tag{5}$$

Deriving Eq. (4) related to temperature gives

$$\frac{d\left(\ln P_2^S\right)}{dT} = \frac{1}{P_2^s}\frac{dP_2^S}{dT} = \frac{B}{T^2} \tag{6}$$

Comparing (5) and (6) gives

$$P_2^s = \frac{\Delta H^{sub}.T}{B\left(v_2^g - v_2^s\right)} \tag{7}$$

Substituting the Eqs. (4) and (6) in the Eq. (3), since the vapor pressure is very low regarding to the pressure P, Eq. (3) is simplified to give

$$\mathrm{F(P)} = \frac{B}{T^2} - \frac{v_2^s P}{RT^2} - \left(\frac{\partial \ln \hat{\phi}_2}{\partial T}\right)_P = 0 \tag{8}$$

The root from Eq. (8) is a crossover pressure. Some general considerations are needed to determine the crossover pressure.

$$\text{Defining } FT(2) = \left(\frac{\partial \ln \hat{\phi}_2}{\partial T}\right)_P \tag{9}$$

and deriving the Eq. (8) related to *P* with constant temperature, and taking into account the Eq. (9), one can obtain:

$$\left(\frac{\partial F(P)}{\partial P}\right)\Bigg|_T = -\frac{v_2^s}{RT^2} - \frac{\partial FT(2)}{\partial P} \tag{10}$$

On the inflexion point of *F(P)*, one must have the descent part of *FT(2)* because,

$$\frac{\partial FT(2)}{\partial P} < 0 \tag{11}$$

On the other hand, the inflexion point of *FT(2)* corresponds the negative part of *F(P)* inclination, so

$$\left(\frac{\partial F(P)}{\partial P}\right)\Bigg|_T = -\frac{v_2^s}{RT^2} < 0 \tag{12}$$

As a consequence one can evaluate the general aspect of the *F(P)*. The two roots, which correspond to two crossover points, are obtained as: the inferior crossover correspond to inferior point of *FT(2)*, the superior crossover correspond to the negative inclusive of *FT(2)* curve. Thus, the inflexion point of *F(P)* corresponds to a negative inclination of *F(T)* curve. One can conclude that the two crossover pressures correspond to the negative inclination of *FT(2)* curve.

It is also important to analyze the influence of parameter B on crossover pressure. Vieira de Melo *et al.* (2009) show that as the parameter B diminishes at constant temperature, the crossover pressure increases. One can analyze the relative contribution related to the terms of the new equation for the determination of crossover pressure, that is, Eq. (8). Considering P higher to the inflexion point of Eq. (8), one has in this case

$$\left(\frac{\partial F(P)}{\partial P}\right) > 0 \tag{13}$$

Therefore, $(-\frac{v_2^s}{RT^2} - \frac{\partial FT(2)}{\partial P}) > 0$ (14)

As seen before, for this pressure *P*, $\frac{\partial FT(2)}{\partial P} < 0$, and one can say that

$$\left|\frac{\partial FT(2)}{\partial P}\right| > \frac{v_2^s}{RT^2} \tag{15}$$

Consequently, from the analyses of a pressure P lower to the inflexion point pressure,

$$\left(\frac{\partial F(P)}{\partial P}\right) < 0 \quad \text{and} \quad -\frac{v_2^s}{RT^2} - \frac{\partial FT(2)}{\partial P} < 0 \tag{16}$$

In relation to the Eq. (16), one has two evaluations to do:

(i) For $\frac{\partial FT(2)}{\partial P} > 0$, the Eq.(16) will be satisfied independently of the relation between the modules of the terms of the inequality;

(ii) To satisfy the inequality $\frac{\partial FT(2)}{\partial P} < 0$, one must has necessarily the following behaviour: $\left|\frac{\partial FT(2)}{\partial P}\right| < \frac{v_2^s}{RT^2}$

So, it can be concluded that for pressures above the upper crossover pressure and pressure below the lower crossover pressure, the values of function *F(P)* must have positives. Considering for the Eq. (8) that:

$$L = -\frac{v_2^s.P}{RT^2} - FT(2) \tag{17}$$

For pressures on this range, $F(P) > 0$ and as a consequence $(B/T^2) < L$. For pressures between the two crossover pressures, $F(P) < 0$ and therefore $(B/T^2) < L$. At both crossover pressures, the value of $F(P)$ is null and $(B/T^2) = L$. For a constant temperature T, the term B/T^2 is constant and the term $v_2^s.P/RT^2$ corresponds to a linear variation with the pressure. At constant temperature, for pressure between upper and lower crossover pressures, the value of is higher than B/T^2.

It can be theoretically shown (Vieira de Melo *et al.*, 2009) that depending on the value of B/T^2 in Eq. (11) the others terms of this equation may be less or more significant. If this first term decreases, the second term seems to affect more and the third one to affect less on the result, and as a consequence the crossover pressure increases. Calculation of upper crossover pressure depends on the sublimation pressure parameters A and B, whose initial input values are determined by two different group contribution procedures. In the first procedure, parameters A and B are initially determined from group contribution, using input information of critical temperature (T_c), critical pressure (P_c) and normal boiling point temperature (T_b), and subsequently internally adjusted. In the second procedure, parameters A and B are initially calculated using the enthalpy of sublimation (ΔH^{sub}) and the melting point temperature (T_f), and subsequently internally adjusted.

3. Results and Discussion

The results obtained in this paper consider only the upper crossover pressure (UCP) for the system CO_2-bixin. To calculate crossover pressure, only solubility experimental data and the molecular structure of the solute are given. The properties of the compounds of this system and the detailed calculation procedure can be found in Vieira de Melo *et al.* (2009). Table 1 presents the calculated values of crossover pressure (*Pco*), sublimation pressure parameter *B* and molar volume of solid solute (v_2^s) for the isotherms of 288, 308 and 328K. The influence of parameter B on the crossover pressure calculated for the system CO_2-bixin at 288, 308 and 328 K is shown in Fig. 1. Increasing B decreases the convergence crossover pressure for each temperature as theoretically can be demonstrated. Although the values of molar volume and sublimation pressure of solid solute are tuned during convergence procedure of crossover pressure calculation, it is important to remark that they are interdependent due to Equation (2). So, increasing sublimation pressure decreases the molar volume of solid. As one could expect, sublimation pressure values are very low.

Table 1. Calculated crossover pressure (*Pco*), sublimation pressure parameter *B* and molar volume of solid solute (v_2^s) for CO2-bixin at 288, 308 and 328K

		Pco (atm)		
Vs (cm3/gmol)	B	T = 288 K	T = 308 K	T = 328K
362.9	15153.5	109.3	169.7	216.2
387.1	14118.9	148.2	205.2	255.2
411.2	14298.0	143.1	203.2	255.1
435.4	14376.6	147.0	209.0	263.2
447.5	13384.4	234.0	275.4	275.4
450.0	13294.0	250.2	286.8	286.8
459.6	13375.9	250.8	288.8	288.8

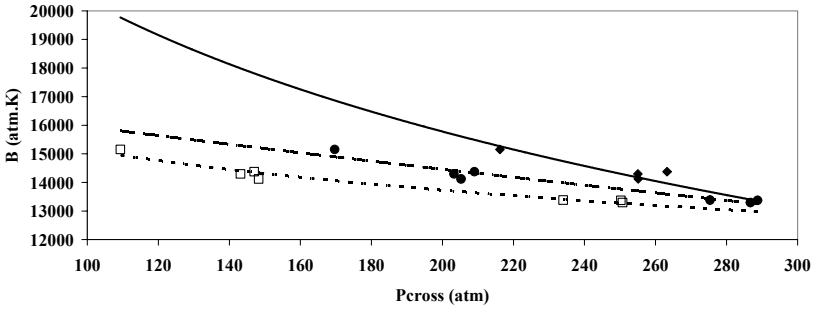


Fig. 1. Influence of parameter B on crossover pressure calculated for the system CO_2-bixin at 288 K (□), 308 K (●) and 328 K (♦). The solid curve, the dashed curve and the dotted-dashed curve indicate the tendency lines.

4. Conclusions

In this paper, a new method to predict crossover pressure based on cubic equations of state is developed and applied to systems containing supercritical CO_2, and bixin. Peng-Robinson-LCVM-UNIFAC equation of state (Boukouvalas *et al.*, 1994) is used in a predictive way and crossover pressure is determined for several conditions. Correlations among many important properties as molar volume and vapor pressure of solid solute are investigated and it is shown that sublimation pressure has particularly an important role on the accurate calculation of crossover pressure. Depending on the slope of sublimation pressure curve, the three terms in the equation for crossover pressure calculation can be less or more significant on the results. Despite the illustration of this method using Peng-Robinson-LCVM-UNIFAC equation of state for CO_2-bixin, it is important to say that the crossover pressure analysis presented in this paper is independent of the equation of state used to calculate the fugacity coefficient of solid solute in the fluid phase. So, this method can be applied to any solid-fluid system used in supercritical fluid extraction.

Acknowledgements

The authors are gratefully acknowledged to PROCAD/CAPES, CNPq, and FAPERJ.

References

F.D. Kelley, E.H. Chimowitz, 1989, Experimental data for the crossover process in a model supercritical system**,** AIChE Journal 35, 981-987.

N.R. Foster, G.S. Gurdial, J.S.L. Yun, K.K. Liong, K.D. Tilly, S.S.T. Ting, H. Singh, 1991, Significance of the crossover pressure in solid-supercritical fluid phase equilibria, Ind. Eng. Chem. Res. 30 1955-1964.

C. Boukouvalas, N. Spiliots, N. Coutsikos, D. Tassios, 1994, Prediction of Vapor-Liquid Equilibrium with the LCVM model: a linear combination of the Vidal and Michelsen mixing rules coupled with the original UNIFAC and the t-mPR equation of state, Fluid Phase Equilib. 92, 75-106.

S.A.B. Vieira de Melo, G.M.N. Costa, A.C.C. Viana, F.L.P.Pessoa, 2009, Solid pure componenet property effects on modelling upper crossover pressure for supercritical fluid process synthesis: A case study for the separation of Annatto pigments using SC-CO_2, J. Supercrit. Fluids, *in press*.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Computational Aspects for Optimization of High Pressure Phase Equilibrium for Polymer Industrial Systems

Gloria M.N. Costa,a,* Yuri Guerrieri,b Simão Kislansky,b

Silvio A.B. Vieira de Melo,a Marcelo Embiruçu,a Fernando L.P. Pessoac

[a] *Programa de Engenharia Industrial - PEI, Escola Politécnica, Universidade Federal da Bahia, R. Aristides Novis, n. 2, Federação, Salvador - BA, 40210-630, Brasil*
[b] *Escola Politécnica, Universidade Federal da Bahia, R Aristides Novis, n. 2, Federação, Salvador - BA, 40210-630, Brasil*
[c] *Departamento de Engenharia Química, Escola de Química, Universidade Federal do Rio de Janeiro, CT, bloco E, S. 209, 21949-900, Rio de Janeiro-RJ, Brasil*
** E-mail address: gmeyberg@ufba.br*

Abstract

Many computational problems occur during optimization of high-pressure phase equilibrium of polymer systems due to the existence of critical points and to the possibility of there being more than two equilibrium phases. In this work the main numerical aspects related to high-pressure flash simulation with the Soave-Redlich-Kwong (SRK) and the Sanchez and Lacombe (SL) equations of state are examined. A LDPE (low-density polyethylene) high pressure flash industrial separator is used to case study of 8 resins. The results show that Sanchez-Lacombe equation of state is more appropriate to overcome the numerical difficulties associated with the complexity of polymer systems.

Keywords: high pressure, phase equilibrium, polymer, optimization.

1. Introduction

Optimization of polymerization process which includes some separation steps depends critically on the accurate knowledge of polymer phase equilibrium. Modeling of thermodynamic properties of polymer mixtures demands considerable efforts due to the large difference in the molecular sizes of polymers and solvents, mainly at high pressures. Reliable simulators are essential to calculate rigorously industrial separator operational conditions. However, many computational problems occur in this kind of system associated with the existence of critical points and their mathematical singularity, and also due to the possibility of there being more than two equilibrium phases (Heidemann, 1983). Several attempts to use EOS to model polymer systems can be found in the literature, but they focus on low pressure ranges and laboratory experimental data (Wang, 2007; Kalospiros and Tassios, 1995). As a consequence, these systems contain low molecular weight polymer and are investigated very far from the typical operational conditions of industrial polymer separation processes. In this work some advances in the computation of high pressure phase equilibrium for polymer systems with equations of state are presented and several numerical aspects are discussed in detail. A high pressure flash separator for polyethylene is used as a case-study to explicit these features. Performances of Sanchez-Lacombe (SL) and Soave-

Redlich-Kwong (SRK) equations of state are compared using industrial experimental data. The SRK equation of state is employed with two different mixing rules: LCVM and Wong-Sandler. Bogdanic-Vidal activity coefficient model is coupled with SRK EOS through these mixing rules as recommended Costa *et al.* (2008). The main numerical aspects examined for each EOS employed are the influence of mass and mole composition in the phase equilibrium, the general optimization strategy for the estimation of the EOS interaction parameters, the algorithms of flash calculation, the influence of low and high polymer concentration, and the convergence criteria at each step calculation.

2. Thermodynamic Models

2.1. Soave-Redlich-Kwong (SRK) equation of state

The Soave-Redlich-Kwong (SRK) equation of state *(Soave, 1972)* is

$$P = \frac{RT}{v-b} - \frac{a(T)}{v(v+b)} \tag{1}$$

where parameters *a* and *b* can be given by Wong-Sandler mixing rules (Wong and Sandler, 1992):

$$a = b \cdot \left(\sum_i \frac{x_i \cdot a_i}{b_i} + \frac{G^E}{\delta} \right) \tag{2}$$

$$b = \frac{\sum_i \sum_j x_i \cdot x_j \cdot \frac{1}{2} \cdot \left[\left(b - a/R \cdot T\right)_i + \left(b - a/R \cdot T\right)_j \right] \left(1 - WS_{ij}\right)}{1 - \sum_i \left(a_i/b_i \cdot R \cdot T\right) - \left(G^E/\delta \cdot R \cdot T\right)} \tag{3}$$

δ is a numerical constant equal to *-ln2* for SRK EOS and WS_{ij} is the interaction parameter.

LCVM mixing rule is given by (Boukouvalas *et al*., 1994):

$$\alpha = \frac{a}{b \cdot R \cdot T} = \sum x_i \cdot \alpha_i + \left(\frac{\lambda}{A_v} + \frac{1-\lambda}{A_M} \right) \cdot \frac{G^E}{R \cdot T} + \frac{1-\lambda}{A_M} \cdot \sum x_i \cdot \ln\left(\frac{b}{b_i} \right) \tag{4}$$

with $A_V = \ln 2$ and $A_M = -0.53$ for SRK EOS and λ is 0.36. The classical linear mixing rule is used for parameter *b*.

LCVM and Wong-Sandler mixing rules need a G^E model and the recommended one is the BV model (Bogdanic and Vidal, 2000), whose segment activity coefficients are calculated through the UNIQUAC model and contains an interaction parameter UQ_{ij}.

2.2. Sanchez-Lacombe (SL) equation of state

The Sanchez and Lacombe EOS (Sanchez and Lacombe, 1976) is given by:

$$\tilde{\rho}^2 + \tilde{P} + \tilde{T} \cdot \left[\ln(1-\tilde{\rho}) + \left(1 - \frac{1}{r}\right) \cdot \tilde{\rho} \right] = 0 \tag{5}$$

$$\tilde{T} = \frac{T}{T^*} \qquad \tilde{P} = \frac{P}{P^*} \qquad \tilde{\rho} = \frac{\rho}{\rho^*} \tag{6}$$

$$T^* = \frac{\varepsilon^*}{k} \qquad P^* = \frac{\varepsilon^*}{v^*} \qquad \rho^* = \frac{MW}{r v^*} \tag{7}$$

where T is the absolute temperature, P is the pressure, ρ is the density, MW is the molecular weight, k is the Boltzmann constant, and r, ε^*, and v^* are pure component parameters related to the corresponding scale factors T^*, P^*, ρ^*, respectively. These scale factors are independent of the molecular size of the polymer. For mixtures, the model parameters are composition dependent. This dependence is expressed in terms of segment fraction of component i. The binary interaction parameters $SL1_{ij}$ and $SL2_{ij}$ are related to the cross parameters.

3. Equilibrium Calculation Methods

Equilibrium flash calculation is used to simulate the high pressure separator and to compare the results of two phase compositions x_i and y_i with industrial data. In flash calculations, the problem is to find β, the vapor phase equilibrium fraction of a mixture with initial composition z_i. This fraction is updated iteratively by first solving:

$$\sum (y_i - x_i) = \sum \left[(K_i - 1) \cdot z_i / (\beta \cdot K_i + 1 - \beta) \right] \tag{8}$$

for β, assuming that the K_i are constant. Then the phase mole fractions are evaluated from:

$$y_i = K_i \cdot x_i = \frac{K_i \cdot z_i}{\beta \cdot K_i + 1 - \beta} \tag{9}$$

The thermodynamic models for the two phases are then used to update the K_i values and the process is iteratively repeated as follows:

$$\ln K_i^{(n+1)} = \ln K_i^{(n)} + \ln \left(f_i^L / f_i^V \right)^{(n)} \tag{10}$$

An important point to highlight regarding the SL equation is the fact that usually the chemical potential is used instead of fugacity. Therefore, Eq. (10) is written as a function of chemical potential instead of fugacity and is expressed as:

$$\ln K_i^{(n+1)} = \ln K_i^{(n)} + \ln \left(\frac{\mu_i^L}{R \cdot T} - \frac{\mu_i^V}{R \cdot T} \right)^{(n)} \tag{11}$$

This equation can be solved for each iterative with a low computational effort. In successive substitution procedures, ratios of mole fractions (K factors) are used in an inner loop to find phase amounts and mole fractions using Eqs. (8) and (9). The K factors are updated in an outer loop until the values of the chemical potentials or fugacities in all coexisting phases are sufficiently close to each other. Sometimes the successive substitution algorithm itself can become oscillatory and divergent. The

analysis of Heidemann and Michelsen (1995) shows that this kind of instability of successive substitution can occur whenever any of the equilibrium phases shows strong negative departures from ideality. Heidemann and Michelsen (1995) suggested a simple "damping" procedure in the K factor updating algorithm that would result in a convergent process. With the strategy used in this work it is not necessary to employ this "damping" procedure regardless the equation of state used.

For polymer systems some numerical aspects are independent of the equations of state employed. One aspect to highlight is that composition must be always expressed as a function of weight fraction instead of mol fraction. Therefore, the mass fraction must be converted in mol fraction to calculate the component fugacity or chemical potential, but the returned composition must be expressed in mass fraction after equilibrium calculation. Different from what consider Koak and Heidemann (1996), we have kept the standard choice of volume fraction in SL equation as the composition variable. Another point regarding to solution of Eq. (8) is that the bisection method is used because it is known that the root β is encompassed by the interval (0, 1). The boundaries of this interval are the initial guess for root finding. Adjustable characteristic parameters of the various models are fitted to industrial operational data, and the Nelder-Mead simplex method (Nelder and Mead, 1964) is applied to minimize the following objective function:

$$OF = \sum_{i=1}^{nc} \left(w_{i,H}^{cal} - w_{i,H}^{\exp} \right)^2 + \sum_{i=1}^{nc} \left(w_{i,L}^{cal} - w_{i,L}^{\exp} \right)^2 \tag{12}$$

where w_i is the mass composition of component i, the subscripts H and L are related to the heavy and light phases, respectively, the superscripts *cal* and *exp* are related to the calculated and experimental values, respectively, and nc is the number of components. Usually, the convergence criterion of this estimation is 10^{-8} for the value of objective function. Some numerical features depend on the equation of state. Due to the high values assumed inside the algorithm of fugacity calculations for polymers at high pressure with SRK equation, an important contribution is regarding the calculation of constant equilibrium K_i. Once this property is monitored in update value by using Eq. (10), it is immaterial to divide both phase fugacities by pressure, and this numerical detail is crucial to solve this type of simulation problem. The constant equilibrium K_i is monitored and updated using Eq. (10), that can be expressed as a function of fugacity or fugacity coefficient. The advantage of using fugacity coefficient is that the pressure is the same for both phases. Therefore, dividing both fugacity coefficients by pressure in Eq. (10) is a simple but key numerical detail to overcome this kind of simulation convergence problem.

4. Results

Eight resins are investigated and their features are presented in Table 1, with the HPS operational conditions. The polymer is assumed to be monodisperse for all systems.

Table 1. LDPE molecular weight and HPS separator operational conditions

Resin	*MW* (g/mol)	T_{HPS} (°C)	P_{HPS} (bar)
LDPE-1	335000	280.29	250.06
LDPE-2	295200	278.80	251.90
LDPE-3	181000	230.29	249.88
LDPE-4	166000	253.03	250.55
LDPE-5	340000	237.68	249.87
LDPE-6	322000	277.17	249.54
LDPE-7	236600	247.92	254.00
LDPE-8	426300	265.33	253.28

Table 2 shows the binary interaction parameters between polymer and each other component: ethylene, ethane, propylene and propane. Due to similarity between molecular interactions the same value of binary interaction parameters is used for all pairs of components.

Table 2. Binary interaction parameters of EOS used for HPS simulation (LDPE resin)

Resin	$SL1_{ij}$	$SL2_{ij}$	UQ_{ij}^{*}	WS_{ij}	UQ_{ij}	λ
LDPE-1	1.5975	1.5075	0.0010	0.9440	0.0010	-0.3663
LDPE-2	1.5975	1.5075	0.0010	0.9296	0.0010	-0.2255
LDPE-3	2.0199	1.0099	0.0010	0.8237	0.0013	0.2235
LDPE-4	2.0199	1.0099	0.0010	0.9000	0.0011	-0.0703
LDPE-5	2.0199	1.0099	-0.0032	-0.4543	0.0028	3.6224
LDPE-6	1.5074	1.5033	0.0068	0.5873	0.0016	0.8922
LDPE-7	2.0000	0.9900	-0.0297	-2.6891	0.0025	3.0426
LDPE-8	1.5299	1.4850	0.0023	0.0560	0.0020	2.0432

Simulations of the HPS for the LDPE resins and comparison of the equation of state performances indicate that the Sanchez-Lacombe EOS is the best choice, as illustrated by Figure 1.

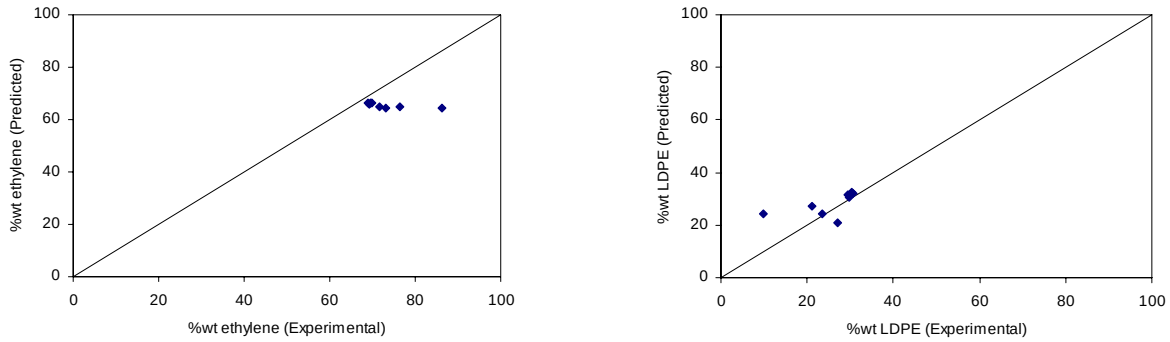


Figure 1. Predicted bottom composition of HPS separator versus experimental values for Sanchez-Lacombe EOS.

5. Conclusion

Simulation of high-pressure separators (HPS) in an industrial low density polyethylene (LDPE) process is used to evaluate Sanchez-Lacombe (SL) and Soave-Redlich-Kwong (SRK) models. Several numerical aspects were examined for each EOS. The results show that Sanchez-Lacombe equation of state is more appropriate to overcome the difficulties associated with the influence of mass and mole composition in the phase equilibrium, the general optimization strategy for the estimation of the EOS interaction parameters, the algorithms of flash calculation, the influence of low and high polymer concentration, and the convergence criteria at each step calculation. SRK-LCVM and SRK-WS are unable to describe the HPS overhead composition, since the results indicate only the presence of ethylene in this stream. Regarding the LDPE composition in the bottom stream, SRK-LCVM and SRK-WS give similar results but are not so accurate as SL.

Acknowledgements

The authors are gratefully acknowledged to FAPESB, CAPES, CNPq and FAPERJ.

References

G. Bogdanic, J. Vidal, 2000, A Segmental Interaction Model for Liquid–Liquid Equilibrium Calculations for Polymer Solutions. Fluid Phase Equilibria, 173, 241.

C. Boukouvalas, N. Spiliots, P. Coutsikos, N. Tzouvaras, D. Tassios, D., 1994, Prediction of Vapor-Liquid Equilibrium with the LCVM Model: a Linear Combination of the Vidal and Michelsen Mixing Rules Coupled with the Original UNIFAC and the t-mPR Equation of State, Fluid Phase Equilibria, 92, 75.

G.M.N. Costa, T. Dias, M. Cardoso, Y. Guerrieri, F.L.P. Pessoa, S.A.B. Vieira de Melo, M. Embiruçu, 2008, Prediction of Vapor-Liquid and Liquid-Liquid Equilibria for Polymer Systems: Comparison of Activity Coefficient Models, Fluid Phase Equilibria, 267, 140.

R.A. Heidemann, M. Michelsen, 1995, Instability of Successive Substitution, Ind. Eng. Chem. Res. 34, 958.

N. Koak, R.A. Heidemann, 1996, Polymer–Solvent Phase Behavior near the Solvent Vapor Pressure, Ind. Eng. Chem. Res. 35, 4301.

G.M. Kontogeorgis, V.I. Harismiadis, A. Fredeslund, D.P. Tassios, 1994, Application of the van der Waals Equation of State to Polymers I. Correlation, Fluid-Phase Equilibria, 96, 65.

J.A. Nelder, R. Mead, 1964, A Simplex Method for Function Minimization, Computer Journal, 7, 308.

I.C. Sanchez, R.H. Lacombe, 1976, An elementary molecular theory of classical fluids: pure fluids, J. Phys. Chem., 80, 2352.

G. Soave, 1972, Equilibrium Constants from a Modified Redlich-Kwong Equation of State, Chem. Eng. Sci., 27, 1197.

L. Wang, 2007, Calculation of vapor-liquid equilibria of polymer solutions and gas solubility in molten polymers based on PSRK equation of state, Fluid Phase Equilibria, 260, 105.

D.S.H. Wong, S.I. Sandler, 1992, A Theoretically Correct Mixing Rule for Cubic Equation of State, AIChE J., 38, 671.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling the hydrodynamics and mass-transfer phenomena for sedimentary rocks used for Flue Gas Desulfurization. The effect of temperature.

Cataldo De Blasio, Jarl Ahlbeck and Tapio Westerlund

Åbo Akademi University , Technical faculty Process Design & Systems Engineering Biskopsgatan , 20500 ABO; E-mail address: cdeblasi@abo.fi

Abstract

Sedimentary rocks, such as limestone, are widely utilized in Flue Gas Desulfurization (FGD) processes. The study of the dissolution for solid particles involved in FGD is therefore significant for process design and plant operation. The rate of dissolution affects the cost of makeup and waste disposal. For this reason a method to test different qualities of raw materials can give us a better understanding of the desulfurization performance. One method to test the reactivity takes into account the utilization of diluted strong acids like hydrochloric acid [1]. In the present study the mass transport phenomena involved in batch stirred tank reactors (BSTRs) are modeled in presence of hydrochloric acid and the experiments were taken at different temperatures in order to demonstrate the reliability of the method. The Surface Renewal Time, also called Surface Time of Exposure (TOE), is a quantity describing the life-time of the separation surface between the solid and liquid faces. The method gives an estimation of the temperature effect over the solid particles dissolution in BSTR and can be used to evaluate the reactivity and the diffusivity values of different raw materials.

Keywords: Gas Desulfurization, Limestone Reactivity, Mathematical modeling, Transport Phenomena.

1. Objective

The present work takes into consideration a generalized form of the penetration theory first introduced by Higbie (1935). The present approach has been utilized for studying the absorption by simultaneous diffusion and chemical reaction for solid particles [5]. It was taken into account the modeling of the mass transport phenomena involved in BSTR systems. By evaluating the physical parameters involved, the ratio of convective to diffusive mass transport and the ratio of momentum and mass diffusivity, it was possible to relate the diffusivities and mass transfer coefficients at three different temperatures for raw materials in terms of a defined TOE. The generalization of the penetration theory for solid particles in BSTR gave us useful results; this encouraged us to utilize the same approach for evaluating the diffusion parameters at different temperatures. The objective of this study was to verify the method for evaluating the quality of solid particles used in flue gas desulfurization [8].

2. Theory

2.1. Main reactions involved

Previous studies [1] have considered the reactivity of calcium carbonate in presence of strong acids like hydrochloric acid. In the following table the main reactions involved are given:

Table 1. Main reactions involved in limestone-acid systems

Fenomena	Reactions
Dissolution of limestone	$CaCO_3(s) \rightarrow Ca^{2+} + CO_3^{2-}$
Carbonate ions react with the hydronium ions according to:	$H_3O^+ + CO_3^{2-} \rightarrow HCO_3^- + H_2O$
Carbonic ions react with hydronium ions	$HCO_3^- + H_3O^+ \rightarrow H_2CO_3 + H_2O$
Final products	$H_2CO_3 \rightarrow CO_2(g) + H_2O$

The system described above can be considered as one overall first order reaction [1]. We denote with C_{ai} the saturated concentration of the solute at the surface (Z = 0), C_a is the concentration of C_a^{2+} ions at a generic distance Z, t designates the time and k_r is the reaction rate constant. D_a represents the diffusivity of the considered dissolving element. The solute concentration into the bulk solution is described by the following differential equation:

$$\frac{\partial C_a}{\partial t} = D_a \frac{\partial^2 C_a}{\partial z^2} - k_r C_a \tag{1}$$

The boundary conditions are:

$$C_a = C_{ai}, z = 0, t > 0$$
$$C_a = 0, z > 0, t = 0$$
$$C_a = 0, z = \infty, t > 0$$

It can be proved, utilizing the Duhamel's theorem and by considering the analogy between heat and mass transfer for the same mixing conditions [2], that the final relation for the concentration of the diffusing component into the bulk solution is described by:

$$C_a = \frac{1}{2} \cdot C_{ai} \cdot \left[\exp\left(-z\sqrt{\frac{k_r}{D_a}} \right) + \exp\left(z\sqrt{\frac{k_r}{D_a}} \right) \right] \cdot \left[erfc\left(\frac{z}{2\sqrt{D_a \cdot t}} - \sqrt{k_r t} \right) \right] \tag{2}$$

The term *erfc[z, t]* in Eq.(2) gives the complementary Error Function of *erf[z, t]*. Considering the analogy between the heat and mass transfer and applying this theory to the present system, the dissolution of solid particles into the bulk solution can be described in terms of mass transfer per unit time as follows:

$$\dot{m} = \left(\pi \cdot d_p^2\right)\left(C_{ai} - C_\infty\right)\left(\frac{\langle k_c \rangle}{d_p}\right)\left[\frac{(3\cdot\pi)^{\frac{2}{3}}}{2^{\frac{7}{2}}\Gamma\left(\frac{4}{3}\right)}\right]\left(\mathrm{Re}\cdot Sc\right)^{1/3} \tag{3}$$

Where C_∞ is the solute concentration in the liquid region, d_p is the diameter corresponding to one particle having the total surface of reaction, k_c is the mass transfer coefficient, The term between the square brackets is considered to be a numerical value. [3], *Re* and *Sc* are respectively the Reynolds and the Schmidt number. In our case the Sherwood number is evaluated from:

$$\langle k_c \rangle = Sh\left(\frac{D_a}{d_p}\right) \tag{4}$$

From Eq.(3) it appear clear that the mass transfer per unit time is a function of the Sherwood number, this result is useful because Sherwood numbers have been evaluated for different kind of chemical reactors. More accurate values were obtained for diverse Schmidt numbers and for our system it is well described as follows [3]:

$$Sh = 2 + \left(0.4\cdot \mathrm{Re}^{0.5} + 0.06\cdot \mathrm{Re}^{2/3}\right)\cdot Sc^{0.4}\left(\frac{\mu_\infty}{\mu_w}\right)^{0.25} \tag{6}$$

The above formula is valid for $0.35 < Re < 8\cdot10^4$ and for $0.7 < Sc < 380$. The viscosity of the liquid medium is considered to be the same near the surface and at a reasonable distance, in this way the term $(\mu_\infty/\mu_w)\ \hat{}\ 0.25 = 1$. At this stage we consider again the general equation for mass transport related to the dissolving component and taking into account the reactor volume V we obtain:

$$V\frac{dC_a}{dt} = -V\cdot k_r\cdot C_a + \langle k_c \rangle\left(C_{ai} - C_a\right)\cdot S \tag{7}$$

Where the term S represents the surface of reaction. With $\tau = v/(kc\cdot S)$ and $k_r = 1/t_r$ for fast reactions, finally we get:

$$C_a = \frac{C_{ai}}{\tau}\left\{1 - \exp\left[-t\left(\frac{1}{t_r} + \frac{1}{\tau}\right)\right]\right\}\cdot\frac{1}{\frac{1}{t_r} + \frac{1}{\tau}} \tag{8}$$

2.2. The surface penetration theory.

It was assumed that the separation layer between two phases is considered to be a sum of differential volumes; in this case we consider that each of these volumes has a different age or life-time at the surface. At the solid liquid interface we consider a velocity profile to be linear, for this reason a particular velocity exists that belongs to a non-zero thickness. This layer has the same behavior of a solid if we consider the mass transport phenomena involved. The mass flux is described by the following formula [4]:

$$\langle N_i \rangle = \frac{1}{L}\int_0^L \sqrt{\frac{D_a \cdot v_i}{\pi \cdot x}}\Delta C \cdot dx \tag{9}$$

Where v_i is the solid-liquid velocity at the interface and L is the integration length along the x axis, ΔC is the already mentioned difference in concentration of the dissolving component. The mean coefficient of mass transfer is:

$$\langle k_c \rangle = 2\sqrt{\frac{D_a \cdot v_i}{\pi \cdot L}} \tag{10}$$

The quantity L/v_i represents a time θ over the integration length; this value can be obtained from experiments and is known as TOE or Surface Renewal Time [4].

3. Materials and methods

For particle size measurements the Malvern 2600 particle sizer system has been used. The equipment includes a batch stirred tank reactor and a laser-beam diffractometer where the sample particles were analyzed in order to get a particle size distribution. The model involved takes into account the variation of the particle size distribution derived from the distribution of the scattered light energy using the Fraunhofer diffraction theory and the spherical model, furthermore the procedure expected synchronized *pH* and particle size measurements. This theory is valid for opaque particles whose radius is large compared with the laser wavelength. For each size range, the limestone fraction in volume was given. In this way, it has been possible to evaluate the change in the sample volume as a function of time. The following table describes the sedimentary rocks used and their provenience, the sample tested showed the best quality in terms of TOE and mass transfer coefficient [8].

Table 2. Tested raw materials: general description

Sample	Description	Provenience
Ljj 01c	Paleoproterozoic magmatic limestone	Halpanen, Finland
Ljj 09c	Paleoproterozoic metamorphic limestone, marble	Parainen, Finland
Ljj-P	1900 million-year-old limestone metamorphosed to marble.	Parainen Finland

The sample indicated as Ljj-P is from the Parainen quarry in south west Finland. The carbonate rock is a 1900 million-year-old limestone metamorphosed to marble during the Svecofennian orogeny 1830 million years ago (metamorphosed limestone). Mineralogically it is almost pure calcite and texturally an even grained marble. The temperatures at which the experiments were conducted were: 25°, 38° and 42° Celsius.

4. Results and discussion

The concentration C_a is a function of: time, distance from the particle's surface, $z = 0$; it is also a function of the reaction rate constant (1/sec). Fig. 1 shows the concentration function for a determinate value of k_r Eq.(2), is this way the total amount dissolved is described by the surface integral over the time and the integration length. Eq.(8) is well described by Fig. 2 for assigned values of the diffusivity and surface of reaction. The mathematical treatment refers to a single shrinking particle.

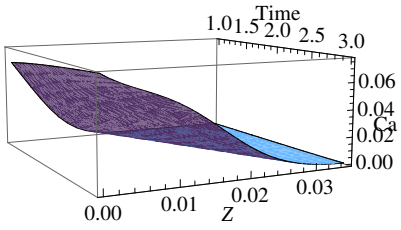


Fig. 1. Concentration in space and time at a determinate value of *kr* Eq.(2).

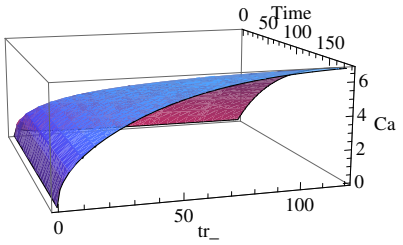


Fig. 2. Three dimensional representation of Eq.(8) C_a is expressed in micro-moles.

The following two dimensional figures are more suitable for the reporting of the Surface Time of Exposure (TOE) and the mass transfer coefficient as a parametric function of the temperature, experiments were conducted at three different temperatures: 25°, 38° and 42° Celsius, indicated respectively with the symbols: *, # and + in Fig. 3, 4, 5, 6. Experiments were conducted at equal time intervals. The mathematical fitting is shown by the dashed lines for each temperature. The results related to samples Ljj-09c and Ljj-P, Table 2, are reported as follows:

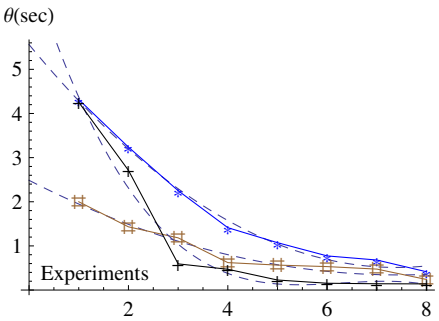


Fig. 3. TOE at 25°C, 38°C, 42°C. Sample: Ljj_P. Range: 150-250µm.

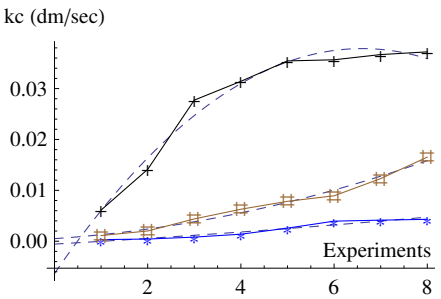


Fig. 4. K_c at 25°C, 38°C, 42°C. Sample: Ljj_P. Range: 150-250µm.

The experiments have been done immediately after the addition of hydrochloric acid to the system, the data reported also show the lag time needed in order to settle TOE and k_c to the minimum and maximum value respectively. Furthermore the derivatives of TOE and k_c at the first seconds indicate dependence from the temperature as well, the Surface Time of Exposure specify a punctual value while the derivative denotes an initial acceleration due to the diffusion phenomena. The following results are also in agreement with the above mentioned principles and in accordance with our expectations.

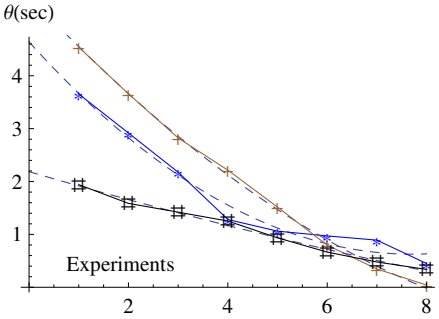


Fig. 5. TOE at 25°C, 38°C, 42°C. Sample:Ljj_09c. Range: 63-106µm

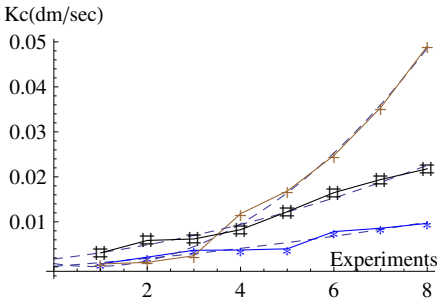


Fig. 6. K_c at 25°C, 38°C, 42°C. Sample: Ljj_09c. Range:63-106µm.

5. Conclusions

The experiments were conducted during a short period of time when the calcium carbonate is able to utilize all the hydronium ions in a fast way. When the concentration of carbonic acid (weak acid) is reasonably high, the second step of reaction shown in table one is less influent. See Fig.3. The hydronium concentration profile was in agreement with this mechanism. By evaluating the time of exposure and the mass transfer coefficient it has been demonstrated that the method is sensible to temperature variations; the experimental results are perfectly in accordance with the theory: at higher temperatures we have a more rapid decrease for the Surface Renewal Time and a greater increase for the mass transport coefficient, furthermore the above plots show also the non perfect mixing obtained among the initial experiments. The slope and the absolute values of TOE and mass transfer coefficient describe accurately the effect of temperature on the dissolution process.

6. Acknowledgements

The authors would like to thank the Fortum Foundation for the financial support. The Heat Engineering Laboratory and the Process Design & Systems Engineering Laboratory, Technical faculty, Åbo Akademi University, Åbo. The department of Geology at University of Turku for the samples provided. Professor Heikki Ruskeepää, University of Turku, Department of Mathematics.

• References

1. Toprac, A.J. G.T. Rochelle (1982). Env. Prog.1, p 52.
2. Carslaw, H. S.; Jaeger, J. C. (1959). Conduction of heat in solids. Oxford: Clarendon Press.
3. R. Byron Bird; E. Stewart; E.N. Lightfoot. (2002) "Transport phenomena" Wiley and sons.
4. R. Higbie. (1935). Transactions of the American Institute of Chemical Engineers 31. pp. 365–389.
5. P.V. Danckwerts (1950), Trans. Faraday Soc., vol. 46, pp. 300-304
6. Kasper Lund, H. Scott Fogler. (1973). "The dissolution of dolomite in hydrochloric acid" Chem. Eng. Sci. Vol. 28.pp. 691-700.
7. P. V. Danckwerts, Unsteady-state diffusion or heat-conduction with moving boundary. Trans. Faraday Soc. , 1950, 46, 701 – 712
8. C. De Blasio et al. (2008). Proceedings to the 8th International conference of Eco Balance, Tokyo. pp. 545-548.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Energy Optimization of a Kraft Pulp Mill

Araceli García, María González, Rodrigo Llano-Ponte, Jalel Labidi

Department of Chemical and Environmental Engineering – University of the Basque Country, Plaza Europa 1, San Sebastián 20018, Spain

Abstract

New regulations, economical and environmental requirements urge industrial companies to continuously improve the efficiency of their production technologies. The application of process integration tools for process retrofit is a must whenever energy or raw material intensive processes are considered. Since pulp and paper processes use a large amount of energy and water, the application of process integration techniques in this area is essential and can contribute to the protection of the environment to a great extent. CadSim Plus simulation software has been used to simulate several sections of a kraft pulp mill. Based on the process flow sheet and the developed simulation, pinch analysis was undertaken using the software Aspen HX-Net 2006. Potential of energy savings has been identified and the feasibility of the associated heat exchangers network (HEN) modification has been evaluated by the process simulation. The combination of both computer tools, has allowed to carry out the study of energy optimization of the process, maintaining the conditions of operation and evaluating the feasibility of the possible changes in the design of the heat exchangers network that allow minimizing the energy consumption.

Keywords: Process simulation, optimization, pinch analysis, heat exchange networks.

1. Introduction

Optimization and integration of industrial processes has become and important research topic, because it allows to analyze the factors that improve the production yield, as well as to study the mechanisms and possible changes that can minimize energy and water consumption and pollutant emissions (Nurmesniemi, 2007; Salama, 2008; Wang, 2009). For this purpose, the process simulation has turned out to be a very useful tool for the study and optimization of industrial processes, because it allows to analyze the behavior of the implied systems and to evaluate the feasibility of real conditions of operation with a previous knowledge of the process performance. Simulation software can be used to study the behavior of several process and equipments (Cardoso, 2009) and to analyze the general development of different processes, optimizing the operation conditions.

The paper industry is one of the most affected by the increasing need to reduce resources consumption since in most cases the production of pulp and paper generates an energy surplus that is often not conveniently used (Jönsson, 2008, Svensson, 2008b). In addition, this sector consumes a large amount of water and the feasibility of circuit closure, allowing optimal use of internal flows, and its influence on the process yield is becoming a subject widely studied. Usually the circuit closure and the use of water from the process imply an increase in the minimum temperature and the cooling requirement inside the plant, so that the processes of heat exchange in the mill are modified (Foo, 2008).

The highest energy consumption of the pulp mills is localized in the recovery zone. The evaporation process can consume up to 30% of the energy generated in the form of steam (Axelsson, 2008), and other equipment, such as paper machine which also requires large amounts of live steam (Sivill, 2005; Geldermann, 2006). The optimization of these processes is a key point to improving the overall process.

On the other hand, a way to optimize the use of energy is to redesign the Heat Exchange Networks (HEN), but also other equipment, such as the recovery boiler, where the optimization of the combustion process of black liquor results in a higher yield (Hektor, 2007). This case is very interesting because, although it is not a main part of the purpose of the pulp and paper production, the recovery process makes the mill self-sufficient of reagents and energy, making profitable the production. So, a by-product, like black liquor, results in an energy resource within the company as its combustion generates high-pressure steam.

Therefore, it is possible to convert a sector such as paper industry, usual consumer of energy and water resources, in a producer not just steam and electricity (Svensson, 2009), but also other high value-added products, such as tall oil and lignin of the black liquor (Svensson, 2008b).

Many authors have used simulation of pulp and paper processes (Castro, 2004; Turon, 2005) in conjunction with other tools to achieve the processes integration and energy optimization. Pinch analysis has been proved to be a powerful technique for this purpose, being employed for several analyses as heat exchanger networks studies, energy savings at different levels of the processes and minimizing water consumption of the mill (Wising, 2005; Nordmann, 2006).

In this work combination of process simulation and Pinch analysis has been used to achieve the energy optimization of kraft pulp mill.

2. Process Description

The studied pulp Kraft mill (Figure1) produces 400 Tones of pulp per day and uses a batch pulping process that comprises several stages. The mixture of wood chips and whit liquor are heated up to 170 ºC in heat exchanger using medium pressure steam (MPV, 15 bar). During the pulping process a part of the black liquor is recirculated and cooled from 144 ºC to approximately 90 ºC. For this purpose two heat exchangers are employed, using condensates and warm water from de water circuit. The obtained pulp from the digesters is then screened and washed with hot condensates (63 ºC aprox.) and a recirculated partial stream of washing liquor, since the other one is used in the digesters for pulp washing during one of the cooking stages.

Eight effects compose the evaporation plant, where the black liquor obtained from pulping process containing 17 % TDS (Total Dissolved Solids) is first preheated using low pressure steam (LPV, 5 bar) and then concentrated (up to 60 % TDS) before its combustion in the recovery boiler, where the resulting inorganic matter is used to reconstitute the white liquor employed in the pulping stage. The mill uses a second furnace burning the wood bark and producing high pressure steam (HPV, 55 bar) which used to produce electricity, MPV and LPV used within the pulp mill. The air required for the combustion is preheated by low and medium pressure steam.

The condensates generated in the evaporation plant and in all the heat exchangers that use live stream, as well as the water consumed in the surface condenser are classified according to their pollutant content (for the condensates: clean, dirty or very dirty) and temperature (warm, up to 45 ºC, or hot, up to 65 ºC) for internal reuse in the pulp mill. Clean condensates are heated by using LPV and used in the recovery boiler and to

adjust the stream extracted from the turbine; hot condensates and water are recirculated to several sections (causticizing, washing …). The water used to condensate the vapor of last evaporation stage is previously cooled in the cooling tower and mixed with a fresh water stream.

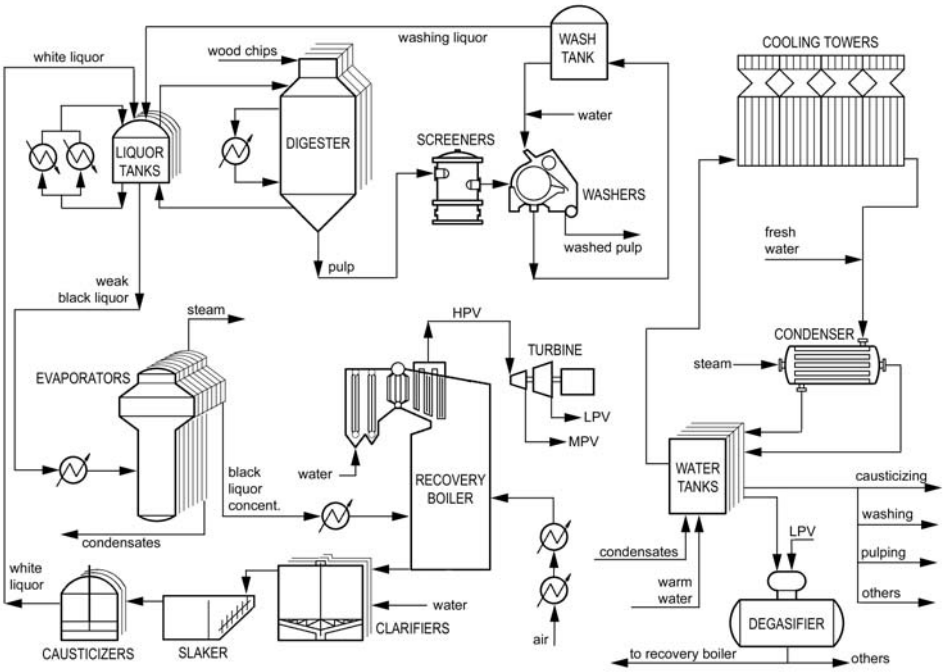


Figure 1. Diagram of the Pulp Mill sections studied

3. Kraft Process Simulation

The simulation of the process described has been developed by means of the software CadSim Plus (Aurel Systems), specifying main streams, equipment characteristics and process conditions. The simulation has been validated with several operation conditions. The results have been collected for a later energy study. The most important data are described in the Table1.

The recovery boiler generates HPV, which feed the turbine producing 28 T/h of MPV at 250 ºC, 62 T/h of LPV and around 7700 kW of electrical power. The live steam obtained is consumed in several sections of the mill (15.7 T/h MPV in pulping process, 14,5 T/h LPV for black liquor evaporation…).

4. Pinch Analysis

The Pinch Analysis has been carried out with the HX-Net software (AspenTech). The data extraction has been made using the process simulation and taking into account only streams with a minimum heat flow of 1 MW. The HEN has been analyzed taking into account the utilities consumption (steam and water), the characteristics of heat exchangers (area) and related costs (operation and investments).

The HEN of the Base Case consists of 9 heat exchangers and represents a total heat requirement of 18.17 MW. This configuration was compared to two case studies:

Table 1. Simulation results for the main sections of the Kraft Pulp Mill.

Pulping process		Recovery Furnace	
Chip load	62 T/h	Air consumption	128 T/h
Pulp production	160 T/h (90 % water)	HPV generation	61 T/h
White Liquor consumption	47 T/h (14 % TDS)	Water feeding	85 T/h
Black Liquor generation	144 T/h (17.8 % TDS)	Smelt production	7.8 T/h
Steam consumption	18.6 T/h	Steam consumption	8.8 T/h
Evaporation process		Causticizing process	
Concentrated Black Liquor	40 T/h (61 % TDS)	CaO consumption	3.1 T/h
Steam consumption	16.1 T/h	Turbine	
Cooling water to condenser	357 T/h (19 ºC)	Power generation	7721 kW

Scenario 1: new design of the HEN. In this case only utilities have been employed to satisfy the heat exchange requirements (live steam for heating and fresh water for cooling).

Scenario 2: resulting from the automatic optimization by means of HX-Net software, which implies a simple reorganization of the streams, changing the flow ratios in the branch designed.

The results of the analysis, energy consumptions (heating and cooling), and number of heat exchangers and related operating – investment costs, as well as the energy generated in the turbine (electricity) are reported in Table 2. The HEN initial configuration and two alternatives are represented in Figure 2.

Table 2. Energy savings and related operating and investments costs for the different scenarios.

	Base Case	Scenario 1		Scenario 2	
		value	% respect Base Case	value	% respect Base Case
Heating (MW)	18.17	18.17	-	18.17	-
Cooling (MW)	-	14.78	-	-	-
Area (m^2)	7327	3574	- 51	7148	-2.5
HE units	8	7	-12	8	-
Investments cost Index (cost)	$1.69 \cdot 10^7$	$7.77 \cdot 10^6$	-54	$1.63 \cdot 10^7$	-3
Operating cost Index (cost/year)	$1.82 \cdot 10^6$	$3.30 \cdot 10^6$	81	$1.82 \cdot 10^6$	-
Total Cost Index (cost/year)	$6.14 \cdot 10^6$	$5.28 \cdot 10^6$	-14	$5.99 \cdot 10^6$	-2

Base Case configuration takes advantages of most of the water generated in the process (condensates and warm water) for cooling some streams of the process. For Scenario 1 in which the water reusability has been removed, using only fresh water for cooling leading to an increase of energy requirement of 14.78 MW (49 %) compared with base case. however, the Scenario 1 represents the better economical option, since it implies a 14 % reduction of the total cost, but entails higher water consumption. The HEN configuration of Scenario 2 results in a minimizing of the total annual cost due the lower need of heat exchange area.

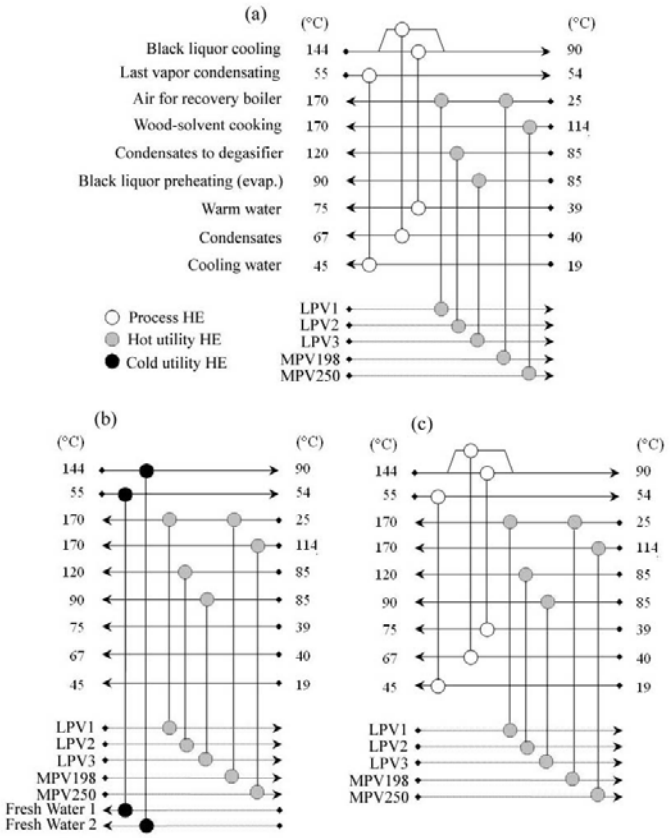


Figure 2. The HEN configuration for different scenarios: (a) Base Case, (b) Scenario 1 (without water reuse), (c) Scenario 2 (HE cost optimization).

5. Conclusions

The combination of simulation and pinch analysis has been proved to be a good tool for the study of energy optimization in real processes.

The comparison between three possible scenarios has been made, concluding that for the studied pulp mill the HEN optimization means a change in the network configuration and subsequently of the process. This can be developed by means of the use of process simulation.

Acknowledgements

The authors would like thank the Department of Agriculture, Fishing and Feeding of the Basque Government for supporting financially this project through the scholarship of Mrs. Garcia.

References

E. Axelsson, M.R. Olsson, T. Bertsson, 2008, Opportunities for Process-Integrated Evaporation in a Pulp Mill and Comparison with a Softwood Model Mill Study, Applied Thermal Engineering, 28, 2100-2107.

M. Cardoso, K. de Oliveira, G.A. Costa, M.L. Passos, 2009, Chemical Process Simulation for minimizing Energy Consumption in Pulp Mills, Applied Energy, 86, 1, 45-51.

J.J. Castro, F.J. Doyle III, 2004, A Pulp Mill Benchmark Problem for Control: Problem Description, Journal of Process Control, 14, 17-29.

D.C.Y. Foo, Y.H. Chew, C.T. Lee, 2008, Minimum Units targeting and Network Evolution for Batch Heat Exchanger Network, Applied Thermal Engineering, 28, 2089-2099.

J. Geldermann, M. Treitz, O. Rentz, 2006, Integrated Technique Assessment based on the Pinch Analysis Approack for the Design of Production Networks, European Journal of Operationa Research, 171, 1020-1032.

E. Hektor, T. Berntsson, 2007, Future CO_2 Removal from Pulp Mills – Process Integration Consequences, Energy Conversion and Management, 48, 3025-3033,

J. Jönsson, I.L. Svensson, T. Berntsson, B. Moshfegh, 2008, Excess Heat from Kraft Pulp Mills: Trade-offs Between Internal and External Use in the Case of Sweden – Part 2: Results for Future Energy Market Scenarios, Energy Policy, 36, 4186-4197.

R. Nordman, T. Berntsson, 2006, Design of Kraft Pulp Mill Hot and Warm Water Systems – A New Method that maximizes Excess Heat, Applied Thermal Engineering, 26, 363-373.

H. Nurmesniemi, R. Pöykiö, R.L. Keiski, 2007, A Case Study of Waste Management at the Northern Finnich Pulp and Paper Mill Complex of Stora Enso Veitsiluoto Mills, Waste Management, 27, 1939-1948.

A.I.A. Salama, 2008, Heat Exchanger Network Synthesis based on Minimum Rule Vairations, Applied Thermal Engineering, 28, 1234-1249.

L. Sivill, P. Ahtila, M. Taimisto, 2005, Thermodynamic Simulation of Dryer Section Heat Recovery in Paper Machines, Applied Thermal Engineering, 25, 1273-1292.

E. Svensson, T. Berntsson, A.B. Strömberg, 2008, Benefits of Using and Optimization Methodology for Identifying Robust Integration Investment under Uncertainty – A Pulp Mill Example, Energy Policy, doi: 10.1016/j.enpol.2008.10.024.

E. Svensson, T. Berntsson, A.B. Strömberg, M. Patriksson, 2009, An Optimization Methodology for Identifying Robust Process Integration Investments under Uncertainty, Energy Policy, 37, 680-685.

I.L. Svensson, J. Jönsson, T. Berntsson, B. Moshfegh, 2008, Excess Heat from Kraft Pulp Mills: Trade-offs Between Internal and External Use in the Case of Sweden – Part 1: Methodology, Energy Policy, 36, 4178-4185.

X. Turon, J. Labidi, J. Paris, 2005, Simulation and Optimisation of a High Grade Coated Paper Mill, Journal of Cleaner Production, 13 (15), 1424-1433.

U. Wising, T. Berntsson, P. Stuart, 2005, The Potential for Energy Savings when Reducing the Water Consumption in a Kraft Pulp Mill, Applied Thermal Engineering, 25, 1057-1066.

Z. Wang, D. Zheng, H. Jin, 2009, Energy Integration of Acetylene and Power Polygeneration by Flowrate-Exergy Diagram, Applied Energy, 86, 372-379.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Theoretical Nucleation Study of the Combined Effect of Seeding and Temperature Profile in Cooling Crystallization

David Widenski[a], Ali Abbas[b], Jose Romagnoli[a]

[a]*Chemical Engineering Department, Louisiana State University, Baton Rouge, USA*
[b]*School of Chemical and Biomolecular Engineering, University of Sydney, Sydney, Australia*

Abstract

Crystallization is a widely used unit operation used for the production of pharmaceuticals, fertilizers, and fine chemicals. One of the most common types of crystallization is cooling crystallization. In cooling crystallization, the solution is cooled to generate supersaturation causing the formation and growth of crystals. The properties of the product are dependent on the shape of the supersaturation curve. Nucleation will occur at high rates if supersaturation increases to excessive levels, this being an undesired outcome. A commonly used crystallization objective is to produce large crystals with minimum fines production under conditions of low nucleation rates. A typical approach to minimize nucleation known as "programmed cooling" is to utilize an optimum temperature profile. Recent research works have favored initial seed characteristics over programmed cooling for the production of unimodal crystal size distributions (CSD). In this paper, a theoretical seed chart is developed for combined seeded-cooling crystallization via fundamental crystallization kinetic modeling of potassium chloride (KCl). Fundamental analysis here shows that joint cooling and seeding optimization of cooling crystallization gives superior performance to just optimizing the seed, and that the current trend of experimentally optimizing the seed is undermined by model-based optimization approaches.

Keywords: crystallization, seed chart, modeling, optimization, cooling.

1. Introduction

Crystallization is a widely used unit operation used for the production of high-purity compounds for pharmaceuticals, fertilizers, and fine chemicals. One of the most commonly used techniques is cooling crystallization, and is the technique that is considered in this paper. In cooling crystallization the solution inside the crystallizer is cooled to a predetermined temperature profile. This cooling causes generation of supersaturation which in turn causes both the formation and growth of crystals. In order to prevent excessive nucleation from occurring it is important to maintain the solution's supersaturation below the primary nucleation metastable limit. Rapid increases in supersaturation exceeding the metastable limit will result in uncontrollable primary nucleation. The first method to minimize nucleation was to implement programmed cooling profiles (Mullin & Nyvlt, 1971; Mullin & Jones, 1974). Although this method helped reduce nucleation, it was not successfully repressed.

Another way to minimize nucleation is to seed the crystallizer with already formed crystals. These crystals minimize nucleation because any supersaturation that is formed is used to grow the seed crystals. However, it is imperative to add the correct initial amount of seed to the crystallizer in order to get sufficiently grown crystals. If the seed loading is too large, the seed will not grow sufficiently and if on the other hand it is too small there will not be enough seed to suppress primary nucleation. Jagadesh et al. (1996) investigated the effect of crystal seed on the final CSD profile. They successfully showed that a suboptimal cooling profile such as natural cooling can be utilized to produce unimodal CSD's for the potash alum system. They introduced the experimental seed chart to assist in the proper selection of seed size and mass. Once either seed size or mass is fixed, the seed chart can be used to find the corresponding seed size or mass that will maximize seed growth. Kubota et al. (2001) performed further studies producing a seed chart for potassium sulphate. They showed that using an optimum amount of potassium sulphate seed, a unimodal CSD can be produced even using suboptimal natural cooling which was something Jones and Mullin (1974) could not achieve with programmed cooling for the same system. Since then Chung et al. (1999), Choong and Smith (2004), Sarker et al. (2006) and Nowee et al (2007), have performed optimization studies to find joint optimal temperature and seed profiles for various systems.

In this paper, a theoretical seed chart is developed for a combined seeded-cooling system via fundamental crystallization kinetic modeling of potassium chloride (KCl). Fundamental analysis here shows how model-based optimization of seed mass, seed size, and the temperature profile gives superior results over the current trend of experimentally optimizing the seed.

2. Crystallization Model

The crystallization model is comprised of a solubility model, a population balance, and nucleation and growth kinetics. The solubility of potassium chloride in water was correlated using a quadratic equation over a temperature range of 0-100 °C from data taken from the CRC Handbook of Chemistry and Physics (Lide, 2006). A population balance for a constant volume crystallizer with growth independent of crystal size, and negligible agglomeration is used. The population balance is discretized into 1000 equal intervals and solved through backward finite differences. The nucleation kinetics was broken down into homogeneous nucleation (1), heterogeneous nucleation (2), surface nucleation (3), and attrition (4) with equations taken from Mersmann (2001). Further information about the parameters and variables used in these equations are detailed in Mersmann (2001). Potassium chloride crystallization is reported as being diffusion limited so crystal growth (5) is modeled as a mass transfer process using a mass transfer coefficient (Lopes & Farelo, 2006).

$$B_{hom} = 1.5 \text{ x } D_{AB}(C\ N_A)^{7/3}\sqrt{\frac{\gamma_{CL}}{kT}}\frac{1}{C_c N_A}\ \exp\left(-\frac{16}{3}\pi\left(\frac{\gamma_{CL}}{kT}\right)^3\left(\frac{1}{C_c N_A}\right)^2\frac{1}{(v\ln S)^2}\right) \quad (1)$$

$$B_{het} = \left(\frac{1}{2\pi}a_{for}d_m HE_{ad}(CN_A)^{\frac{7}{3}}\sqrt{\frac{f\gamma_{CL}}{kT}}V_m\right)\times \left(\frac{D_{surf}\ sin\,\theta}{r_c}HE_{ad}d_m^{\frac{3}{2}}(CN_A)^{\frac{1}{6}} + 3\pi D_{AB}(1-cos\,\theta)\right)exp\left(-\frac{4}{3}\pi f\frac{\gamma_{CL}}{kT}r_c^2\right) \quad (2)$$

$$B_{surf} = 9E\frac{D_{AB}}{d_m^4 L_{32}}\varphi_T \, exp\left(-\pi\frac{(\gamma_{CL}d_m^2/kT)^2}{v\,ln\,S}\right) \quad (3)$$

$$B_{attrit} = 7x10^{-4}\varphi_T\frac{H_V^5}{\mu^3}\left(\frac{\Gamma}{K}\right)^{-3}\frac{\pi^2\rho_c\bar{\varepsilon}N_v}{2k_vPo}\frac{N_{a,eff}}{N_{a,tot}}\eta_w^3\eta_g \quad (4)$$

$$G = \frac{D_{AB}}{3\rho_c\bar{L}}\left(2 + 0.8\left(\frac{\bar{\varepsilon}\bar{L}\rho_s^3}{\eta^3}\right)^{1/5} Sc^{1/3}\right)\Delta C \quad (5)$$

In order to investigate the effect of the temperature profile, three different temperature profiles were considered. These three profiles represent three popular cooling profiles, linear, natural, and programmed cooling, and are shown in Figure 1. For each temperature profile the temperature was decreased from 55 °C to 10 °C over 30 min.

To investigate the effect of the size of the seed, three different seed sizes were evaluated 75, 250, and 500 microns with various mass loadings. The seed loadings were carefully chosen to occur where the critical seed loading would be for each different seed size. Since it has been proven by Kubota et al. (2001) that the critical seed loading is dependent on the seed size, the seed loadings were specified differently for each seed size. The 75, 250, and 500 micron seed loadings ranged from 0.1-1, 1-40, and 10-100 g of KCL per kg of H_2O respectively. The seed was then added at the beginning of each simulation before cooling was initiated. Stochastic experimental simulations comprised of 25 seed loadings were subsequently executed using the gPROMS modeling package (Process Systems Enterprise, UK) for each unique size and temperature profile combination.

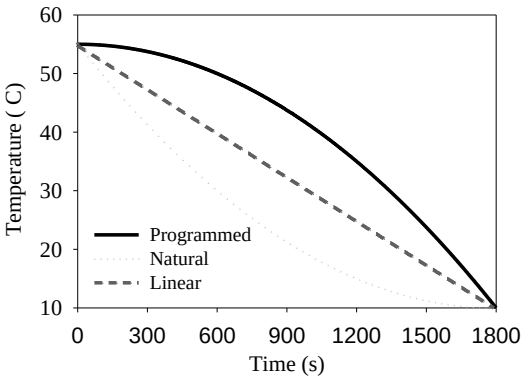


Figure 1: Implemented Temperature Profiles

3. Seed Chart Results

By organizing the data into a seed chart, as shown in Figure 2, it can be seen that the simulations produced a very similar result to the experimental seed charts published by Jagadesh et al. (1999) and Kubota et al. (2001). Each data point on the seed chart represents the ratio of the final volume mean size to its initial volume mean size for a given seed size, seed loading, and temperature profile. Similar to their seed charts, ours shows that the smaller seed sizes grow more compared to their initial size and require less loading than the larger sizes. It also shows that the temperature profile does not

have an effect on the final crystal size until the seed loading approaches the critical seed amount for that profile. Crystal growth follows the ideal growth line until the seed loading becomes insufficient and nucleation occurs. This causes the deviation of the curves away from the ideal growth line. Since this deviation is dependent on the temperature profile it follows that the critical seed amount is a function of the temperature profile. This is because some cooling profiles are superior to others at minimizing nucleation and maximizing the final crystal mean size. Thus, if the objective is to maximize the growth of the crystals then both the temperature profile and the seed loading are important.

While inspecting the seed chart it may seem that it is always better to seed small sizes. This is not always true. It is true that small seed sizes will grow more than the larger seed sizes due to the limited amount of mass available in the solution. However, if the objective is to make large crystals it may be necessary to seed moderate to high seed sizes at the expense of higher seed loading. For example the 75 micron seed was only able to grow to 330 microns before the temperature profile affected the size. However, the 250 and 500 microns seed sizes were able to grow to 425 and 630 microns respectively regardless of the temperature profile.

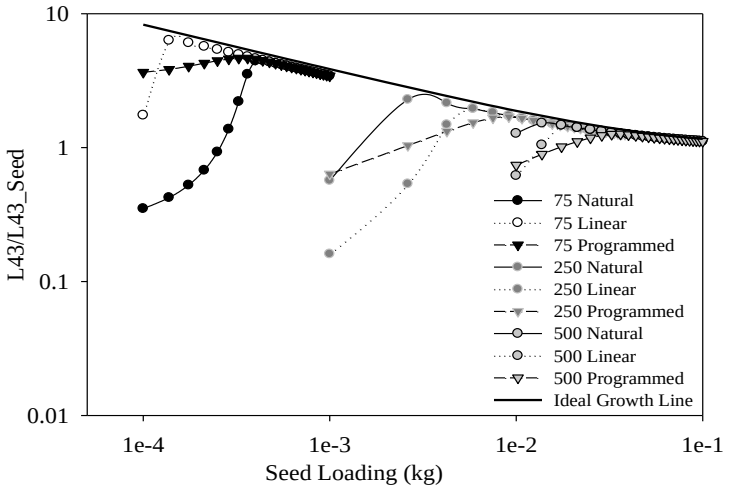


Figure 2: Seed Chart of KCl for Different Seed Sizes and Cooling Profiles.

4. Joint Seeding and Cooling Optimization

As shown earlier there are benefits to optimizing the temperature profile, namely larger crystal size and less seed loading. The chosen objective function in this study is the minimization of the zeroth moment. This is the same as minimizing the total amount of nucleated particles. This objective function should ensure a unimodal distribution. The four constraints used are the desired final crystal size, the maximum cooling rate, and the initial and ending temperatures. The optimization determines two control variables, the initial seed loading and the temperature profile. The temperature profile was discretized with 30 one minute control intervals. The three optimized temperature profiles for the 75, 250, and 500 micron seeds are displayed in Figure 3.

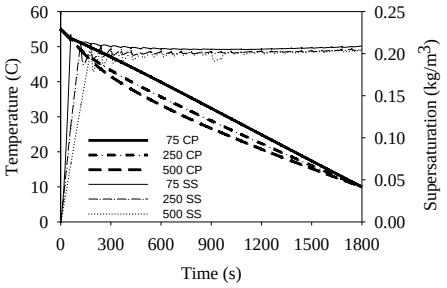


Figure 3: Optimized Temperature & Supersaturation Profiles

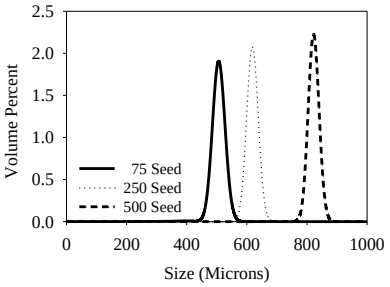


Figure 4: Optimized Volume Percent CSD

All three profiles are very similar in shape. They all have a quench cooling section followed by linear cooling. The initial quench cooling has the effect of quickly raising the supersaturation to the metastable limit, whereas the subsequent linear cooling keeps it at that limit until the end of the batch. This ensures maximum growth over the batch. All three seed sizes grew to larger sizes with the optimized temperature profiles than with the previous profiles used in generating the seed chart as shown in Table 1. The optimized size is 7-8% larger than the size possible from using the best seed chart temperature profile/ seed mass combination, and 23-33% larger than seeding the crystallizer at the critical seed mass, which corresponds to temperature profile independent operation. In addition, all three final volume percent CSD's shown in Figure 4 have unimodal distributions which show nucleation was successfully suppressed. Due to the optimized temperature profiles less seed loading was also required. These optimization studies show the benefit of optimizing the temperature profile and the seed conditions, and that joint optimization of the seed loading and temperature profile is superior to the seed chart in creating unimodal distributions of large mean size.

Table 1: Optimization Results

Seed Size (microns)	Seed Loading (kg)	Optimized L_{43} (microns)	Seed Chart Best L_{43} (microns)	Temperature Profile Independent L_{43} (microns)
75	1.26×10^{-4}	505	474	336
250	2.71×10^{-3}	617	572	425
500	1.11×10^{-2}	822	761	631

5. Conclusion

When the seed chart was introduced in the late 1990's it was a very useful tool. However, with the recent advances of both crystallization modeling and computational efficiency its usefulness is undermined. The disadvantage of the seed chart is that it does not optimize the temperature profile, and many experiments are required to

generate it. If another temperature profile is desired then more experiments need to be done to add that temperature profile to the seed chart. This is laborious and time consuming. However, an accurate crystallization model developed from carefully planned experiments cures these limitations. The crystallization model will work for any temperature profile or seed size. The crystallization model can be used to optimize both the seed loading and the temperature profile for any objective function which the seed chart cannot do. Experimental time spent creating a seed chart is better spent toward the development of a crystallization model. In conclusion, the seed chart has been a very useful tool for crystallization processes, but the advantages in crystallization modeling and model-based optimization make it the preferred method for the future.

Acknowledgement

This work was jointly supported by the Australian Academy of Science and the National Science Foundation's East Asia and Pacific Summer Institute Program.

References

Choong, K.L.; Smith, R., 2004, Optimization of Batch Cooling Crystallization, Chemical Engineering Science, 59, 313-327.

Chung, S.R.; Ma, D.L.; Braatz, R.D., 1999, Optimal Seeding in Batch Crystallizers, Canadian Journal of Chemical Engineering, 77, 590-596.

Hojjati H.; Rohani S., 2005, Cooling and Seeding Effect on Supersaturation and Final Crystal Size Distribution (CSD) of Ammonium Sulphate in a Batch Crystallizer, Chemical Engineering and Processing, 44, 949-957.

Jagadesh D., Kubota N., Yokota M., Sato A., Tavare N.S., 1996, Large and Mono-Sized Product Crystals from Natural Cooling Mode Batch Crystallizer, Journal of Chemical Engineering of Japan, 29, 5, 865-873.

Jagadesh D., Kubota N., Yokota M., Doki N., Sato A., 1999, Seeding Effect on Batch Crystallization of Potassium Sulphate under Natural Cooling Mode and a Simple Method of Crystallizer, Journal of Chemical Engineering of Japan, 32, 4, 514-520.

Jones, A.G.; Mullin, J.W., 1974, Programmed Cooling Crystallization of Potassium Sulphate Solutions, Chemical Engineering Science, 29, 105-118.

Jones, A.G., Optimal Operation of Batch Cooling Crystallization, 1974, Chemical Engineering Science, 29,1075-1087.

Kubota N., Doki N., Yokota M., Sato A., 2001, Seeding Policy in Batch Cooling Crystallization, Powder Technology, 121, 31-38.

Lide, D.R., 2006, CRC Handbook of Chemistry and Physics 87th ed., Taylor & Francis Group: Boca Raton, FL.

Lopes, A.; Farelo. F, 2006, Growth Kinetics of Potassium Chloride-I: Pure Aqueous Solutions, Journal of Crystal Growth, 290, 213-219.

Mersmann, A., 2001, Crystallization Technology Handbook, 2nd ed., Marcel Dekker: New York City, NY.

Mullin, J.W.; Nyvlt, J., 1971, Programmed Cooling Crystallization of Batch Crystallizers, Chemical Engineering Science, 26, 369-377.

Nowee S.M., Abbas A., Romagnoli J.A., 2007, Optimization in seeded cooling crystallization: A parameter estimation and dynamic optimization study, Chemical Engineering and Processing, 46, 1096-1106.

Sarker, D.; Rohani, S.; Jutan, A., 2006, Multi-Objective Optimization of Seeded Batch Crystallization Processes, Chemical Engineering Science 2006, 61, 5282-5295.

Xie. W.; Rohani, S.; Phoenix, A., Dynamic Modeling and Optimization of a Seeded Batch Cooling Crystallizer, 2001, Chemical Engineering Communications, 187, 229-248.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Decomposition Based Algorithm for the Design of Multipurpose Batch Facilities Using Economic Assessments

Tânia Rute X. M. Pinto[1,2], Ana Paula F. D. Barbosa-Povoa[2] and Augusto Q. Novais[1]

[1]*DMS – Departamento de Modelação e Simulação, INETI, Lisboa, Portugal*
Estrada do Paço do Lumiar, 1649-038 Lisboa, Portugal
E-mail address: tania.pinto@ineti.pt
[2]*CEG- Centro de Estudos de Gestão, IST, Lisboa, Portugal*

Abstract

In this paper we propose a multi-objective decomposition algorithm for the design of multipurpose batch plants, which avoids the direct solution of the fully detailed MILP model. Most of such design problems involve the maximization of the total revenue, as well as the minimization of the total cost. The way to deal with these two terms simultaneously is either to combine them into a single criterion (e.g., profit), or to define the efficient frontier that offers the optimal solutions by multi-objective optimization. In this work the latter approach, while more elaborate, was adopted, since the exploration of this frontier enables the decision maker to evaluate different alternative solutions. A combination of the proposed decomposition algorithm and the ε-constraint method is employed, which supports the application of this approach to perform economic assessments. The proposed algorithm allows the identification of the plant topologies, scheduling, equipment design and storage policies, subject to the plant's cost minimization and revenue maximization. A comparative analysis is presented between the detailed model proposed by Pinto *et al.* (2008b) and the proposed algorithm.

Keywords: Scheduling, Design, ε-constraint, Multipurpose batch facilities.

1. Introduction

In multipurpose batch facilities a wide variety of products can be produced via different processing recipes, by sharing all available resources, such as equipment, raw material, intermediates and utilities. In order to ensure that any resource in the design can be utilized as efficiently as possible, an adequate representation is necessary in order to address such type of problems without creating ambiguities in the process/plant mathematical model. The Resource-Task Network (RTN) is one of the possible adequate representations to describe the design of multipurpose batch plants as suggested by Pinto *et al*. (2008a).

Like most real-world problems, the design of multipurpose batch facilities involves multiple objectives, while most of the existing literature on the design problem has been focused on single objectives (Barbosa-Povoa 2007). Therefore, the multi-objective optimization of such problems requires the development of an adequate modelling approach, as a pre-requisite for the resulting models to be useful as decision making tools where trade-offs among objectives can be investigated. In order to guarantee

optimal solutions, most of the published mathematical design formulations consider a large number of potential equipment items, out of which a selection is made of those that will be incorporated into the optimal plant configuration. This factor and the diversity of products' recipes, gives rise to large MILP problems and consequently to an increasing computational burden. For this reason, effective solution tools are still an open area of research in the design of batch plants.

In this paper, we aim to overcome some of these computational difficulties and also to avoid the direct solution of the fully detailed design and scheduling MILP model. With that aim, we propose a decomposition algorithm that exploits the hierarchical structure of the problem: the original detailed design and scheduling model is decomposed into an upper-level (UL) and a lower-level (LL) models; the UL uses profit maximization as the objective function and selects the so-called complicated variables, which in the current case are the design variables responsible for equipment selection; the LL comprises both the equipment design and the scheduling problem.

Applying the ε-constraint method (Chankong and Haimes, 1983), the procedure iterates until the Pareto-optimum surface is completed. This allows the identification of a range of plant topologies, design facilities and storage policies that minimize the total cost of the system while maximizing revenue, subject to product demands and operational restrictions.

2. Decomposition Algorithm

To avoid the direct solution of the MILP model, we propose a decomposition algorithm (DA) that exploits the hierarchical structure of design/scheduling models, and employ it together with the ε−constraint method. As referred to previously, the original detailed design and scheduling model is decomposed into one UL design problem and one LL design plus scheduling problem. The UL determines the necessary equipment to satisfy the demand along the entire time horizon, with the design constraints being omitted, leading to a relaxation of the original, DM, or detailed problem and therefore to an upper-bound on profit. In the LL, the DM is solved by fixing the equipment obtained from the UL. The LL corresponds to a sub-problem of the original MILP problem, in a reduced space and at iteration *i*. It considers the selected subset of equipment obtained from the UL, producing a lower-bound on profit. The procedure iterates until the difference between the upper-bound and the lower-bound is less than a specified tolerance. To expedite the search, integer cuts are added to the UL to exclude previous solutions.

The DA is subject to an objective constraint function that defines the maximum cost available for the optimal solution. The initial restriction takes the value of the necessary cost to generate the maximum revenue (C_{max}). This restriction is decreased by small values, $\Delta\varepsilon$, and successive optimization steps ($i=1,\ldots,n$) undertaken until the intended range is covered and the Pareto front defined.

2.1. Integer Cut

If the upper-bound obtained from the UL and the lower-bound obtained from the LL do not lie within a pre-defined tolerance, it is necessary to obtain a new solution from the UL. Then an integer cut on the binary design variable, Δ_r^{UL}, where *r* characterizes the equipment obtained from the UL model, is added to the UL problem so as to avoid the solution in the previous iteration *i*. For iteration $s > k$, this is defined as:

$$\sum_{r\in U_1^k} \Delta_r^{UL} - \sum_{r\in U_0^k} \Delta_r^{UL} \leq \left|U_1^k\right| - 1 \qquad (1)$$

Where $U_0^k = \ r \mid \Delta_r^{UL} = 0$ and $U_1^k = \ r \mid \Delta_r^{UL} = 1$. The U_0^k and U_1^k are obtained from the optimal UL solution, in terms of the assignment variable in iteration *k*.

2.2. ε-constraint

In order to obtain the Pareto front, we used the ε-constraint method combined with the decomposition algorithm. The ε-constraint requires a cost constraint which is activated in both models (UL and LL) and imposes limits on the objective functions. The detailed model, LL, uses the objective functions FO1, which defines the revenue, and FO2, which reflects the cost associated with both equipment and operational tasks. To obtain each point in the Pareto front, the cost constraint is activated and decreased by $\Delta\varepsilon$:

$$F02 = C_{\max} - i\,\Delta\varepsilon \qquad i=0,\ldots,n. \tag{2}$$

The final decomposition algorithm is illustrated in Figure 1.

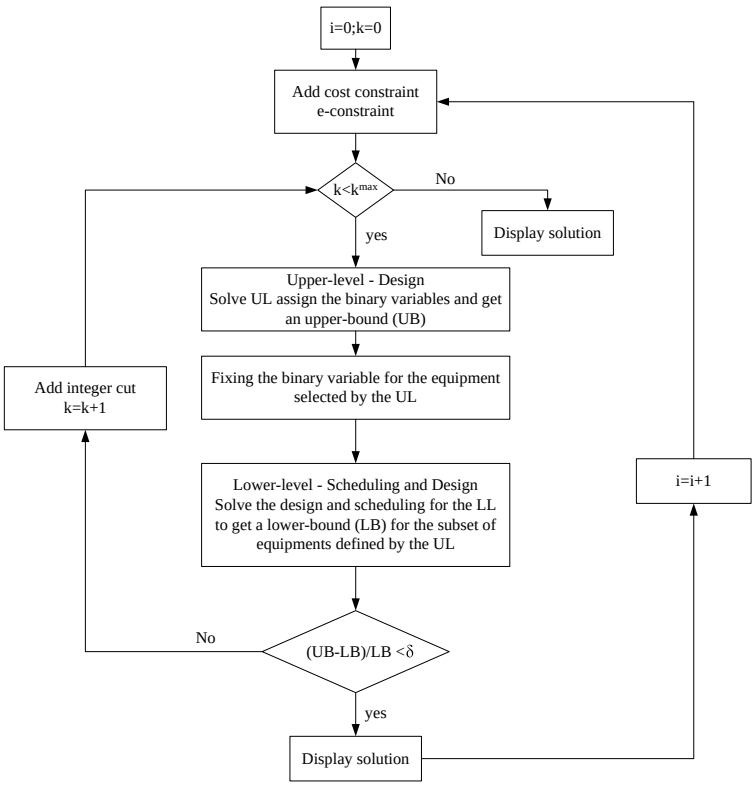


Figure 1: Flow chart for the decomposition algorithm.

2.3. Algorithm Steps

An overview of the proposed decomposition follows:

1. The indices *k* and *i* control, respectively, the inner loop for estimating individual points on the Pareto front and the outer loop for generating successive points.

Initially these are set at *k*=0 and *i*=0, while the upper-bound at UB=∞, the lower-bound at LB=- ∞ and a value δ assumed for the optimality tolerance.

2. Within the inner loop the cost constraint is activated; the MILP aggregated model solved for the upper-bound; the following sets defined:

$$U_0^k = \{ r \mid \Delta_r^{UL} = 0 \}\ ,\ U_1^k = \{ r \mid \Delta_r^{UL} = 1 \}\ ;\ \text{and}\ \ \Delta_r = \Delta_r^{UL} \quad \text{for} \quad r \in U_1^k\ ;$$

3. Still within the inner loop, the LL is solved; an optimal solution obtained, which is a lower-bound of the problem that defines the equipment design, scheduling and storage policy for each point of the Pareto front.

4. The convergence criterion $\frac{UB - LB}{LB} < \delta$ is tested.

If it is satisfied the solution corresponding to the lower-bound is the optimal solution and the outer loop is entered, where *i* defines the next point on the Pareto front.

5. If this convergence criterion is not satisfied an integer cut given by $\sum_{r \in U_1^k} \Delta_r^{UL} - \sum_{r \in U_0^k} \Delta_r^{UL} \leq \left| U_1^k \right| - 1$ is added to the UL and the inner loop controlled by *k* is again calculated.

3. Example results

In this example, the design of a multipurpose batch plant for a non-periodic mode of operation is performed. A production of [0; 170] tons of products S5, [0; 166] tons of S9 and S10, [0; 270] tons of products S6 and [0; 143] tons of products S11 is defined. Three raw materials, S1, S2 and S7, are used over the horizon of 24 h. The materials S5 and S6 are both intermediate and final products. S3, S4 and S8 are intermediate material. S3 and S8 are unstable and S4 is storable in V4 if necessary (Pinto *et al.* 2008b). There are available six main reactors (R1 to R6) and nine dedicated vessels. In terms of equipment suitability, only reactors R1 and R2 may carry out the two processing tasks, T1 and T2, while each storage vessel and reactors R3, R4, R5 and R6 are dedicated to a single State/Task. T1 may process S1 during 2 hours in R1 or R2; T2 may processes S2 during 2 hours in R1 or R2; T3 may process 0.5 of S3 and S4 during 4 hours in R3; T4 process 0.5 of S3 and S4 during 2 hours in R4; T5 process S6 during 1 hour to produce 0.3 of the final product S11 and 0.7 of S8 in R5, and finally T6 processes S8 during 1 hour in reactor R6 to produce the final products S9 and S10. The connections capacity range is assumed between 0 to 200 [m.u./m^2] at a fix/variable cost of 0.1/ 0.01 [10^3c.u.]. The capacity of R1, R2, R5 and R6, range from 0 to 150 [m.u./m^2], while the others range from 0 to 200 [m.u./m^2] (m.u. and c.u. are, respectively, mass and currency units).

3.1. Comparative analysis for the Example

In this section a comparative analysis of the performance is made, between the DM previously presented (Pinto *et al.* 2008b) and the currently proposed algorithm.

The optimality gap used for the UL was 7% followed by an additional 5000 seconds. For the LL and DM it was used a 5% of optimality gap and an additional 5000 seconds time for solution polishing. The results obtained for these specifications correspond to an approximation to the optimal set. The models characteristics are presented in Table **1**. Points A, B, C, D and E are those where there is a change caused by the addition of one or more main equipment units to the previous topology (Pinto *et al.* 2008b). Table 2 presents the optimal design of the main equipment for each point in terms of capacities, while Table 3 presents the final product and their quantity.

Table 1: Model characteristics.

Model	Bin.Var.	Tot.Var.	Equat.
DM	612	2582	4378
UL	228	1809	2478
LL	606	2583	4379

Table 2: Design for the main equipment using the decomposition algorithm.

Equipment	A	B	C	D	E
R1	76.2	93.3	103.4	141.2	120.3
R3	76.2	62.2	103.4	141.2	180.5
R4	-	155.5	129.3	140.5	159.1
R5	-	-	-	21.8	138.2
R6	-	-	-	15.3	96.8
V4	-	-	-	-	120.3
V5	76.2	-	51.72	170	170
V6	-	155.5	258.6	270	270
V9/V10	-	-	-	7.6	145.1
V11	-	-	-	6.5	124.4

Table 3: Quantities produced for each final product using the decomposition algorithm.

Final Products	A	B	C	D	E
S5	76.2	-	-	170	170
S6	-	155.5	258.6	270	270
S9	-	-	-	7.6	145.1
S10	-	-	-	7.6	145.1
S11	-	-	-	6.5	124.4

A comparative analysis between results obtained for the DM model (Pinto *et al.* 2008b) versus the DA model is shown in Table 4. The first column characterizes the different points on the Pareto curve. There were topology changes only in points A to E. The second column quantifies the relative performance of the two models. The one with the highest performance is identified on the third column. From the gap column it can be noted that the execution for the DM and the LL ends when the optimality gap reaches (5%). In run 3, the DM model reached the maximum CPU time available without attaining the optimal gap. Looking at the best performance column, we verify that in most runs the DA model shows a better performance than DM, except for runs 1 and 9, where the latter took around 12% and 10%, i.e. less CPU time than DA. For all runs the DA converged in the first iteration. Despite the computational results, the equipment design and amounts of final products remained equal for both models.

Table 4: Comparative analysis.

Run/Point	$(CPU_{DA} - CPU_{DM}) / CPU_{DA} \times 100$	Best performance	DM Cpu (s)	DM Gap (%)	DA Cpu (s)	LL Gap (%)	Iteration
1	12.145	DM	2.17	2.15	2.47	3.6	1
2/E	-261.27	DA	461.92	4.99	127.86	4.99	1
3	-256.691	DA	5000	5.77	1401.77	4.99	1
4	-168.921	DA	210.27	4.99	78.19	4.99	1
5	-324.125	DA	227.84	4.99	53.72	4.99	1
6/D	-102.481	DA	1843.45	4.99	910.43	5.0	1
7	-131.622	DA	596.45	5.0	257.51	5.0	1
8/C	-87.147	DA	415.0	4.99	221.75	4.99	1
9/B	10.271	DM	103.75	4.99	115.63	4.99	1
10/A	-99.78	DA	217.78	4.99	109	4.99	1

4. Conclusions

A multi-objective decomposition algorithm is proposed for the design of multipurpose batch plants. The model aims to overcome some performance difficulties from the direct solution of the fully detailed MILP model. A combination of the decomposition algorithm with the ε–constraint is applied allowing the definition of a range of topologies, schedules, design facilities and storage profiles in the vicinity of the efficient frontier.

The performance of the proposed methodology is illustrated with an example. As seen in Table 4 the proposed model presents an all-round better performance than the detailed one. Despite these results some future work must be undertaken to improve the optimality gap and the model performance.

References

Barbosa-Povoa, A. P. (2007). "A Critical Review on the Design and Retrofit of Batch Plants." Computers & Chemical Engineering 31(7): 833-855.

Pinto, T., A. Barbosa-Povoa and A. Q. Novais (2008a). "Design of Multipurpose Batch Plants: A Comparative Analysis between the Stn, M-Stn, and Rtn Representations and Formulations." Industrial & Engineering Chemistry Research 47(16): 6025-6044.

Pinto, T. R., A. Barbosa-Povoa and A. Q. Novais (2008b). Multi-Objective Design of Multipurpose Batch Facilities Using Economic Assesments. 18th European Symposium on Computer Aided Process Engineering (ESCAPE-18), Lyon, France.

Chankong, V.; Haimes, Y. (1983). Multiobjective Decision Making Theory and Methodology. Elsevier. New York.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Electrochemical engineering modelling of the electrodes kinetic properties during two-phase sustainable electrolysis

Ph. Mandin[1,*], R. Wüthrich[2] & H. Roustan[3]

[1]*LECA UMR 7575 CNRS-ENSCP-Paris6, ENSCP, 11 rue Pierre et Marie Curie, 75231 Paris cedex, France*
[2]*Department of Mechanical & Industrial Engineering, Concordia University, Montreal, Canada*
[3]*Alcan – Centre de Recherche de Voreppe -725 rue Aristide Bergès - BP 27 – 38341 Voreppe Cedex, France*

Abstract

During two-phase electrolysis for aluminium, hydrogen or fluorine industrial production, there are bubbles which are created at electrodes which imply a quite important electrical properties and electrochemical processes disturbance. Bubbles are motion sources for the electrolysis cell flow, and then hydrodynamic properties are strongly coupled with species transport and electrical performances. Bubbles presence modifies these global and local properties: the electrolysis cell and the current density distribution are modified. This disturbance can lead to the modification of the local current density and to anode effects for example. There is few works concerning the local modelling of electrochemical processes during a two-phase electrolysis process due to the difficulty of measurements in such electrolytes. Produced bubbles involve a great hydrodynamic acceleration at the electrode interface which leads to great chemical composition modification and, sometimes, also the mechanical the electrode consumption due to both erosion and flow accelerated corrosion. Nevertheless, effects like the anode effect, particularly expensive on the point of the process efficiency, should need a better understanding.
The goal of this proposition is to present the electrochemical engineering modelling of two-phase electrolysis properties at electrode vicinity. This work is due to the necessity of a better knowledge of the actual interface condition during electrolysis, for example to have a better process optimisation or electrode consumption prevention. The numerical simulation is performed from the single bubble scale to the macroscopic electrochemical cell. The present work shows both scales theoretical modelling and also performances changes during the two-phase electrolysis processes.

Keywords: Two-phase Electrolysis Modelling Electrochemical Engineering.

1. Introduction

Gas release and induced fluid flow at electrodes are characteristic for several electrochemical processes such as aluminium, fluorine, chlorine and hydrogen production. Many authors have investigated these processes. The two-phase phenomena at gas evolving electrodes are in general neglected because of the major difficulty to be correctly taken it into account. With the increasing interest in hydrogen production,

clean, sustainable aluminium production and fluorine production for nuclear industry, these processes have to be revisited for further optimization.
The classical alkaline water electrolysis has been chosen as a representative two-phase electrolysis process:

$$2\ H_2O + 2\ e^- \rightarrow H_2 + 2\ OH^-_{aq} \qquad E° = -0.83\ V/ENH \qquad (R1)$$

$$2\ OH^-_{aq} \rightarrow 0.5\ O_2 + H_2O + 2\ e^- \qquad E° = 0.40V/ENH \qquad (R2)$$

Note that the alkaline water electrolysis, with minimal cell voltage equal to 1.23V, is generally chosen as the reference process for industrial hydrogen production. Many researchers have studied and exposed the difficulties in the two-phase electrolysis processes.The main problem is the existence of several effects, surface and volume effects, at the bubble scale but as well at the macroscopic electrochemical cell or reactor scale. This is the reason why zero gravity experiments are planfied to avoid natural convection.

2. Experimental set-up

Figure 1 shows the used experimental one dimensional set-up with general geometrical parameters definition. The configuration is axis symmetric for all input parameters chosen. The counter electrode has a large diameter D = 70 mm and height H'=17 mm with a geometric surface area $S=\pi\ D{\cdot}H$ = 440 mm^2 if H=h=2mm (height of the electrolyte free surface). The stainless steel 316L, working pin electrode is cylindrical with diameter d = 0.4 mm and a geometric surface area $s= 0.25\ \pi\ d^2 + \pi\ d{\cdot}h$ = 5.5 mm^2 for an immerged length h=2mm.

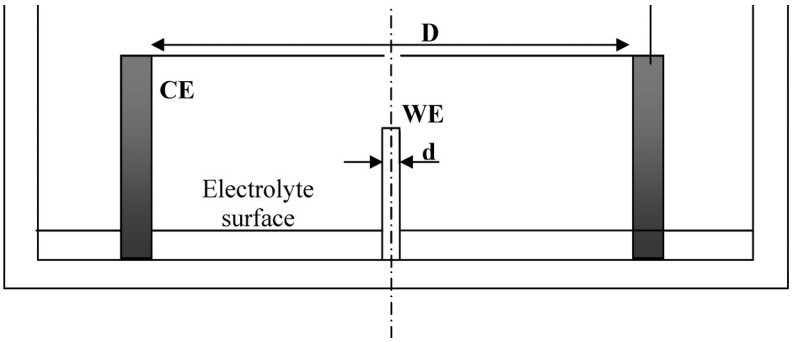


Figure 1: experimental set-up.

The cell voltage scan rate dV/dt, the electrolyte composition and temperature, the working electrode material and the counter electrode diameter factors have been explored.

3. Experimental results

Figure 2 shows an exemple of steady two-phase boundary layer which take place at the working electrode vicinity. The shape is said like an "inverted pyramid" due to the Archimede vertical forces applied on each bubbles, like the weight force.

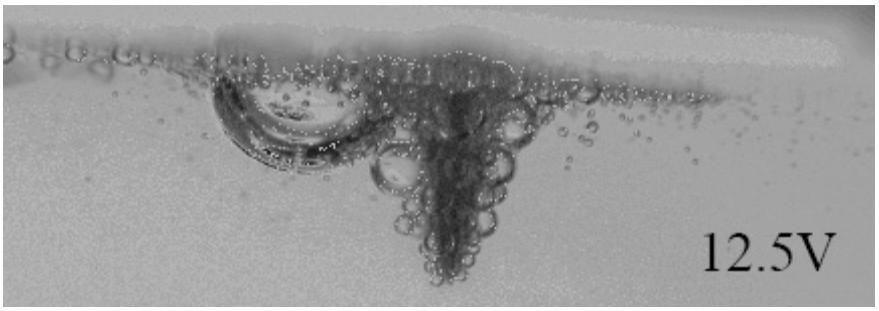


Figure 2: Bubble accumulation at 12.5 V around the working electrode during alkaline water electrolysis

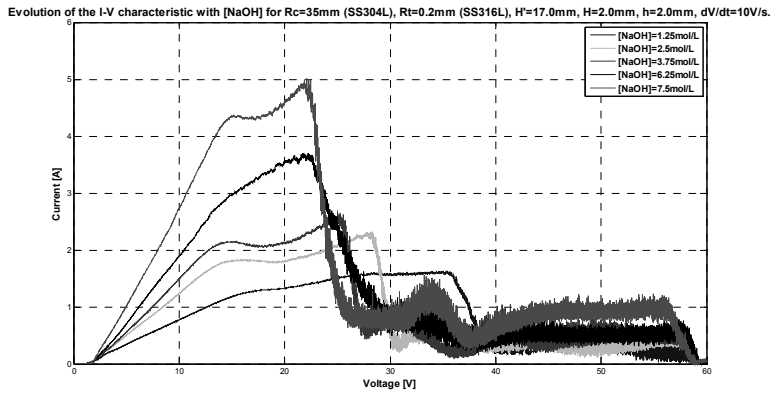


Figure 3: I-V characteristic evolution with the electrolyte NaOH concentration with stainless steel SS316L working electrode, D=70 mm, d=0.4 mm, h=H=2mm and dV/dt=10 V s^{-1}.

Figure 3 shows the effect of the electrolyte concentration upon the experimental resistance R_{exp}. The larger the concentration, the smaller the measured resistance is.
Table I shows these results quantitatively. The Bruggeman modelling is used to model the electrical resistance:

$$R_{theo} = \ln(1+2\delta/d)/(2\,\pi.h\sigma)+\ln(D/(d+2\delta))/(2\pi h.\sigma) \quad (1)$$

With: δ is the average two-phase boundary layer thickness (m), d and h the working electrode diameter and height (m), D the counter electrode diameter (m) and σ the electrical conductivity (S m^{-1}) given by Bruggeman law:

$$\sigma = \sigma^{\circ}.(1-\varepsilon)^{1.5} \quad (2)$$

with σ° the without gas (or bulk) liquid electrolyte conductivity (S m^{-1}) and ε the volume void fraction.
The minimum resistance is obtained for δ=0 m and is called R_{theo}, which is reported in table I.

As shown in table I, the under estimated theoretical resistance R_{theo} largely over-estimate the experimentally measured electrical resistance R_{exp} for all concentration explored. This primary modelling (only ohmic resistance description) is not sufficient to be realistic. For the two-phase electrolysis processes, the bubble motion rotational and translational, in the boundary layer, increases a lot the mass transfer in the working and

counter electrodes. The bubbles promote the mass transport. The equivalent impedance is not only ohmic, like for the Warbur impedance.

Table I: Evolution of the electrical resistance with the electrolyte concentration.

NaOH %wt	5	10	15	25	30
C° (mol/L)	1.25	2.5	3.75	6.25	7.5
$\sigma°$ (S m^{-1}) [I. Zaytsev (17)]	27	35	41	37	31
R_{exp} (Ω)	11.7	7.1	6	4.4	3
R_{theo} (Ω)	15.2	11.7	10	11.1	13.3

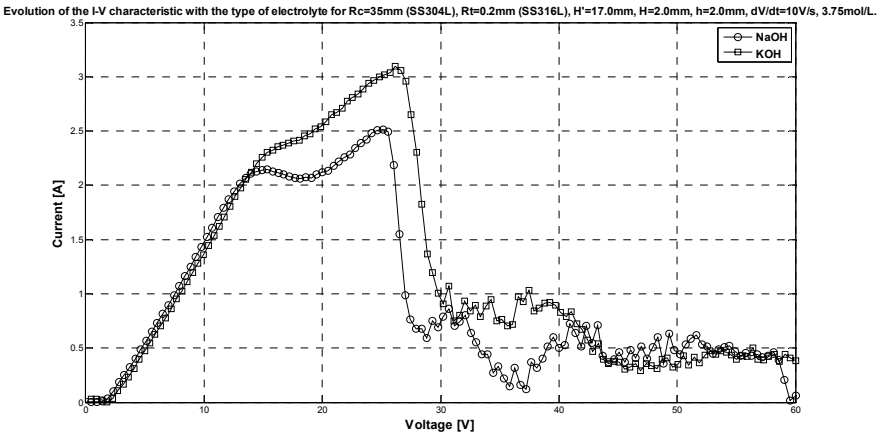


Figure 4: I-V characteristic evolution with the electrolyte species NaOH or KOH; concentration is 3.75 M with stainless steel SS316L working electrode, D=70 mm, d=0.4 mm, h=H=2mm and dV/dt=10 V s^{-1}.

Figure 4 shows that the electrical resistance is non dependent, for an identical molar concentration, of the electrolyte nature NaOH or KOH. The electrolysis cell resistance is non dependent with the electrical conductivity $\sigma°$, which is not in accord with theoretical expression (1). According with this expression the resistance should decrease with the electrical conductivity $\sigma°$.

The table II shows the evolution with concentration of the electrical conductivity $\sigma°$ for both species. This conductivity reaches a maximum at electrolyte concentration about 3.75 M. For smaller concentration, both conductivities are approximatively equal. But at 3.75 M, there is 34% difference between the conductivity values. Nevertheless, this large difference leads to identical resistance: once again, the ohmic primary description of the charge transport in inacurate for concentration around 3.75 M.

Table II: evolution of the electrical resistance with the electrolyte composition

Concentration [mol/L]	1.25	2.5	3.75	6.25	7.5
Conductivity NaOH [S/m]	27	35	41	37	31
Conductivity KOH [S/m]	28	39	58	58	44

Figure 5 shows an other contradiction of the ohmic-primary theoretical formulation for the electrical resistance. The electrolysis cell resistance is non dependent with the counter electrode diameter D, which is not in accord with theoretical expression (1). According with this expression the resistance should increase with diameter D.

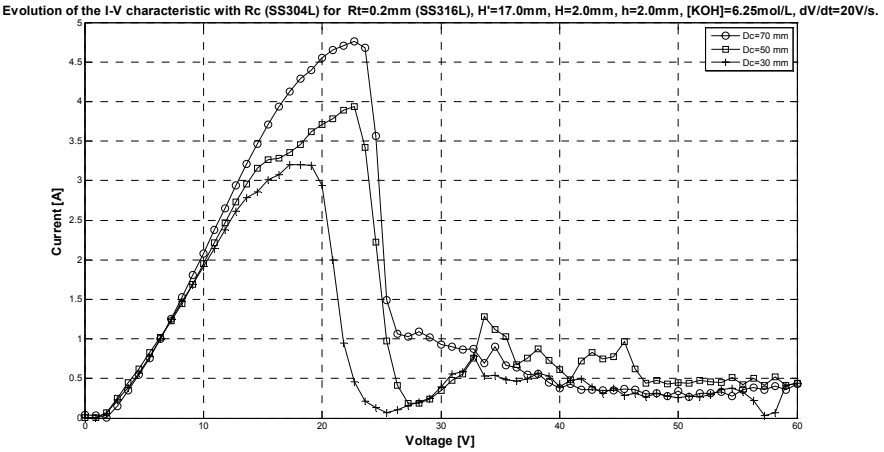


Figure 5: I-V characteristic evolution with the counter electrode diameter D: the electrolyte is NaOH the concentration is 6.25 M, with stainless steel working electrode, d=0.4 mm, h=H=2mm and dV/dt=10 V s^{-1}.

4. Conclusion

As it has been shown in this work, the electrical resistance obtained with 1G experiments doesn't obey the ohmic-primary theoretical modelling given in equation (1). Nevertheless, as shown in the second work of this session, for zero gravity experiments, this modelling is in good accord with experimental observation. Then, for two-phase electrolysis under normal Earth gravity, the induced natural flow has strong effect on the mass transport at electrodes. This effect is macroscopic (in the whole electrolysis cell), but also microscopic, at the bubble scale.

References

K. Azumi, T. Mizuno, T. Akimoto, and T. Ohmori, 1999, *J. Electrochem. Soc.* 146, p3374
J. Eigeldinger, H. Vogt, 2000, *Electrochimica Acta*, 45, 27 pp4449-4456
C. Gabrielli, F. Huet, R.P. Nogueira, 2005, *Electrochimica Acta*, 50, 18, pp3726-3736
L.J.J. Janssen, J.G Hoogland, 1970, *Electrochimica Acta*, 15, p1020
L.J.J Janssen, C.W.M.P Sillen, E. Barendrecht and S.J.D Van Stralen, 1984, *Electrochimica Acta*, 29, 5, pp633-642
H. Kellogg, 1950, *Journal of the Electrochemical Society*, 97, 4
S.M. Korobeinikov, A.V. Melekhov, Yu.N. Sinikh and Yu.G. Soloveichik, 2001, *High Temperature*, 39, 3, pp368-372

Ph. Mandin, J. Hamburger, S. Bessou, G. Picard, 2005, *Electrochimica Acta*, 51, 6, pp1140-1156
Ph. Mandin, H. Roustan, R. Wüthrich, J. Hamburger & G. Picard, 2007, *Simulation of Electrochemical Processes II*, Transactions on Engineering Sciences, WIT Press p73
Ph. Mandin, A. Ait Aissa, H. Roustan, J. Hamburger, G. Picard, 2008, *Chemical Engineering and Processing: Process intensification*, 47, pp1926-1932
H. Matsushima, T. Nishida, Y. Konishi, Y. Fukunaka, Y. Ito, K. Kuribayashi, 2003, *Electrochimica Acta*, 48, 28, pp4119-4125
H. Matsushima, Y. Fukunaka, K. Kuribayashi, 2006, *Electrochimica Acta*, 51, 20, pp4190-4198
H. Vogt, Ö. Aras, R. J. Balzer, 2004, *International Journal of Heat and Mass Transfer*, 47, 4, pp787-795
H. Vogt, R.J. Balzer, 2005, *Electrochimica Acta*, 50, 10, pp2073-2079
R. Wüthrich, Ch. Comninellis, H. Bleuler, 2005, *Electrochimica Acta*, 50, pp5242-5246
R Wüthrich, L.A. Hof, A. Lal, K. Fujisaki, H. Bleuler, Ph. Mandin, G. Picard, 2005, *J of Micromech. Microeng.*, 15, pp268-275
I. Zaytsev, G. Aseyev, 1992, *Properties of Aqueous Solution of Electrolytes*, CRC Press, Boca Raton, Ann Arbor, London, Tokyo

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Extending mass transfer correlations for packed three-phase distillation simulation: A hierarchical parameter estimation methodology

Liang Chen,[a,b] Shuqing Wang,[a] Jens-Uwe Repke,[b] Günter Wozny[b]

[a] *State Key Lab of Industrial Control Technology, Institute of Cyber-Systems and Control, Zhejiang University, Hangzhou 310027, P.R. China*
[b] *Institute of Process and Plant Technology, TU Berlin, Berlin10623 , Germany*

Abstract

Simulation of three-phase (vapor-liquid-liquid) distillation is still lacking of reliable mass transfer correlations to describe the rate-based (or nonequilibrium, NEQ) mass transfer process. The objective of this paper is to extend existing mass transfer correlations for three-phase distillation simulation in packed column. The transfer parameters in these correlations are classified into two categories: packing-related parameters and flow-related parameters, which are estimated using a (double-layer) hierarchical methodology based on our comprehensive experimental researches. The experimental database consists of 162 different data sets (both two-phase and three-phase distillations) of n-Butanol/Water/n-Propanol system for three different packings (Montzpack B1-350, Raschig Super-Ring 0.3, Rombopak 9M). The estimation problem with multi-data-set and multi-packing turns out to be a large-scale optimization subjected to a large number of nonlinear model equations. Three widely used correlations in packed column (i.e. Billet & Schultes correlation, Shi & Mersmann correlation, Rocha correlation) are adjusted for rate-based mass transfer calculation of three-phase distillation. Substantial improvements on the simulation predictive ability are observed by using the NEQ model with the adjusted correlations.

Keywords: three-phase distillation, parameter estimation, mass transfer correlation.

1. Introduction

There are many examples of extractive and azeotropic distillation processes, in which binary azeotropes are broken by adding a third component that creates a vapor-liquid-liquid (three-phase) heterogeneous distillation. Modeling three-phase distillation, however, is a complicated task due to the strong influence of the second liquid on the mass and heat transfer (Repke & Wozny, 2000).

Traditional method of modeling three-phase distillation assumes equilibrium condition among all three existing phases, which however is seldom realized in real processes. The non-equilibrium (NEQ) condition is dominated by the rate-based inter-phase mass transfer rates, which can be calculated based on film model and theory of multi-component mass transfer (Taylor & Krishna, 1993), i.e.

$$\dot{V}_i = N_i a_{ph} = \left(J_i + z_i \sum N_i\right) a_{ph} \tag{1}$$

The molar diffusion flux J_i is the most dominating term of mass transfer flux N_i, and can be computed by Taylor and Krishna (1993) method using the Maxwell-Stefan equation.

$$\mathbf{J} = -c_t[\mathbf{k}]\nabla\mathbf{z} = c_t[\mathbf{B}]^{-1}[\mathbf{\Gamma}](\mathbf{z}^{\mathbf{I}} - \mathbf{z}) \tag{2}$$

where the thermodynamic matrix [**Γ**] can be derived from certain activity coefficient model (NRTL, UNIQUAC etc.), and the [**B**] matrix has the elements

$$B_{ij} = -z_i\left(\frac{1}{\beta_{ij}} - \frac{1}{\beta_{in}}\right) \text{ and } B_{ii} = \frac{z_i}{\beta_{in}} + \sum_{k=1(i\neq k)}^{n}\frac{z_k}{\beta_{ik}} \quad i,j = 1,\cdots,n-1 \quad (i \neq j) \qquad (3)$$

where the binary mass transfer coefficients β can be calculated using empirical mass transfer correlations in multi-component distillations. For packed columns, one widely used correlation to describe vapor/liquid mass transfer on random packing was developed by Shi & Mersmann (1985). The first overall investigation about structured (gauze) packing was conducted by Bravo et al. (1985). The results were called SRP (I) model, which was lately improved and updated by Rocha et al. (1993, 1996), i.e. SRP (II) model. An attractive characteristic of SRP (II) model is that the mass transfer and hydraulic performance of packing are related by liquid holdup. Billet & Schultes (1999) also considered this point in their model for mass transfer calculation, filled with either random or structured packings, in the entire loading range. Detailed formulations of the above mass transfer correlations are given in table 1.

Table 1. A summary of mass transfer correlation for packed column

Name	β_V, β_L, a_e
Shi & Mersmann	$\frac{a_e}{a_P} = 0.76 m d_h^{1.1}\frac{U_L^{\theta_1}\nu_L^{\theta_2}}{1-0.93\cos\gamma}\left(\frac{\rho_L}{\sigma_L g}\right)^{\theta_3}\frac{a_P^{0.2}}{\varepsilon^{0.6}}$ $\beta_V = K_V D_V / d_h (Re_V)^{\theta_4}(Sc_V)^{\theta_5}$, $\beta_L = 1.19\sqrt{\frac{D_L}{d_h}}\left(\frac{U_L^{1.2} g^{1.3}\sigma_L^{0.3}\varepsilon^{1.2}(1-0.93\cos\gamma)^2}{\nu_L^{1.3}\rho_L^{0.3}a_P^{2.4}}\right)^{\theta_6}$
Rocha, SRP II	$\frac{a_e}{a_p} = F_{SE}\frac{29.12(We_L Fr_L)^{\theta_1} S^{0.359}}{Re_L^{\theta_2}\varepsilon^{0.6}(1-0.93\cos\gamma)(\sin\theta)^{0.3}}$ $\beta_V = 0.054 D_V / d_h (Re_V)^{\theta_3}(Sc_V)^{\theta_4}$, $\beta_L = 2\sqrt{\frac{D_L}{\pi\cdot t_{res}}} = 2\left(\frac{C_E D_L}{\pi\cdot d_h}\frac{U_L}{\varepsilon h_L \sin\theta}\right)^{\theta_5}$
Billet & Schultes	$\frac{a_e}{a_p} = 1.5\frac{1}{\sqrt{a_p d_h}}\mathrm{Re}_L^{\theta_1} We_L^{\theta_2} Fr_L^{\theta_3}$ $\beta_V = C_V\sqrt{\frac{a_p}{(\varepsilon - h_L)d_h}} D_V (Re_V)^{\theta_4}(Sc_V)^{\theta_5}$, $\beta_L = 1.514 C_L\left(\frac{U_L}{h_L}\right)^{\theta_6}\left(\frac{D_L}{d_h}\right)^{\theta_7}$

These correlations, however, were originally developed for vapor-liquid (two-phase) absorption or distillation. No mass transfer correlation for three-phase packed distillation can be found in literature. In this paper, three most used correlations, (i.e. Rocha correlation, Shi & Mersmann correlation, Billet & Schultes correlation), are adjusted for three-phase mass transfer calculation. A set of parameters used in the correlations is re-estimated by a hierarchical estimation methodology based on our comprehensive experimental researches of two-phase and three-phase distillations.

2. Parameter Estimation Problem

2.1. Experimental Design

To estimate the transfer parameters for three-phase distillation, a laboratory-scale packed column (Fig. 1, left) was built up in the Institute of Process Engineering (TU Berlin) to produce representative experimental database. The column has an inner diameter of 0.1m and whole height of 6.2m. The column was operated under total reflux and can achieve a maximum reboiler heat duty of 30KW (F-factor up to 2.8 $Pa^{0.5}$).

In this study, the experimental database contains 162 different experiments of n-Butanol/Water/n-Propanol system for three different packings (Montzpack B1-350, Raschig Super-Ring 0.3, Rombopak 9M). For the same experimental setup, both the two-phase and three-phase distillations experiments were carried out. Four vapour and four liquid samples are taken from suitable locations of the column.

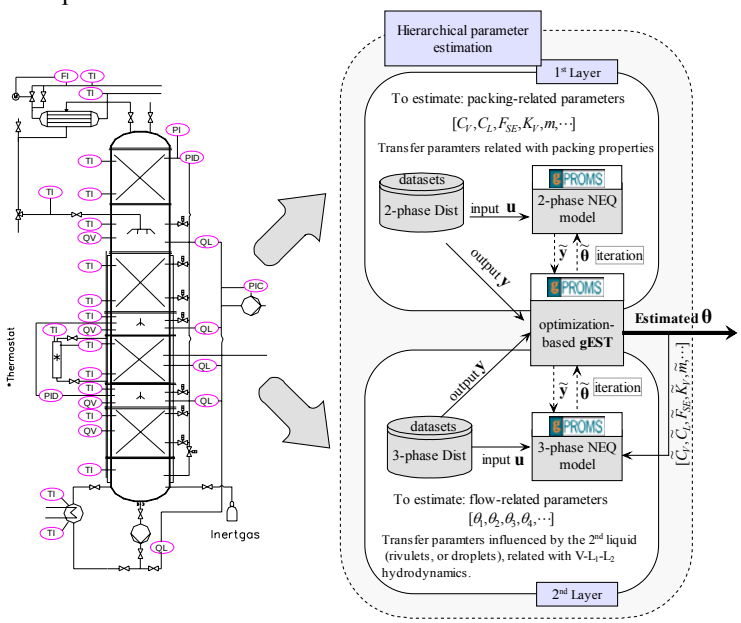


Fig. 1. Experiment details and estimation strategy

2.2. Hierarchical Methodology

Considering three-phase distillation, there are two kinds of transfer parameters should be determined based on the above experimental design.

1. "Packing-related" parameters (e.g. parameters such as C_V, C_L, F_{SE}, K_V, m in the correlations in table 1) are used to express the influence of geometric packing properties on mass transfer. For certain packings, the values of packing-related parameters are available; for other packings, however, these values are still not available and have to be experimental determined.
2. "Flow-related" parameters (as exponential $[\theta_1, \theta_2, \theta_3, \cdots]$ in table 1) reflect directly the influence of flow-behavior and hydrodynamics. In three-phase distillation, the 2nd liquid appears and could change the hydrodynamics in a dramatic way. Hence, these parameters have to be re-estimated using three-phase distillation database.

To estimate both kinds of transfer parameters, a hierarchical (two-layer) parameter estimation methodology (Fig. 1, right) is proposed in this paper. The packing-related parameters only reflect the influence of packing properties on mass transfer, and should be identical for two- and three-phase distillation. Hence, two-phase experimental data are utilized to estimate these parameters in the 1st layer. The estimated parameters are then passed into 2nd layer as given values to estimate the exponential flow-related parameters $[\theta_1, \theta_2, \theta_3, \cdots]$. These parameters account for the three-phase hydrodynamics. They are estimated based on experimental data sets of three-phase distillations.

The parameter estimations are carried out in gPROMS®. NEQ models of two-phase distillation and three-phase distillation are developed for simulation. The mass transfer correlations in table 1 are used for rate-based mass transfer calculation. The gEST®

entity is used to integrate simulation model and experimental database, and to minimize the deviations between them. Detailed mathematical formulation of the principle model-based parameter estimation is discussed in next section.

2.3. Mathematical Formulation

The main idea of the parameter estimation is to formulate a multi-data-set optimization problem subjected to the NEQ model equation system. A general parameter estimation problem with multiple data sets is formulated as in Eq. (4). To estimate the parameters $\boldsymbol{\theta}$, there are normally several experimental data sets of output variables $\mathbf{y}$ ($\tilde{y}$, simulation predictions), and the input (manipulated) variables $\mathbf{u}$. There are also many unmeasurable state variables $\mathbf{x}$ in the model equations $\mathbf{g}$ and inequality constrains $\mathbf{h}$. In this study, we assume the output variables $\mathbf{y}$ are subjected to measurement error, while the manipulated variables $\mathbf{u}$ are free from error (i.e., only the covariance matrix $\boldsymbol{\omega}$ of output variables was considered in the maximum likelihood objective function $\boldsymbol{\Phi}$).

$$\begin{aligned} &\boldsymbol{\Phi}(\boldsymbol{\theta}) = \frac{N}{2}\ln(2\pi) + \frac{1}{2}\min_{\boldsymbol{\theta}} \sum_{i=1}^{NDS}[\sum_{j=1}^{NV}\ln(\omega_{i,j}^2) + (\mathbf{y_i} - \tilde{\mathbf{y}}_\mathbf{i})^T \boldsymbol{\omega}_\mathbf{i}^{-1}(\mathbf{y_i} - \tilde{\mathbf{y}}_\mathbf{i})] \\ &s.t. \\ &\quad \mathbf{g_i}(\mathbf{x_i}, \mathbf{y_i}, \mathbf{u_i}, \boldsymbol{\theta}) = \mathbf{0} \\ &\quad \mathbf{h_i}(\mathbf{x_i}, \mathbf{y_i}, \mathbf{u_i}, \boldsymbol{\theta}) \geq \mathbf{0} \quad i = 1, \cdots, NDS \\ &\quad \boldsymbol{\theta}^L \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}^U \end{aligned} \tag{4}$$

Considering the proposed hierarchical parameter estimation, here

$$\boldsymbol{\theta} = \begin{cases} \boldsymbol{\theta}_1 = [C_V, C_L]^T, or\ F_{SE}, or\ [K_V, m]^T & 1^{st}\ \text{layer estimation} \\ \boldsymbol{\theta}_2 = [\theta_1, \theta_2, \theta_3, \theta_4, \cdots]^T & 2^{nd}\ \text{layer estimation} \end{cases} \tag{5}$$

The output variables are

$$\mathbf{y_j} = [y_{in,1,j}, y_{in,2,j}, y_{in,3,j}, y_{out,1,j}, y_{out,2,j}, y_{out,3,j}]^T \tag{6}$$

where $\mathbf{y_{in}}$ and $\mathbf{y_{out}}$ are the vapor concentrations (experimental measurements) for each packing section and each component, respectively.

The parameter estimation turns out to be a large-scale problem (Chen, Repke, Wozny, 2008). Using the hierarchical estimation methodology the problem is divided into four sub-problems with smaller problem dimension, and the coupling between packing-related and flow-related parameters is removed. This is the benefit of the strategy.

3. Results and Discussion

3.1. Case Study: Extending Rocha Correlation

As an illustrative example, Rocha correlation (SRP II model) is adjusted using the hierarchical parameter estimation methodology.

1st Layer Estimation. In the SRP II model, all packing-related contributions to mass transfer are attributed in parameter F_{SE}. A number of 30 data sets (two-phase distillation, F-factor, 0.2-2.6 $Pa^{0.5}$) are adopted to estimate F_{SE} for three packings (10 experiments for each packing). The estimation problem contains about 3,000 variables (equations). The estimated values are listed in table 2.

Table 2. Estimation results of packing-related parameters

Packing	F_{SE}	95% t-value	Standard deviation	HETP (exp)
Montzpack B1-350	0.313	8.5	0.018	0.326m
Raschig super-ring 0.3	0.484	7.8	0.031	0.288m
Rombopak 9M	0.489	15.4	0.016	0.286m

The estimated results show a rather high 95% t-value compared with the reference 95% t-value 1.68, and rather small standard deviation, which suggest reasonable acceptance of the estimation. The results indicate that Super-ring 0.3 and Rombopak 9M have larger effective interfacial areas than Montzpack B1-350. This conclusion agrees well

with the experiment measurement of HETP values (table 2), which also implies better separation efficiency by using Super-ring and Rombopak.

2nd Layer Estimation. There are five exponential flow-related parameters, i.e. $[\theta_1, \theta_2, \theta_3, \theta_4, \theta_5]$. 60 data sets of three-phase distillation are used to estimate the flow-related parameters. It's a large-scale parameter estimation problem include around 44,000 variables (equations). The CPU time needed is about 1500s. The results are listed in table 3 for different F-factor operations. Enhancement factors f_E are defined (Fig. 2) to visualize the enhancement ($\pm$) achieved by new parameters.

Table 3. Estimation results of flow-related parameters

	a_e		β_V		β_L
	θ_1	θ_2	θ_3	θ_4	θ_5
Three-phase, L①	0.150	0.210	0.743	0.040	0.489
Three-phase, H①	0.179	0.216	0.810	0.276	0.457
Three-phase, A①	0.163	0.197	0.769	0.241	0.468

①: L, low F-factor (1.0-2.2Pa$^{0.5}$); H, high F-factor (2.2-2.8Pa$^{0.5}$); A, all F-factor (1.0-2.8Pa$^{0.5}$)

Mass Transfer Coefficients. For vapor side mass transfer coefficients, they have +10% enhancement under high F-factor conditions and –30% under low F-factor conditions (see Fig. 2). For liquid side mass transfer coefficients, positive enhancement is observed for all experimental cases.

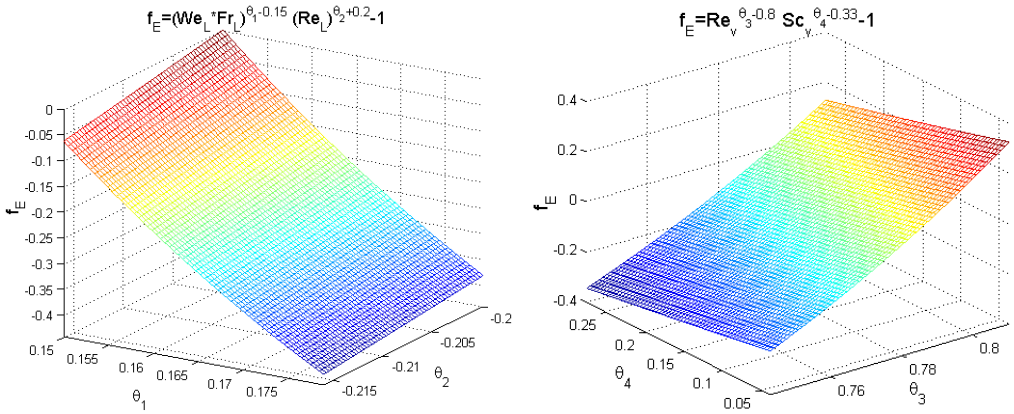


Fig. 2. Enhancement of transfer parameters (left: a_e, right: β_V)

Effective Interfacial Area. Compared with two-phase distillation, an obvious reduction of a_e is found. However, under low F-factor, f_E is only –5%, indicating Rocha correlation can be used for low-factor three-phase distillation; under high F-factor, f_E is about 30%-40%, indicating the unreliability of Rocha correlation. In these cases, the newly estimated flow-related parameters should be preferred.

3.2. Summary and Remarks

Rocha correlation has been adjusted for simulation of three-phase distillation. The updated parameters are summarized in table 4. To verify the benefit of using the adjusted correlation, a comparison of simulation (sim) results with another 48 experimental (exp) data is shown in Fig. 4. Obviously, the results using the adjusted correlation show much better prediction for light component n-Propanol and Butanol.

Table 4. An extension of Rocha correlation

	θ_1	θ_2	θ_3	θ_4	θ_5	$[F_{SE}]$		
2-phase	0.15	0.20	0.80	0.33	0.50	Super-ring 0.3	Montzpack B1-350	Rombopak 9M
3-phase	0.163	0.197	0.769	0.241	0.468	0.484	0.313	0.489

In a similar way, Shi & Mersmann correlation has been successfully adjusted. The results are listed in table 5. The adjusted parameters for Billet & Schultes correlation are summarized in table 6.

Table 5. An extension of Shi & Mersmann correlation

	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	$[K_V, m]$		
2-phase	0.40	0.20	0.15	0.67	0.33	0.17	Super-ring 0.3	Montzpack B1-350	Rombopak 9M
3-phase	0.417	0.199	0.149	0.651	0.111	0.165	[0.441,159]	[0.422,145]	[0.473,166]

Table 6. An extension of Billet & Schultes correlation

	θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7	$[C_V, C_L]$		
2-phase	-0.2	0.75	-0.45	0.75	0.33	0.5	0.5	Super-ring 0.3	Montzpack B1-350	Rombopak 9M
3-phase	-0.167	0.799	-0.457	0.758	0.333	0.485	0.460	[0.595,1.603]	[0.458,1.246]	[0.542,1.482]

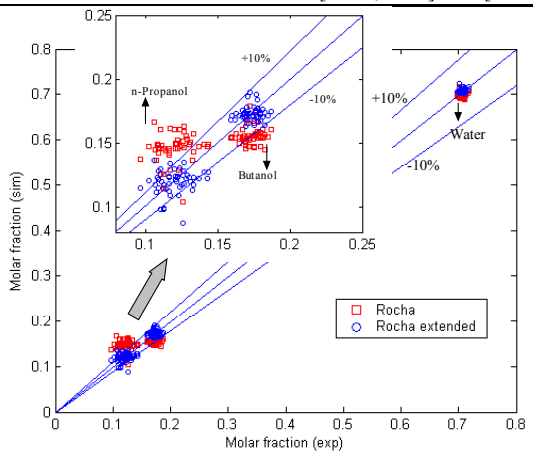


Fig. 4. Comparison of simulated and experimental results for vapor molar fraction

4. Conclusion

A hierarchical methodology is proposed in this paper to extend the transfer correlations for three-phase mass transfer calculation. A set of parameters used in the correlation is re-estimated based on our comprehensive experimental research on both two-phase and three-phase distillations in packed column. Three most widely used correlations, (i.e. Rocha correlation, Shi & Mersmann correlation, Billet & Schultes correlation), are adjusted. Substantial improvements on the NEQ model prediction have been observed.

References

J.-U. Repke, G Wozny, 2002, Experimental investigations of three-phase distillation in packed column, *Chem. Eng. & Tech*., 25, 513-519

R. Taylor, R. Krishna, 1993, Multicomponent Mass Transfer, John Wiley & Sons, New York.

M.G. Shi, A. Mersmann, 1985, Effective interficail area in packed column, Ger. Chem. Eng., 8, 87-96

J. L. Bravo, J. A. Rocha, J. R. Fair, 1985, Mass Transfer in Gauze Packings, *Hydrocarbon Processing,* 64, 1, 91-95

J. A. Rocha, J. L. Bravo, J. R. Fair, 1996, Distillation Columns Containing Structured Packings: A Comprehensive Model For Their Performance. 2. Mass-Transfer Model, *Ind. Eng. Chem. Res.*, 35, 5, 1660-1667

R. Billet, R. Schultes, 1999, Prediction of Mass Transfer Columns with Dumped and Arranged Packing – Updated Summary of The Calculation Method of Billet & Schultes, *Trans IChemE, Part A, Chem. Eng. Res. Des.*, 77, A6, 498-504.

O. Villain, J.-U. Repke, G. Wozny, Performance characterization of three-phase operated column, AIChE annual Meeting, 2003

L. Chen, J.-U. Repke, G. Wozny, 2008, nonequilibrium model based large-scale parameter estimation for three-phase distillaion in packed column, Chem. & Proc. Eng., submitted

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A mathematical model for water removal in the press section of a paper manufacture industry

Daniela M.Devienne Drummond[a], Maria Teresa Moreira Rodrigues[b], Ignacio Grossmann[c], Reginald Guirardello[d]

[a,d]*UNICAMP/FEQ/DPQ,Campinas,Brazil*
[b]*UNICAMP/FEQ/DTF,Campinas,Brazil*
[c]*Carnegie Mellon University,Pittsburgh,USA*

Abstract

A production optimization problem concerned with the water removal in the press section in a paper machine is considered in the present paper. The proposed model seeks to determine the planning of production of paper in order to minimize a cost function that consists of replacement of the felts in the press section, cost of energy to operate the press and cost of energy in the drying section. The proposed model corresponds a mixed-integer nonlinear programming (MINLP) where the most important decisions in the paper machine are: a) the sequence of paper to produce, or when to produce the paper, b) the need to exchange the felts, and c) when to exchange the felts. Numerical examples are presented to illustrate the performance of the model. This work was developed considering a real case with data from a Brazilian paper plant.

Keywords: Optimization, Schedule, GAMS, MINLP

1. Introduction

With increased interest in optimizing parts of paper machines in recent years, there has been much work addressing to solve those problems. Most of the works that has been reported have focused in optimizing different parts in a paper producing mill, for example in paper-converting Westerlund et al (1980). Also, several authors have studied water removal in the press section, and we can find references using as a basis pressing defined for one nip of vertical flow Wahlstrom (1960). Different models and mathematical models for that part of paper machine are reported in Kerekes et al (1991). The process for the manufacture of paper is common to any plant, initially preparing the mass of the paper, following the formation sector, going to the press section for water removal the web of paper, and finally going to the drying sector for drying the web with hot air. In a paper machine, specifically in the press section, we have two important parts, the felts whose function is to carry the wet web through the press nip Farouk (1991) and the nip which is the zone of contact between two rolls, where the water is transferred from the web to the felt. In the press section, which is the object of study of this work, the average life of a felt is different, depending on the type of the felt, but in general it is around 35-45 days. We consider in this work the press section, which is an important part of the machine, affecting the properties of the paper, as well as having an impact on the final cost of manufacture. Low efficiency of this section causes difficulties like reduction of the tensile strength, increase in the steam consumption in the drying section of the machine, and in many cases, the reduction of the productivity due to reduction of the speed of the machine. A reduction of 1% of humidity in the leaf results in a reduction of steam consumption in the order of 4.5%. This work described in

this paper has the novelty of minimizing the replacement of the felts, by using an optimal sequence of paper production, in order to improve the water removal in press section. The model for optimizing the press section has as an objective to obtain better sequence of production of the reels, aiming at the increasing the water removal for the press section.

2. Press section

The sector of press in a paper machine is responsible for the removal of approximately 18 20% of water of the sheet of paper, this represents about 12% of the total cost of water removal in the machine. The biggest amount of water that can be removed of the paper, before this enters in the drying sector represents a great economy, since the sector of responsible drying and for 78% of the cost of water removal in the sheet of paper. Its main function, then, is to remove the maximum amount of water contained in the paper for half mechanics. The humid prensagem is based on the compression mechanics of the leaf and paper, in contact with felt between two coils, combined with the suction of the water for the inferior coil. Bigger removal determines efficiency of the section better, leading in account the economic aspect of the operation and the final quality of the production. The way with that the felt and the leaf enter in the NIP (contact zone enters the coils) of prensagem, both load obtain a restrained water film for the superficial tension of its respective structures fibrosa. The water extration in the press is folloied by a compression of the leaf, in the ticket between the two coils. The prensagem directly is related with, the applied pressures, the width or area of the NIP which is submitted to the compression, the time of permanence under pressure of the leaf on the felts and the capacity of retention and water conduction for the felts.

3. Model

First of all we divide the programming for "reels in the end of the process", independently of the time of processing of each one. Since in this work we do not assume unexpected changes or a plant recess for unexpected problems in the felts, the total time T is exactly independently of the processing order. However, the final humidity of each processed reel depends on the one that was dried before, which influences the felt used for drying. We define the binary variable for production of the reels as follows:

$$x_{ik} = \begin{cases} 1 & \text{for roll } i \text{ process in interval } k \\ 0 & \text{otherwise} \end{cases} \tag{01}$$

The following constraints them hold: each reel (or group of reels) can only be processed in one time interval; each time interval can only process one reel at a time (or group of reels). These assignment constraints are, given by:

$$\sum_{k=1}^{N} x_{ik} = 1 \qquad i = 1,\ldots,N \tag{02}$$

$$\sum_{i=1}^{N} x_{ik} = 1 \qquad k = 1,\ldots,N \tag{03}$$

Considering that at the beginning of the production all the felts are new, we define the binary variable y_{jk} for representing potential replacement of a felt in position *j* at the beginning at interval *k*. Each position is one of the press machines (there are four of them). We then define the binary variable for exchanges of felts as follows:

$$y_{jk} = \begin{cases} 1 & \text{if the felt in position } j \text{ was changed at the begining of interval } k \\ 0 & \text{otherwise} \end{cases}$$

Since the production starts with new felts (at the beginning of interval $k = 1$):

$$y_{j1} = 1 \qquad j = 1,2,3,4 \tag{04}$$

In this model, the objective function includes the cost of new felts, the cost of energy in the press section, and the cost of energy in the drying section. Since we are using a fixed total time horizon, the objective function uses the mass of water removed from the reels in the drying section at each interval k, m_k, and not the mass flow rate, $\dot{m}_k$.

The objective function is then given by:

$$\min z = \sum_{k=1}^{N}\sum_{j=1}^{4} CF_j \cdot y_{jk} + \sum_{k=1}^{N}\sum_{i=1}^{N} CP_{ik} \cdot x_{ik} + \sum_{k=1}^{N} cs_k \cdot m_k \tag{05}$$

The first two terms in the right hand side of Equation (05) are linear. The most complicated term is the last one. The model was then rearranged as follows:

$$\min z = \sum_{k=1}^{N}\sum_{j=1}^{4} CF_j \cdot y_{jk} + \sum_{k=1}^{N}\sum_{i=1}^{N} CP_{i,k} \cdot x_{ik} + \sum_{k=1}^{N}\sum_{i=1}^{N} CS_{ik} \cdot ma_{ik} - \sum_{k=1}^{N}\sum_{i=1}^{N} CS_{ik} \cdot U_i^s \cdot A_i \cdot x \tag{06}$$

Subject to restrictions (02)-(03) and to:

$$t_k = \sum_{i=1}^{N} T_i \cdot x_{ik} \tag{07}$$

$$\bar{\tau}_{jk} = \theta_{jk} + 0.5 \cdot t_k \tag{08}$$

$$\tau_{jk} = \theta_{jk} + t_k \qquad k = 1,\ldots,N \tag{09}$$

$$\theta_{jk} = \tau_{j,k-1}^{(2)} \tag{10}$$

$$\tau_{j,k-1} = \tau_{j,k-1}^{(1)} + \tau_{j,k-1}^{(2)} \tag{11}$$

$$0 \le \tau_{j,k-1}^{(1)} \le M_1 \cdot y_{jk} \tag{12}$$

$$0 \le \tau_{j,k-1}^{(2)} \le M_1 \cdot \left(1 - y_{jk}\right) \tag{13}$$

$$v_k = \sum_{i=1}^{N} \left(\ln\left(U_i^e \cdot A_i\right)\right) \cdot x_{ik} - \sum_{j=1}^{4} \ln\left(1 + \exp\left(u_{jk}\right)\right) \tag{14}$$

$$u_{jk} = \sum_{i=1}^{N} \left(\ln \lambda_{ij}\right) \cdot x_{ik} - \alpha_j \cdot \bar{\tau}_{jk} + \ln \xi_{0,j} - \ln\left(1 - \left(1 - \xi_{0,j}\right) \cdot \exp\left(-\alpha_j \cdot \bar{\tau}_{jk}\right)\right) \quad (15)$$

$$ma_{ik} = m_{ik}^{(1)} \quad (16)$$

$$m_{ik}^{(1)} + m_{ik}^{(2)} = \exp\left(v_k\right) \quad (17)$$

$$0 \le m_{ik}^{(1)} \le M_2 \cdot x_{ik} \quad (18)$$

$$0 \le m_{ik}^{(2)} \le M_2 \cdot \left(1 - x_{ik}\right) \quad (19)$$

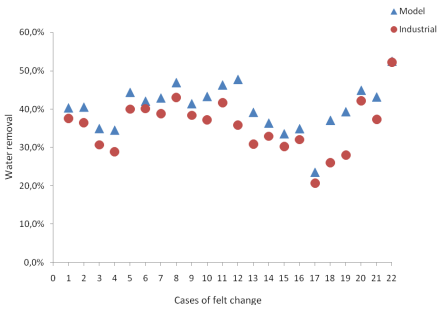
$$v_k \ge \sum_{i=1}^{N} \left(\ln\left(U_i^s \cdot A_i\right)\right) \cdot x_{ik} \quad (20)$$

The MINLP model is then formulated as the minimization of the objective function given by equation (06), subject to the constraints given by equations (02) - (03), and (07)-(20).

4. Results

4.1. Water removal

The challenge faced by the proposed mathematical model was the reduction in the replacement of felts through the selection at an optimal sequence of the reel production, in this way increasing the water removal in the press section and thereby reducing the cost of energy in the drying sector. This represents an annual increase of the order of 5% in the water removal, which is a significant result.



4.2. Cost reduction

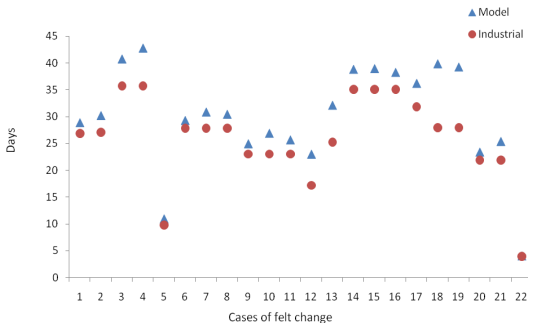
From the results we can also see a clear total cost reduction in the operation of the plant, the total cost was reduced by 4%. Some costs have not changed, like the costs in the press reduction, but the cost reduction in the drying section and the lower cost in the replacement of felts were the main reason for the total cost reduction.

Table 1: Costs for the optimized problem and the industrial case study

Cost	Model	Industrial
Felts	US$ 5,178,960.03	US$ 5,833,774.48
Press	US$ 16,420.31	US$ 16,420.31
Drying	US$ 23,538,536.13	US$ 24,082,059.66
Total	US$ 28,733,917.03	US$ 29,932,254.45

4.3. Felt lifetime and replacement

The lifetime of the felts can also have a significant impact on the cost. The main issue in the replacement of the felts is to find a point where they can stay in the machine without losing much their capacity of absorption. Through the proposed model, we obtained an average increase of 15% in the useful life of the felt, using the optimal sequence predicted from the mathematical model. Hence, besides the increase of the water removal, we can have an increase of the useful life of the felt. And about felts replacement, we obtained a reduction of 46% in average in the reduction of exchange of the felts, clearly due to the increase of lifetime of the felts, because we have fewer replacements of felts in the machine.



6. Conclusions

The optimization to reduce costs and replacement of the felts were considered in the present paper. An MINLP model has been proposed to determine the replacement of felts in the press section in a paper machine. It was shown that the problem could be efficiently solved yielding better results. The improvement obtained for a case study with 22 cases of the felt resulted in a 5% increase in water removal, 15% increase in felt lifetime and 4% reduction in the total cost.

Nomenclature

Sets

i reel

j felt

k interval

Variables

m_k amount of water in the entrance of the press section

ma_{ik} mass of water in the exit of pressing of the reel i in interval k

t_k time of interval k

u_{jk} auxiliary variable for the extraction coefficient of water in pressing

v_k auxiliary variable for water in the exit of the pressing in interval k

x_{ik} reel i processed in interval k

y_{jk} felt in press j exchange in initial time k

τ_{jk} time of life of the felt *j* in interval k

$\overline{\tau}_{jk}$ average of the time of the felt *j* in interval k

θ_{jk} initial time of the felt *j* in interval k

Parameters

A_i total area of the processed reels

CS_{ik} steam cost in the drying section ($/mass of water) in interval k

CF_j cost of buy/change ($/felt)

CP_{ik} energy cost expense in the press section of i ($/reel) in the interval k

M_1 big number (time)

M_2 big number (mass of water)

T_i time to process reel i

U_i^e humidity of the paper in the entrance of the press section (mass of water/area)

U_i^s humidity of the paper in the exit of the drying sector (mass of water/area)

α_j coefficient of reduction of the useful life of the felt

λ_{ij} extraction factor of the new felt j with reel i paper

$\xi_{0,j}$ volumetric fraction of fiber in the new felt j

References

Westerlund,T, Harjunkoski.I, Isaksson.J., Solving a production optimization problem in a paper-converting mill with MILP, Computer Chem.Eng (22):563-570 (1980).
Wahlstrom,P.B. Pulp Paper Mag.Can. 61(8):T379, (1960) .
Kerekes RJ, McDonald JD, A decreasing permeability model of wet pressing: theory. Tappi Joumal.74(12):150-156 (1991).
Farouk El-Hossehy, An equation for press felt compression, Tappi Journal p193 – 196(1991).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Application of Water Source Diagram for Multiple Sources and Multiple Contaminants Processes

Rafael B. S. Carvalho[2], Alexandre F. Cordeiro[2], Eduardo M. Queiroz[1], Fernando L. P. Pessoa[1]
[1]*DEQ/EQ/UFRJ – Sala 209, Bloco E, Escola de Química, Centro de Tecnologia, Cidade Universitária, Ilha do Fundão, Universidade Federal do Rio de Janeiro, Brazil 21949-900*
Email address: pessoa@eq.ufrj.br
[2]*CHEMTECH – Rua da Quitanda, 50, Centro, 20011-030, Rio de Janeiro, RJ, Brazil*

Abstract

The minimization of the use of water source and, therefore, the reduction of effluent generation, has been an increasing concern in the chemical industry. Because of that, heuristical algorithms have been developed trying to supply systemic approach to the mass transfer network optimization`s problem. The main objective of these heuristical algorithms is to get a mass transfer network [1], where contaminants of rich streams are transferred to other poor streams, in an optimized structure of linked mass transfer operations. This optimization is measured in terms of pure or low concentration water source consumption, in relation to a structure in which only these sources are used. In this network, the streams have a maximum inlet and outlet concentrations to be respected, which characterizes mathematical restrictions in an optimization context. These restrictions make the best solution attainment difficult. A distinguished method to attainment mass transfer networks is the Water Source Diagram (WSD). This method uses hand calculations being, particularly, a useful tool for processes engineers. From WSD, it is possible to consider a variety of situations as reuse, multiple water sources, loss of water throughout the process, outflow restrictions, regeneration and recycle [2]. The WSD has resulted satisfactory solutions for the single contaminant case [2]. For the multi-contaminant one, a similar approach increased of the mass transfer ratio concept has been proposal. This approach was used in some simple cases [2], revealing it satisfactory. However, its application is still necessary in complex cases to prove the method`s robustness and efficiency. The aim of the present work is to apply the Water Sources Diagram methodology, as well as a proposal variation, in multiples contaminants and multiple water sources problems, to compare its results with more complex approaches presented in literature. As a base case it was used the solution proposed by Jin-Kuk Kim [3] to the problem proposed by Alva-Argaez[4]. The results obtained from WSD methodology was about 1% higher than base case in terms of mass flow used and about 5% in terms of costs. Considering the difference in complexity between the methodology used in the base case and the one presented in this study, we can concluded that WSD generate satisfactory results with less work, becoming very useful to problems that requires fast solutions and as a initial estimation in more complex ones.

Keywords: minimization, water source diagram, mass balance network.

1. Introduction

The sources of water usage minimization implies in a change of attitude when industries have to carry out how much effluent is treated. Typically, the "end of pipe" approach is used in new processes project. The concern is focused only in the treatment of generated flow lines, however the global concern with the conscientious use of the sources and the search for more economic processes has modified the profile performance of industries as for the use and treatment of the water used in processes.

1.1. Heuristical Algorithms

The main objective of the heuristical algorithms is getting mass transfer networks[1], where series of rich contaminants lines are transferred to another poor lines series, in an optimized structure of linked operations of mass transfer. This optimization is measured in terms of the source of pure water consumption, or low concentration water source, comparing with a structure in which only clean sources are used. In this network, the maximum concentrations of entrance and exit lines have to be respected. In the optimization method context, they figure mathematical restrictions. These restrictions make the process to find the optimal solution a harder work.

1.2. Water Source Diagram(WSD)

The Water Source Diagram is a simple methods that brings acceptables solutions and it can be applied where water is treated and supplied. This method uses simple calculations that can be carried through manually, being a particularly useful tool for engineers of processes. From the WSD, it is possible to consider a variety of situations as reuse, multiple water sources, loss of water throughout the process, restrictions of outflow, regeneration and recycle [2].

1.3. Multiple Contaminants Problem

The WSD has resulted satisfactory for the case of exists only one contaminant [2]. In case of multiple contaminants, there is a similar approach, however with additional systematic concept as "mass transfer ratio", has been proposal. This method was used in some simple cases [4], showing also satisfactory. However, it is still necessary apply it in more complex cases so that its robustness and efficiency can be proven.

In this context, Jin-Kuk Kim et al. [3] had studied a multiple contaminants system composed by ten mass transfer operations, four contaminants and two different sources of water. In this study, not only the mass integration and minimization of the water use had been studied, as well as the energy integration and the minimization of the use of hot and cold utilities. In this context, the model could be formulated through non- linear entire mixing programming, it has being used in such a way, an adequate but complex strategy of optimization. In this process, two basic cases had been studied: one considering the possibility of mixing lines at the end of the mass transfer operations (BASE1) and another, with direct thermal exchange between lines (BASE2), otherwise, the mixture of these lines for temperature adjustment is not allowed. These situations still had been compared with a base case where the *pinch* methodology for mass and energy were applied(BASE3).

2. Problem Description

The problem studied in this work contains ten operations of mass transf, four contaminants, respectively called contaminant A, B, C and D, and two water sources. The first source (Source A) contains null concentration of the four contaminants, the first source is constituted of pure water. The second source contains a few and equalized amount of all contaminants. In this, each contaminant is present in equal amount of ten

parts per million (Source B -10 ppm). There costs are respectively 0,5 and 0,1 $/ton. Jin-Kuk Kim [3] shows the the range of concentrations of each operation, hence called opportunity table.

3. Applied Solutions

They had been elaborated at all four distinct solutions for the problem; two using methodology WSD for current multiple contaminants (WSD -MC1 and WSD -MC2), one using analysis for inspection (IA) and the fourth implementing a modified WSD methodology (WSDR-MC). It is assumed that each mass transfer operation has a constant load removal of contaminants. The opportunity table represents the information of entrance for the models of synthesis of mass transfer networks. For the developed heuristical methods, a detailed analysis of the opportunity table becomes necessary. In this process, each choice can influence directly in the final result, where the flowcharts are generated from the interconnections of the operations. As premise is assumed, that there is no possibilities of incorporation of regenerators that could reduce the concentration of contaminants supplying new sources. The opportunity table can be rearranged of graphical to visualize all the intervals and opportunities at reusing lines as flow. A schematic graphic is built as a stair structure representing all the operations and components commanded in the possible intervals of concentration. An important propertie that can be observed is the monotonically, where the order is guaranteed according to the sequence of operations. The sequence follows the increasing of concentration of entrance and exit. In this way, a selection of the operations becomes necessary that can promote this monotonically.

The first rule for the selection of the operations involves the choice of a reference component. In this work, it was chosen as reference that one that after the ordinance and the exclusion of the operations that violate the monotonically, presents the biggest involved operation number. In case of same operation number, other criteria of sattle of matter had been considered, as outflow, levels of concentration etc. This increases the possibilities to occur a bigger integration between the involved operations. Table 3 illustrates, for each contaminant, the sequence operations capable to take care of the monotonically criteria.

Table 1. Sequence of the operations that guarantee of monotonically for component

COMPONENT	OP.SEQUENCE	NUMBER OF OPERATION.
A	10,8,3,2,4 10,9,3,2,4 10,9,7,5,6	5
B	10,7,8,4,5,6/2 10,9,3,4,5,6/2	6
C	10,8,4,2/5/9	4
D	**10,6,4,2,8,7,1**	**7**

It's necessary to choose one contaminant as reference, initially it is the D that presents 7 commanded operations. So far they had not been considered in the heuristic, factors as the outflow, that can make impracticable a flowchart on account of the low one reuses.

Selected the reference component, the second stage is the choice of the reference operatio. The reference operation must be that one that contains the lesser concentrations of entrance and exit, and that it has reuse potential in all the other

operations, in other words, it is the first operation of the described order previously. The reference operation is the code 10. Therefore, component D as reference and operations 10 and 6 will be developed two proposals of solution using.

Only evaluating the operations that will go to be integrated by the diagram. All the operations limits should be adjusted based on the reference operation. The mass transfer ratios are calculated considering the ratio between the differences in each operation and in the reference operation, for each component. But, adjusting the inlet concentration, the denominator is given by the differential of reference component's concentration, in the reference operation, whereas, in the outlet adjustment of concentration, the denominator is defined by the delta of reference component´s concentration in the operation that is being shift. In the calculation of displacement what he looks himself it is the concentration of the component of reference in the evaluated operation. For in such a way, use of the equation becomes:

Equation 1. Calculation of the concentrations shift of the reference component in the reference operation

$$C_{XR}^{Op} = \frac{(Cf_X^{Op} - Ci_X^{Op})}{\Delta ad} + Cad_{XR}^{Op}$$

Where:

Δad Delta (ton) to be defined. It is 1(one), in case of beeing calculated the inlet flowrate, and 2(two) when based on outlet concentration.

Cad Concentration (ppm), to be defined. It is the same as the used source in the reference operation, or the same as the inlet shift concentration, to get the shift outlet.

After the calculation of the shifts concentration for each operations that limit the possibility of reusing the reference operation., the procedure considered for Pessoa et al[5] for an only contaminant is executed. In this procedure, it is calculated in each interval, how much it must be transferred of water to the lines, having itself in sight the available sources, and prioritizing the reuse of the bigger concentration sources.

4. Results and Discussion

From the results obtained it was easy observed that appear situations in which is much more advantageous to use a clean source instead of reuse some dirty line, and cause a forward necessity of dilution, because of the components constraint. In this context, the methodology presented in the WSD-MC differs from the simple ones for the proposal of a new heuristic. This would come to answer the following question: "Which is it more advantageous to directly use water sources in the operation or a forward dilution for concentration adjustment ?". This advantage is translated by the amount of water to be used in the two situations. This rule comes to supply the lack of method WSD-MC evaluating the concentrations of all the components. Thus, the following definitions are necessary:

Operation source: Operation that, the principle, has available water for allocation;

White operation: Operation that, the principle, can be white of a water allocation;

Critical concentration Difference: It is the biggest difference of concentration, amongst all the contaminants (except the one of reference) enters the concentration of exit of the operation source and the concentration of entrance of the white operation;

Critical component: Component that supplies the concentrations of the difference of critical concentration.

Performing a simple mass balance, it is possible to determine which operation should be done, checking the lower final consumption. The entire permitted transfer match should be tried in pairs, but it is not necessary try all the matches, because the monotonically is guaranteed. The tables presented below (Table 2) shows the results for each applied solutions, considering respectively the sources flow rates and costs associates.

Table 2. Source flow rate consumption and respective costs

	FLOWRATE (ton/h)						
	BASE1	**BASE2**	**BASE3**	**WSD-MC1**	**WSD -MC2**	**AI**	**WSDR -MC**
SOURCE 0 ppm	432,46	432,46	432,46	478,30	432,10	431,20	464,36
SOURCE 10 ppm	178,86	179,05	179,54	138,81	8463,66	186,67	155,43
TOTAL	611,32	611,51	612,00	617,11	8895,76	617,87	619,79
COST ($/year)							
RELATIVE COST (%)	100,00%	100,01%	100,03%	453,80%	108,08%	100,06%	105,81%

It can be observed that the base cases continue to have less cost. A great discrepancy between solution WSD-MC1 and the cases is also observed. Moreover, the solution evaluated by inspection (AI) presented best results amongst the news evaluated. However, considering a systemic approach of reproducibility, the solution WSDR-MC appeared as the most interesting. Getting the best results in the family of WSD solutions.

Concerning the previously sections explains, to maintain the monotonically is an extremely important factor for the applied method WSD with multiple contaminants. Its application is associated with the use of the reference component concept, that turn its great responsible for the simplicity of the method. The proposal of the WSD -MC include the selection of an adjusted reference component, setting up the mass transfer correlations of other components, it is possible to find an optimal network with an analogous procedure to the WSD for simple contaminants. The unique constraint regarding the simple contaminant method is the guarantee a monotonic sequence of operation However, the results obtained from solutions 1 and 2 (WSD-MC1 and 2) evidence that considering only one reference compound, omitting the other concentrations constraints, takes the solution to a non-minimum water allocation. Analyzing just mass transfer correlations takes, in both solutions, the necessity of dilution in operational lines to close mass balance and follow the inlet and outlet concentrations constraint.The method does not ponder the better water relocation, to use a clean source or reuse some another outline. The method WSDR-MC brings up this concept and takes in consideration the advantageous possibility of directly water source allocation in the operations. Through the insertion of the concept of critical concentration difference it is possible to calculate the amount of used water source to dilute a possible reused line and compare it within the amount of source spends directly in the operation. This proposal heuristic complement the method WSD-MC prevents the necessity of water source usage for simple dilution at the end the built network, to close mass balance.

The solution obtained for inspection analysis, although it has the best result amongst the solutions carried through in this study, it is not acceptable as a reasonable process to find it out, because of it is not based in a systematic approach, and can`t be replied for other problems.

5. Conclusion

From the study presented it can be concluded that the use of the heuristic presented in solution WSDR-MC keeps the simplicity of the Water Sources Diagram method and presents a superior lightly result. In addition, it prevents the dilution after the set up of the network operation connections. Although these method`s results do not surpass the base cases, its simplicity and the small difference within the cited results, it guarantees a great applicability in optimization chemical processes. A detailed WSDR-MC method should be developed and it shows promissory..

References

[1] El-Halwagi MM, Manousiouthakis V. Synthesis of mass exchange networks. AIChE J 1989;35(8):1233e44.

[2] Juliana F. S. Gomes, Eduardo M. Queiroz, Fernando L. P. Pessoa, Design Procedure for Water/Wastewater Minimization: Single Contaminant. J. of Cleaner Production, 2007;15:474-485

[3] Boondarik Leewongtanawit, Jin-Kuk Kim, Synthesis and Optimization of Heat-Integrated Multiple-Contaminant Water Systems. Chem. Eng. And Proc. 2008;47:670-694

[4] A. Alva-Argaez, Integrated Design of Water Systems, 1999, Ph.D. Thesis, UMIST,Manchester, UK, 1999.

[5] Fernando L. P. Pessoa, Eduardo M. Queiroz, Notas de aula de Integração de Processos, Escola de Química, 2008.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Industrial Experience on the Implementation of Real Time On Line Energy Management Systems in Sugar and Alcohol Industry

Douglas C. Mariani [a], Marcos A. Kihn [b], Carlos A. Ruiz [b]

[a] *Soteica do Brasil, R. Dr. Cid de Castro Prado, 108, Planalto Paulista, CEP 04064-040, São Paulo – SP, Brazil*
[b] *Soteica LLC, 16010 Barke's Point Lane, Suíte 260, Zip Code 77079, Houston – TX, EUA*
E-mail address: douglas.mariani@soteica.com.br

Abstract

Sugar and alcohol market are increasing, due humanity needs to find alternatives for fossil fuels dependence. Brazil joins climactic and technologic features to place itself at the forefront of this industry. Sugar-alcohol industry is professionalizing and rising portfolio of product and the electric energy became an important product of these new ones. In this context, an increase in energy efficiency and actions to maximize this product shows it essential for these sites become competitive and reach their goals. Visual Mesa is a process on-line analysis tool to support decisions on energy reduction costs. This work describes how Soteica and Usaçúcar teams have modeled the site and implemented this new way to evaluate energy on this kind of industry. A full model of the energy system has been developed. All the constraints have been included and the model is continually being updated with live data. Performance monitoring was done and it includes the tracking of equipment efficiencies by utilizing updated data for its continuous calculation. By auditing the energy system, imbalances can be identified and reduced. Planning for a better operation of the energy system by performing case studies is usually done by using the validated model. As a result of the project, new sensors have been installed and a completely new way to evaluate energy on the site was implemented.

Keywords: sugar and alcohol industry, on-line energy analysis software, electric energy.

1. Introduction

In the last 30 years it was become clear that here are humanity concerns about the scarcity of fossil fuel supplements. It gets worse with the growth of the worldwide consumption and the geopolitics changes of the period.
This scenery intensified the humanity needs of a substitution of the energetic matrix, based on the oil, by a renewable, less pollutant and that cut down the carbon emission on the atmosphere.
Sugar and alcohol market are in constant expansion and it shows as an alternative to Brazil dependence of oil.
Brazil has the suitable climatic and technologic features to the development of an industry that produces a clean fuel, according to this new order and that stand on the leadership of the sector, as we can notice nowadays.

To maintain this leadership, sugar-alcohol industries are focusing in process optimization and costs reduction and one of the most important points of optimization is energy balance as described in the work of Sartori et al (2007), who developed a mathematical model to improve industries profits and better use of energy potential.
On this change of scenery, the sugar and alcohol industry has also been changing and improving its portfolio and adding new products, as a sale of electric energy to the energy supply network and carbon credits appear as a promising market in the next 20 years in Brazil as described by La Rovere et al (1994).
The alcohol and industry engagement on the electric energy production is very important to maintain the supply and increase the fraction of energy produced (in %) by sustainable process. Lora et al (2007) shows that it is an innovation in Brazilian market that was supported by the government and electricity companies and achieve d approximately 4,4 % of the total electric energy amount generated and this participation is growing.
The energy comes from the steam generated in the boilers that burn sugar cane bagasse from the milling process.
For the sugar and alcohol industry carries on producing energy and it is economically suitable, the energy balance must be monitored and improved.
To achieve this goal, Soteica used its on-line energy management system, Visual Mesa, which is the worldwide market leader on petrochemical sector to supervise and optimize the energetic balance in the sugar and alcohol industry.
The work of Ruiz et al (2005) describes how this tool is being used successfully around the world to optimize Energy Systems in petrochemical industries.
This is the first time that a tool for on-line management system is used to manage energy balance in real time on the sugar and alcohol market.
Soteica and Usaçúcar join their teams to develop a model for an industrial plant that produced sugar, alcohol and electric energy.
This work consists of experiences exchanged between the teams, one that knows about energy management and the other that knows about sugar-alcohol production technologies.

2. Project objectives

The main objective of the project was to have a tool available for the on line optimization, auditing and monitoring of the energy system of Tapejara Plant.
The secondary objective is to gain knowledge, develop a methodology and mapping the difficulties on the implementation of this kind of software in sugar-alcohol industries.

3. Usaçúcar Company

Usaçúcar is a company with about 50 years of experience in sugar and alcohol production and today at the forefront sector in the State of Paraná, where it is located in Brazil.
In Usaçúcar Group there are also trading companies and a distribution center at Paranaguá's Port, where the production is exported to many countries, mainly to Asia.
In the productive arm of company it has 7 productive units distributed in the North of Paraná State, in Brazil. Two of these productive units have alcohol as unique product and others units work with sugar too.
As the entire sugar-alcohol sector in Brazil, Usaçúcar is booming and improving its portfolio and one of those new products is the electric energy.

This improvement on the portfolio took optimization focus to the plants energy balance and how to achieve contracts and product specification, producing energy to export.
Nowadays Usaçúcar Group has two plants that are producing electric energy and selling to distribution companies. One of them was chosen to place the first implementation of a real time on line energy management system in sugar and alcohol industry.

4. Tapejara Plant

The plant chosen for the implementation was the Tapejara Plant. This plant has a capacity to process about 19,5 kt of sugar cane per day and produces 62.000 m^3 of alcohol and 316 kt of sugar each season. It is the largest productive plant of Usaçúcar Group.
Based on these numbers and in the expectation of higher profits, Tapejara Plant was chosen for the first implementation.
The layout can be organized in 7 sectors. A similar division was used in model building.

- **Steam generation:** This sector has two boilers that burn sugar cane bagasse and produce superheated steam with pressure of 65 bar and a temperature of 400 °C.
  Boiler 3 can produce 100 t/h of steam and Boiler 4 produces 300 t/h of steam. Boiler 4 is the biggest biomass boiler operating in the world.
  All steam produced here feed electric generator or is used in productive process.
- **Electric generation:** Electric generation sector has two generators that are fed with the superheated steam produced by the boilers, generates electric energy and reduces steam to 21 bar. The steam that was produced here will be used in the milling process.
- **Sugar cane preparation and milling:** This sector is a consumer of steam at 21 bar and it is used to operate machines that prepare and process the sugar cane to remove juice that will be used in other stages of production process.
- **Evaporators:** Tapejara Plant has two lines of evaporators, this equipment is used to concentrate sugar cane juice and specify the flow, raising the brix to about 65 °C.
  Evaporation is the plants energy core, because here is consumed steam at a high pressure and generates steam at a low pressure.
- **Crystallization:** Crystallization is a sector that consumes steam at a low pressure and the consumption depends on the amount of equipments that is on-line at each moment.
- **Distillation:** As with the crystallization, the distillation sector is a low pressure consumer and depends on the alcohol types that are produced.

5. Visual MESA

Visual MESA software was developed to work with global optimization of energy systems in an industrial plant and lead with its restrictions in order to bring the operation of system utilities for the lowest possible cost. It is a computer program designed to model steam, Boiler Feed Water, condensate, fuel, electrical systems and CO_2 emission costs.

It operates on-line with Plant Information System via standard OPC interface and it is also capable of dealing with historical data used in the optimization.
Visual MESA is installed for two types of use, the Stand Alone use and Client Server use:

- **Client Server use:** The purpose of this installations is to share the solutions, supporting multiple users. Visual MESA server runs as a service on a PC and gives solutions automatically with no interruption every 5 minutes, writing results on the plant information system and generating reports. With this architecture any PC connected to the plant network can be configured to access the model and the reports. So teams from operation to management, including engineering monitoring is able to evaluate and use the results supplied by the tool.
- **Stand Alone use:** The purpose of this installation is for individual users able to run cases studies on their own PC, using a snapshot of the current data or historical data. Data can be achieved automatically taken from the plant information system via standard OPC Historian Data Access. This kind of use also supports "What If" studies that are useful to assess revamps and changes in production schema.

In terms of optimization, Visual MESA has mathematical features and built-on optimization routines that make it possible to calculate how to the steam runs and electrical systems operate at a minimum overall cost while still meeting the required plant steam demands and other constraints.
The software determines which boilers or steam generators should make incremental steam and which turbines or letdown valves will most efficiently let the steam down between pressure levels. It uses a successive quadratic programming (SQP) optimizer, which is grouped into four levels:

- **Level 1:** includes the pressure control-related devices; boilers, letdown valves and vents are optimized to minimize cost.
- **Level 2:** adds the optimization of other continuous variables, including turbo generators and extraction/ induction/ condensing turbines.
- **Level 3:** adds turbine-motor switching optimization, like discrete variables.
- **Level 4:** optimizes variables that are potentially unsafe or affect process yield.

The program's objective function is able to optimize the total operating cost of the system, which is described in equation (1).

$$\text{Operating cost} = \text{total fuel cost} + \text{total electricity cost} + \text{inlet costs} \quad (1)$$

The SQP optimizer's job is to minimize this objective function, subject to operating constraints in the system. Total fuel cost is determined by the fuel use of each boiler and combustion turbine, multiplied by their respective fuel prices. Total electricity cost is determined by the electrical use of each motor, load and generator, multiplied by their respective electricity prices. The model takes into account the fact that electricity costs vary throughout the day.
Inlet costs normally include charges for demineralised water coming into the system, but can cover other arbitrary costs too.

6. Implementation Description

Soteica's teams in Brazil and Argentina in partnership have worked together with Usaçúcar – Tapejara Plant, Operations and Process teams, in the implementation project on the following activities:

- Layout analysis,
- Data collection,
- Linking of Visual Mesa to the Plant Information System,
- Build of Interface to the Historic Plant Information Data,
- Model building,
- Build on-line optimization,
- Model adherence analysis,
- Reports generation

The model built is the most important step of the implementation and is the basis for others steps.

It contemplates an overall and detailed description of the Energy System that includes bagasse stock, steam qualities, boilers, water, condensate, valves and electrical system.

Tapejara´s Plant was divided into 8 main areas on the Visual MESA model.

- **Steam generation:** that includes 2 boilers, one of them that produces 300 ton/h of superheated steam and the other produces 100 ton/h.
- **Reduction valves 67->21:** modeling of valves that reduce steam of 65 bar to 21 bar. That is a valve where condensate water is input in the steam and the pressure was reduced.
- **Reduction valves 21->1,5:** modeling of valves that reduce steam of 21 bar to 1,5 bar.
- **Electricity generation:** comprises two turbines to generate electricity to feed plants consume and exportation to energy supply network.
- **Sugar cane processing**: comprises turbines that move equipment of sugar cane preparation and milling. These equipments consume steam of 21 bar
- **Evaporators:** evaporation sector is composed by 2 evaporators lines, A and B. these lines consume steam of 1,5 bar and produce vegetable steam, that is the steam generated by sugar cane juice, of various levels of pressure below 1,5 bar that will feed distillation and crystallization sectors.
- **Crystallization:** modeled as a consumer of vegetable steam.
- **Distillation:** modeled as a consumer of vegetable steam as crystallization sector.

A main view of the models and the areas can be seen in Figure 1. By navigating through the model, each individual Unit of the system can be monitored in detail.

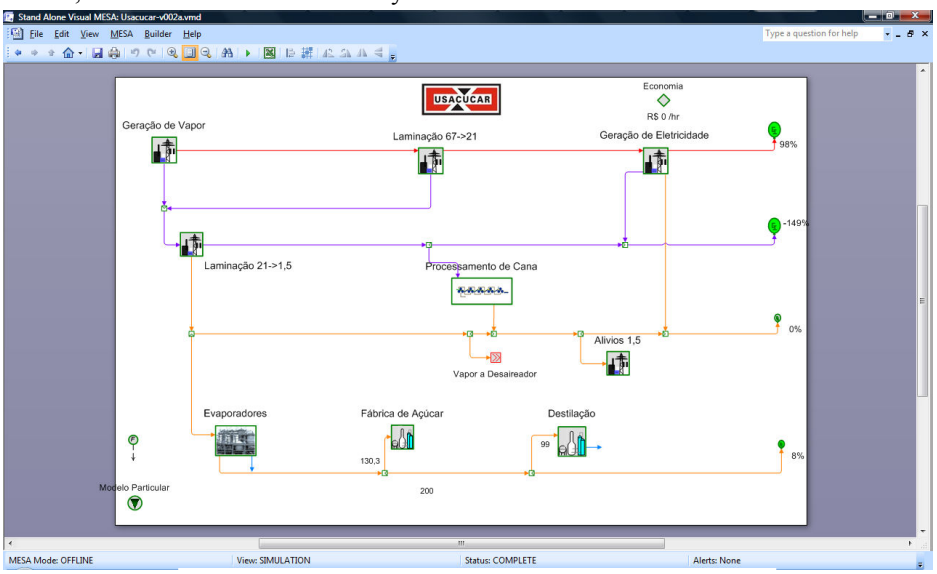


Figure 1: Main view of Tapejara Plant model

Electricity contracts and details of purchase and sale have been included in the model that made the model capable to evaluate energy trade.
Bagasse stock purchasing is modeled in order to manage the purchase of others Usaçúcar Group units and others companies too, focused on maximizing the production of electric energy.
During the implementation lacks of measurement was identified and programmed to be installed in the maintenance period.

7. Model adherence analysis

The model was run on-line under test conditions at the end of the 2008 season and obtained good results in adherence and ability to represent and optimize energy process.
First profits achieved in test phase are around 90,00 USD/h. Taking into account these profits and a season of 220 working days, it has estimated an annual profit of 475 kUSD.
The plant instrumentation will be improved on an annual maintenance and this work will give more data to the model, making the optimization more efficient and improving the reliability of the system.

8. Conclusions

The target was achieved, the model was built and the first implementation on sugar-alcohol industry was in course and presented initial results within the expected in terms of implementation time and costs.
With the implementation of new measurements, the model will be on-line in 2009 season and it will be the most important tool for monitoring and optimization of energy balance on Tapejara´s Plant.
This work will be used as a basis for new implementations of the software on others units in the Usaçúcar Group and others groups of the sugar-alcohol sector.

References

La Rovere, E. L., Legey, L. F. L, Miguez, J.D.G.. Alternative energy strategies for abatement fo carbon emissions in Brazil. Energy Police, v. 22, p 914-924. 1994.

Lora, E. S., Andrade, R. V.. Biomass as energy source in Brazil. Renewable & Sustainable Energy Reviews, 2008.

Ruiz, D., Manprim, J., Ruiz, C, Nelson, D., Rosene, G. Utilities Systems On-Line Optimization and Monitoring: Experiences from the Real World. European Symposium on Computer Aided Process Engineerin., 2005.

Sartori, M. M. P., Florentino, H. O.. Energy balance optimization of sugarcane crop residual biomass. Energy, v, 32, p 1745-1748. 2007.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Isothermal or Non Isothermal Modeling in a Pellet Adsorbent: Application to Adsorption Heat Pumps

Ludovic Montastruc[a,b], Pascal Floquet[a],Volker Mayer[c], Iordan Nikov[b], Serge Domenech[a]

[a] *Université de Toulouse, Laboratoire de Génie Chimique, U.M.R. 5503 BP 1301, 5 rue Paulin Talabot 31106 Toulouse Cedex 1, France*
b ProBioGEM Laboratory, Polytech'lille, Av. Langevin, 59665, Villeneuve d'Asc,q France
c Laboratoire Paul Painlevé, U.M.R. CNRS 8524, U.F.R. de Mathématiques 59 655 Villeneuve d'Ascq Cedex, France

Abstract

Understanding the interaction between a fluid and a solid phase is of fundamental importance to design of an adsorption process. Because the heat effects associated with adsorption are comparatively large, the assumption of isothermal behaviour is a valid approximation only when uptake rates are relatively slow. In this article, we propose to determine when it is needed to choose the isothermal or non isothermal assumptions according to two physical parameters α (ratio convection/capacity) and β (ratio quantity of energy/capacity). The proposed problem was solved by a mathematical method in the Laplace domain.

Keywords: Adsorption, Non Isothermal, Energy, Pellet.

1. Introduction

A basic adsorption heat pump cycle consists of four main parts: an adsorber, which is a container filled with an adsorbent (such as zeolite, active carbon, silica gel, etc.); a condenser; an evaporator; and an expansion valve. Basically, adsorption heat pump operates by cycling adsorbate between adsorber, condenser, and evaporator. In the adsorption heat pump cycle, adsorption phenomena play the same role of mechanical power, so that the working fluid can be circulated in the cycle without any mechanical power (Demir et al., 2008).

In recent years, the importance of adsorption heat pumps and adsorption refrigeration systems has increased since these kinds of systems can directly utilize the primary thermal energy sources and additionally the waste heat generated in various industrial processes. The important advantages of the adsorption heat pumps can be described as follows (Demir et al., 2008):

- they can operate with thermal driving energy sources such as waste heat, solar, and geothermal energies, etc.;
- they can work with low temperature driving energy sources;
- they do not require moving parts for circulation of working fluid;
- they have long life time;
- they operate without noise and vibration;
- they have simple principle of working;
- they do not require frequent maintenance;

- they are environmental friendly since they do not contain any hazardous materials for environment; and
- they can be employed as thermal energy storage device.

On the contrary, the main disadvantages of the adsorption heat pump systems are as follows (Demir et al., 2008):

- they have low COP values;
- they are intermittently working principles;
- they require high technology and special designs to maintain high vacuum; and
- they have large volume and weight relative to traditional mechanical heat pump systems

AHPs are thermally driven and can, therefore, make use of waste heat as well as thermal renewable energy resources. Moreover, AHPs can utilize high energy fuels more effectively than modern condensing boiler heating systems and, consequently, a reasonable amount of the resulting environmental pollution could be avoided. One of the most interesting working pairs for AHPs is zeolite-water. This working pair is non-poisonous, non-flammable and, moreover, has no ozone depletion potential. The adsorption of water vapour into a porous solid adsorbent, such as zeolite, results in the release of the heat of adsorption. Accordingly, the temperature of the adsorbent particle rises, which reduces its water uptake capacity. Therefore, the heat of adsorption has to be removed before a further vapour mass transfer into the adsorbent particle may take place. The evaluation of the heat and mass transfer rates accompanying such an adsorption process is therefore a very important aspect in designing and optimizing the operation of an adsorption heat pump.

In this study, a methodology to choose a good adsorbent in particular according to its thermodynamic properties is presented. We focused our attention on the interaction fluid/solid. The simplest case to consider is a single microporous adsorbent particle exposed to a step change in sorbate concentration at the external surface of the particle at time zero. Heat transfer is assumed to be sufficiently rapid, relative to the sorption rate, so that temperature gradients both through the particle and between particle and surrounding fluid are negligible. Two cases were developed in this article, firstly the isothermal conditions and secondly adsorption behaviour with temperature effect. A mathematical development is presented for each case. According to two developed parameters α (ratio convection/capacity) and β (ratio quantity of energy/capacity), the difference between isothermal and non-isothermal was analysed.

2. Isothermal Assumptions

During adsorption, heat generated is partly transferred to the surrounding fluid and partly accumulated in the particle leading to an increase of the temperature. The calculation of the temperature uptake curves is based on mathematical model of different complexity depending on the simplifying assumptions being made. Since adsorption is exothermic and the heat of sorption must be dissipated by the heat transfer, there is, in general, a difference in temperature between an adsorbent particle and the ambient fluid when sorption is taking place. Whether or not this temperature difference is significant depends on the relative rates of mass and heat transfer. By the simple theoretical analysis it may be shown that in a batch adsorption experiment it is the

dissipation of heat from the external surface of the adsorbent sample, rather than the conduction of heat within the adsorbent, which is generally the rate limiting heat transfer process. The conditions under which heat transfer resistance may be neglected and the system treated as isothermal case.
Firstly, we considered some isothermal conditions in the pellet and the isotherm was $q = q_0$. The isotherm is supposed to relate the concentration q inside the pellet (or at the external boundary) to the concentration of the external fluid. Another basic assumption is that the pellet is treated as a homogenous medium (contact with the fluid occurs only at the external boundary). So, the diffusivity Dc is constant in the pellet.

The system was modeled by the following equations (1):

$$\begin{cases} \dfrac{\partial q}{\partial t} = D_c\left(\dfrac{\partial^2 q}{\partial r^2} + \dfrac{2}{r}\dfrac{\partial q}{\partial r}\right) & \forall t > 0, \forall r\ (0^+ < r < r_c) \qquad (1-a) \\ q(r, t=0) = q'_0 & \forall r\ (0^+ < r < r_c) \qquad (1-b) \\ q(r_c, t) = q_0 & \forall t > 0 \qquad (1-c) \\ \left.\dfrac{\partial q}{\partial r}\right|_{r=0^+} = 0 & \forall t > 0 \qquad (1-d) \end{cases}$$

The value $q(r_c,t)$ is equal to $q_0 \neq 0$, for solving the system analytically.
With $\tau = D_c t/r_c$, the analytic solution is given by the familiar expression

$$\frac{m_t}{m_\infty} = 1 - \frac{6}{\pi^2}\sum_{n=1}^{\infty}\frac{e^{-n^2\pi^2\tau}}{n^2} \qquad (2)$$

Where m_t represents the quantity adsorbed in the pellet at time t and m_∞ the quantity adsorbed in the pellet at final time. We denote $\overline{Q}(\tau) = m_t / m_\infty$. In practice m_t/m_∞ will be approximated by (Ruthven et al. 1980): $\overline{Q}_N(\tau) = 1 - \dfrac{6}{\pi^2}\sum_{n=1}^{N}\dfrac{e^{-n^2\pi^2\tau}}{n^2}$

Let us check by the following simple error estimation:

$$\begin{aligned} \left|\overline{Q}(\tau) - \overline{Q}_N(\tau)\right| &= \frac{6}{\pi^2}\sum_{n=N+1}^{\infty} e^{\frac{-n^2\pi^2\tau}{n^2}} \le \frac{6}{\pi^2}\sum_{n=N+1}^{\infty}\frac{1}{n^2} \\ &\le \frac{6}{\pi^2}\int_N^{\infty}\frac{dx}{x^2} = \frac{6}{\pi^2 N} \end{aligned} \qquad (3)$$

This inequality gives for $N=5$ an absolute error of approximately 0.12. But, notice that, in reality, the error is much smaller at least if τ stays away from zero. Indeed, if $\tau \ge \tau_0 \ge 0$, then the upper estimation can be improved to the following exponential error bound:

$$\left|\overline{Q}(\tau)-\overline{Q}_N(\tau)\right|=\frac{6}{\pi^2}\int_N^{\infty}\frac{e^{-n^2\pi^2\tau_0}}{x^2}dx\le\frac{3}{\pi^4\tau_0}\frac{e^{-N^2\pi^2\tau_0}}{N^3} \quad (4)$$

For example, if the reduced time is at least $\tau_0 = 0.01$ and if again $N = 5$ then we get for the absolute error a much better approximation of 0.000418.

3. Non Isothermal Assumptions

In this case, we assume that intracrystalline diffusion is the only significant resistance to the mass transfer. The sorbate concentration at the surface of each particle in the sample is therefore always in equilibrium, at the temperature of the sample with the sorbate concentration in the ambient fluid . In addition, the thermal conductivity is assumed to maintain an essentially uniform temperature throughout the entire adsorbent particle. The rate of heat transfer between the adsorbent and the ambient fluid which is assumed to obey Newton's law, was finite, so that during the course of the transient adsorption there was a time-dependent temperature difference between the adsorbent and ambient fluid. The system was subjected at time zero, to a small difference step change in ambient sorbate concentration from a previously established equilibrium condition. We assume that concentration equilibrium between fluid phase and adsorbed phase is established quickly at the free surface of the pellet and the equilibrium adsorbed phase concentration is assumed to vary linearly with both temperature and fluid phase concentration. Subject to these approximations, the system may be described by the following equation *(10)* in reduced value ($Q=q(t)/q\infty$), $\eta=r/r_c$ and $\tau=D_c t/r^2_c$).
In opposite to the previous part, the reduced value permitted to find an analytic solution in the Laplace domain.

$$\frac{\partial Q}{\partial\tau}=\frac{1}{\eta^2}\frac{\partial}{\partial\eta}\left(\eta^2\frac{\partial Q}{\partial\tau}\right) \quad (5\text{-a})$$

$$Q(\eta,\tau)=0,(\tau<0)\qquad Q(\eta,0^+)=0,(si0\le\eta<1) \quad (5\text{-b})$$

$$Q(1,\tau)=Qs(T) \quad (5\text{-c})$$

$$\left.\frac{\partial Q}{\partial\eta}\right|_{\eta=0}=0 \quad (5\text{-d})$$

$$\rho(-\Delta H)\frac{d\overline{Q}}{d\tau}=\frac{\rho Cp}{q_\infty-q_0}\frac{dT}{d\tau}+\frac{ha}{q_\infty-q_0}\frac{(T-T_0)}{D/r_c} \quad (5\text{-e})$$

The equilibrium relationship at the crystal surface is assumed to be linear with respect to both concentration and temperature:

$$Qs=1+\frac{\partial q^*}{\partial T}\left(\frac{T-T_0}{q_\infty-q_0}\right) \quad (5\text{-f})$$

The final result is similar to Ruthven(1984), and the mathematical discussion is described in L. Montastruc (2009) then:

$$\overline{Q}(\tau) = 1 + \sum_{n=1}^{\infty} \frac{-9\left(\frac{q_n \cot(q_n) - 1}{q_n^2}\right)^2 e^{-q_n^2 \tau}}{\frac{1}{\beta} + \frac{3}{2}\left[(q_n \cot(q_n))\left(\frac{(q_n \cot(q_n)) - 1}{q_n^2}\right) + 1\right]} \qquad q_n \in \Re \quad (6)$$

where q_n are the roots of the equation

$$\left(\alpha - q_n^2 - 3\beta\right)\sin(q_n) + 3\beta q_n \cos(q_n) = 0 \qquad (7)$$

With $\alpha = \frac{ha}{\rho Cp}\frac{r_c}{D}$ and $\beta = \frac{\Delta H}{Cp}\left(\frac{\partial q^*}{\partial T}\right)$

The parameters α and β were in a first approach variables used to solve the PDE system. These two parameters represent respectively the ratio between the conductivity and the heat capacity and the ratio between the quantity of energy and heat capacity.

4. Non Isothermal and Isothermal Comparison

In the non-isothermal model, the uptake curve commonly shows a rapid initial uptake followed by a slow approach to equilibrium because the external heat transfer controlled the adsorption. Let us recall that the isothermal model represents the ideal condition in the heat adsorption process.

In fact, in the isothermal case, the equilibrium was reached very quickly. The maximum difference between the two models is obtained in the first phase of the batch adsorption. This phenomenon is due to the fast temperature rising when the adsorption starts. The heat of adsorption produced in the pellet is not transferred into the fluid due to the external heat transfer. In the figure 1, different graphs represent the square distance between the adsorbed quantity with an isothermal model and a non isothermal assumption at different reduced time. To calculate, the N choice terms is 20, the α value is comprised between 0 to 40 and β between 0 to 10. If α is near zero and β is large, we observed that when the reduced time is lower that 0.25, the calculated errors increased with the reduced time. And when the reduced time is higher than 0.25, the calculated errors decreased.

To increase the process efficiency, it is necessary to raise the external heat transfer. It can be modified by the hydrodynamics around the pellet. The values of α and β can determine the "efficiency" of the heat adsorption process. In fact, if the β value is small and /or near to zero, the adsorption curve can be analyzed as in the isothermal case. But it means that the heat pump process cannot be operated because it will not produce enough heat. The α value is more significant, because it represents the ratio between the convection and the thermal capacity. To reach the ideal case, α should be very large. The external heat transfer will then not be a limiting phenomenon, and the adsorption is closed to the isothermal case.

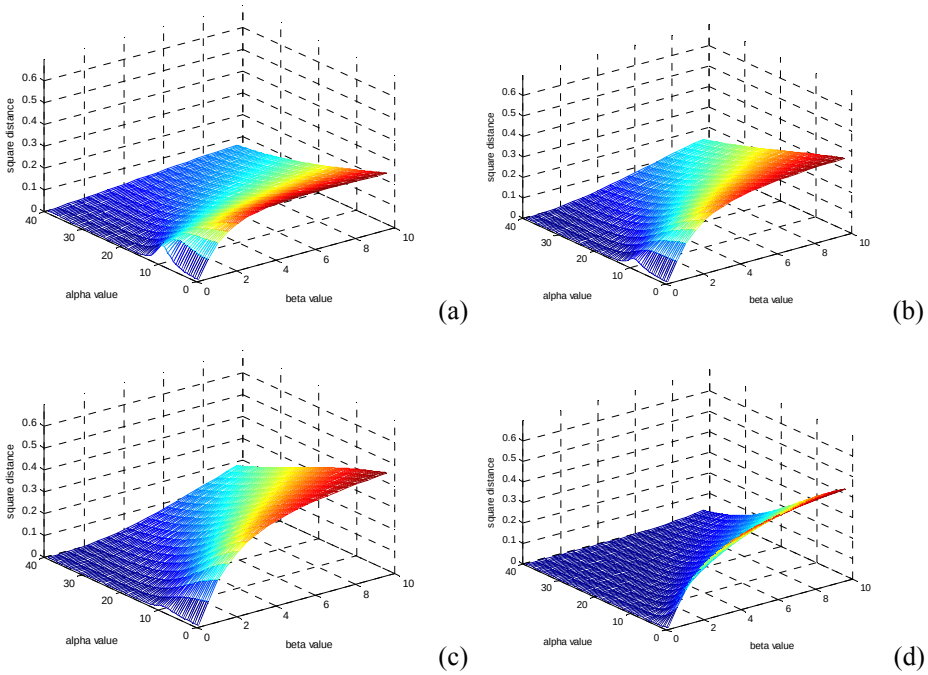


Figure 1. Square distance between the isothermal and the non isothermal model for different α value between 0 to 40 and for different β value between 0 to 10((a) reduced time = 0.01, (b) reduced time = 0.05, (c) reduced time = 0.1, (d) reduced time = 0.5)

5. Conclusion

In this article, we show the importance of the pellet adsorption study before designing a heat adsorption process. The difference between the two models - isothermal or non isothermal - could be measured by the value of two parameters α and β. In fact, when $\alpha \to \infty$ (infinitely high heat transfer coefficient) or $\beta \to 0$ (infinitely large heat capacity), the limiting case was isothermal. When diffusion was rapid (α small) the kinetics of sorption was controlled entirely by heat transfer. If the adsorption process might be used as a heat pump, it shall be represented by an isotherm model with α and β as higher as possible. Then, these two parameters could determine if the couple adsorbant/ adsorbate can be used in the case of an adsorption heat pump process.

References

Demir, H., M., Mobedi and S. Ülkü, A review on adsorption heat pump: Problems and solutions, Renewable and Sustainable Energy Reviews, 12 9, 2381-2403, 2008

L. Montastruc, P. Floquet, V. Mayer, I Nikov, S. Domenech , Kinetic modelling of isothermal or non isothermal in a pellet: application to adsorption heat pumps International Journal Energy Research (in Press, 2009)

Ruthven D.M., Principle of adsorption and adsorption processes, John Wiley & Sons, 1984

Ruthven D.M., L.P. Lee and H. Yucel, Kinetic of non-isothermal sorption in molecular sieve crystals, AIChE J. 26, 16–23, 1980

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

MP4SO: A Model-Partitioning Software for Simulation and Optimization

Alejandro O. Domancich[a,b], Melisa Maidana[b], Patricia M. Hoch[a,c], Nélida B. Brignole[a,b], Ignacio Ponzoni[a,b]

[a]*Planta Piloto de Ingeniería Química - CONICET*
Complejo CRIBABB, Camino La Carrindanga km.7 CC 717, Bahía Blanca, Argentina.
[b]*Laboratorio de Investigación y Desarrollo en Computación Científica (LIDeCC)*
Universidad Nacional del Sur, Av. Alem 1253, 8000, Bahía Blanca, Argentina.
[c]*Departamento de Ingeniería Química - Universidad Nacional del Sur, Av. Alem 1253, 8000, Bahía Blanca, Argentina.*
Email address:{adomancich, p.hoch}@plapiqui.edu.ar{ip,nbb}@cs.uns.edu.ar

Abstract

In this paper a software tool, which was called a Model Partitioning for Simulation and Optimization (MP4SO), is presented. It carries out the partitioning and solving of industrial-process models to be employed for simulation and optimization. It integrates GAMS' computational functionalities with the matrix-decomposition capabilities of the Direct Method. *MP4SO* was tested for the simulation and optimization of several industrial cases. The results show the advantages of using this partitioning-model approach in Process-Systems Engineering.

Keywords: Matrix Partitioning, Graph Theory, Simulation, Optimization, GAMS.

1. Introduction

The inherent complexity in the mathematical models employed for the simulation and optimization of industrial processes hinders their analysis and computational resolution tasks. A well-known strategy used to reduce the complexity of algebraic systems is their decomposition into simpler subsystems of equations [Duff *et al.*, 1986]. However, most of the methods based on this strategy are guided towards the resolution of square or lineal systems, thus limiting their application range. In particular, the simulation of industrial processes requires the mathematical resolution of square equation systems (non-singular systems), where the model contains the same number of equations and variables, being the solution that satisfies the system unique [Perkins, 1983]. On the other hand, optimization models are rectangular (singular), where the number of variables is greater than the number of equations, and they yield multiple possible solutions. It is important to stress that the models are strongly non-linear in both cases. These are the reasons why, if we wish to carry out a decomposition of these mathematical models, we need algorithms that allow us to work with square or rectangular models, with either linear or non-linear equations.

Recently, a new method for structural decomposition of equation systems for observability analysis was published. This algorithm, known as Direct Method (DM) [Ponzoni *et al.*, 2004], has characteristics that overcome the limitations presented by other decomposition algorithms. The DM, which is based on graph theory, carries out a structural reordering of the incidence matrix corresponding to the steady-state model of a process. This restructuring leads to a block-lower triangular form, which implicitly defines a block-decomposition and

the precedence order among blocks for the efficient resolution of the system of equations associated to the matrix. In contrast with other partitioning methods [Pothen & Fan, 1990; Gustavson, 1976; Madron, 1992, Duff & Reid, 1978], its main difference is that the DM allows the decomposition of both structurally singular and non-singular matrices. They may correspond to either linear or non-linear models. In this way, DM can be useful for the treatment of high-dimensional problems, as those that arise in process engineering. On the other hand, GAMS (General Algebraic Modeling System) [Brooke *et al.*, 2004] is a commercial equation-oriented tool for modeling, simulation and optimization of chemical processes. This software owns a great variety of mathematical solvers and it offers facilities to solve block subsystems, although it does not possess an integrated technique to carry out the decomposition. Then, it would be desirable to have a software that integrates the potentialities of the DM with the commercial software GAMS.
In this paper, we present a software called *MP4SO* (*Model Partitioning for Simulation and Optimization*), which is based on the equation-oriented strategy. *MP4SO* consists in a user-friendly tool that works in an interoperative way with GAMS, and bases its operation on the DM. The application of *MP4SO* to big problems facilitates overcoming the main disadvantages that the equation-oriented strategy naturally exhibits. Firstly, the partitioning of a complex system into subsystems achieves a decrease in the computational resources required to reach the solution. Secondly, the model partitioning facilitates the process analysis that should be carried out by the user. Some variables always have to be initialized for simulation or optimization studies. In these cases, this *MP4SO* feature allows an effective identification of the initial variables, by analyzing subsystems and their precedence order for the calculation. This feature is useful since the cumbersome procedure of initializing all the model variables through trial and error becomes easier.
In the next section, we present a brief description of the Direct Method as well as its integration with GAMS. Section 3 describes the *MP4SO*'s main components. Finally, in section 4 we present the performance analysis of the software and the general conclusions.

2. Using DM and GAMS for the simulation and optimization of industrial processes

A brief description of the DM is presented here. We also detail how the integration with GAMS was designed to use the capabilities of both tools so as to constitute *MP4SO*.

2.1. Decomposition of mathematical models using the Direct Method

The DM [Ponzoni et al., 2004] carries out a reordering of the structural incidence matrix corresponding to the steady-state model of a process, taking it to a block-lower triangular form. In a few words, this method is based on a graph-theory node classification technique similar to the one proposed by Dulmage-Mendelsohn[1963], but the reordering matrix pattern obtained by the DM is different and more suitable for our industrial application than Dulmage-Mendelsohn's method. In order to avoid a lengthy account, the main differences between both techniques as well as the DM algorithm are omitted here. Nevertheless, a detailed explanation about this topic can be found in Ponzoni et al. (2004).
The DM carries out a reordering of the structural incidence matrix corresponding to the steady-state model of a process, taking it to a lower-triangular block form. In this way, a subsystem decomposition is obtained, as well as a precedence order to carry out an efficient resolution of them. The technique constitutes a very robust and time-efficient matrix-partitioning approach. It is robust since it can be applied to any type of matrices, whatever their structural pattern, increasing their efficiency as problems grow in dimension and complexity. Therefore, it is possible to claim that this technique constitutes a solid foundation for the development of a methodology for the resolution of systems of

sparse non-linear equations. This reordering simplifies the resolution of the original system. Moreover, it is even applicable to rectangular systems.
The DM yields different kinds of variables according to the feasibility of their calculation:

- Observable variables: unmeasured variables that can be evaluated from the available measurements by using the equations.
- Unobservable variables: the remaining unmeasured variables.

Transferring this classification from the observability analysis to the mathematical resolution of models, observable variables are those calculated from the square blocks detected by the algorithm (simulation), while unobservable variables belong to rectangular blocks with more variables than equations (optimization).

2.2. Subsystem resolution using the Direct Method and GAMS
MP4SO is applied to solving simulation and optimization problems. Its operation is based on the application of the DM, followed by the resolution of the subsystems, by using some of the solvers provided by GAMS, which depend on the kind of problem to be solved. In a first step, the user feeds the mathematical model to *MP4SO*, by using GAMS notation. Then, the model is partitioned by means of the DM, thus obtaining a decomposition in subsystems and their resolution order. In a second stage, *MP4SO* invokes GAMS internally with the previously obtained decomposition. This invocation can be carried out for simulations (square systems) or optimizations (rectangular systems). Finally, the system is solved using built-in GAMS solvers, by carrying out the resolution of each subsystem.
From a mathematical point of view, the differences between a simulation and an optimization problem can be found in the structure of the equation system obtained for each case. The simulation of an industrial process involves the resolution of a square equation system with a unique solution. In terms of the structural analysis made by the observability algorithm, the same one gives a sequence of square subsystems conformed by the equations and variables of the model, which are completely classified as observable variables. Then, the simulation of this model implies carrying out the block-to-block resolution of all these subsystems. On the other hand, the optimization of a model gives a different structural analysis as a result, since this is a problem with more variables than equations. When analyzing the observability in these cases, a sequence of square blocks conformed by observable variables (independent of the optimization variables), followed by a rectangular block (singular) is obtained, which is integrated by observable variables belonging to previously computed blocks, but also unobservable variables; giving a block with multiple feasible solutions as a result. The model optimization should be carried out in two stages: a first one where the square blocks, which are independent of the optimization variables, are solved; and a second one, where the rectangular block is solved.

3. *MP4SO*: A software tool for the integration of DM and GAMS

Next, a description of the main menus that integrate *MP4SO* is included.
Model: This item controls the model analysis and summarizes its information clearly. It is made up with the following items:

- *Preprocessing*: It analyzes the mathematical model, specified with GAMS notation, by extracting the data required by the DM in order to carry out the decomposition.
- *Dimension*: It informs the model dimensions.
- *Parameters*: It shows a list with the model parameters, as well as their values.
- *Variables*: It shows a list with the model variables, as well as their values.

From this menu the user has the possibility of modifying the values of the parameters and variables. Then *MP4SO* automatically changes them on the model.

Partition: This menu allows the user to run the DM and carries out the model partitioning, clearly showing the obtained results. It has the following items:

- *Run*: It starts the partitioning using the Direct Method.
- *Subsystems*: In a simple way it informs the equations and variables that integrate each one of the obtained blocks. The blocks are colored in light blue and white (see Fig. 1).
- *Equations*: It shows the complete list of equations, following the order given by the DM.
- *Variables*: It shows a list of the variables arranged in accordance with the obtained partition. It also shows bounds and initial values. The user can modify these values here.

Optimization/Simulation: Depending on the nature of the problem (a simulation or an optimization), the user will use one of these drop-down menus. The menu options are:

- *Run*: It runs GAMS in order to solve the model.
- *Results*: It shows the corresponding GAMS output file.
- *Subsystems*: It shows all the blocks of the model, coloring in red (see Fig. 1) those blocks that, for any reason, could not be solved by GAMS (infeasible blocks).
- *Variables*: It shows the calculated values for the variables.

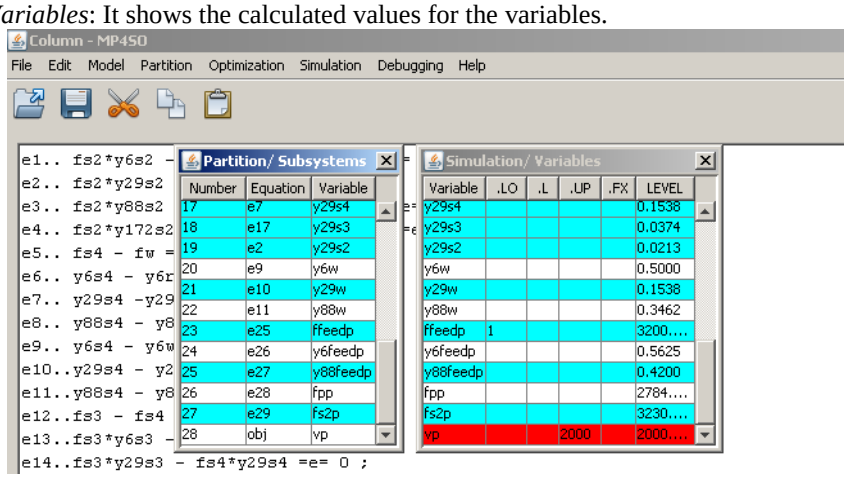


Figure 1: *MP4SO* screen capture. The figure shows the model written in the text area (background), as well as different windows (forefront) showing the DM partition (on the left) and the results obtained with GAMS (on the right).

Debugging: It is oriented to the treatment of infeasible subsystems, allowing the user to focus only on the infeasible blocks. For this purpose, *MP4SO* firstly solves the feasible part of the model, saving the results. Then, the user runs the problematic subsystem, changing initial values of variables until the block converges. Once the convergence is reached, the model is executed again looking for new infeasible blocks. It has the following items:

- *Start Debugging*: It saves the model in two different files. One with the part of the model that converged and the other one, with the infeasible block.
- *Run Feasible Part*: It runs the first file in GAMS, saving the obtained results.
- *Variables*: It shows a table with the infeasible-block variables, as well as their bounds and initial points. The user can change these values before trying to solve the block again.
- *Run Infeasible Part*: It runs the infeasible blocks in GAMS, using for the variables that were already calculated the results saved from the feasible part execution.

4. Performance software analysis and conclusions

In order to test *MP4SO*, cases of diverse dimensions and complexity were run. They correspond to mathematical models of real processes, which were in some cases simulated

and in others, optimized. For each of them, brief descriptions of the process and model dimensions are included, followed by the results obtained. For the simulation models, we detail the decomposition obtained by the software, as well as the percentage of variables that should be initialized in order to carry out the resolution in blocks. For the optimization, we specify the square blocks that *MP4SO* generated, as well as the dimensions of the optimization block. We also included data about both the percentage of the model that was assigned to simulation (square blocks) and the percentage of variables that should be initialized in order to reach the convergence of the whole model.

4.1. Process Simulation

4.1.1. Catalytic cracking of gasoil

This problem deals with the determination of the reaction coefficients for the catalytic cracking of gas oil into gas and other byproducts. This model is taken from the COPS benchmarking suite [Cops, 2009]. The mathematical model chosen to represent the plant consisted of 1299 equations and variables. Table 1 shows the results.

Nº Eq.	**Nº Var.**	**Blocks**			**% Initial. Variables**
		1x1	**12x12**	**285x285**	
1299	1299	102	76	1	**21**

Table 1. Resulting decomposition for the simulation of the catalytic cracking of gasoil.

4.1.2. Reactive distillation column

The reactive distillation column constitutes a clear example of process intensification. It can be presented as multifunctional reactors, where the reactive and separation tasks are combined into a single unit, thus reducing investment costs. The reactive distillation column is modeled as a plate column, using reactive and non-reactive stages where appropriate. A model description can be found in Domancich *et al.* (2009). The mathematical model of this section is made up of 1615 non-linear equations and variables. Table 2 shows the results.

Nº Eq.	**Nº Var.**	**Blocks**				**% Initial. Variables**
		1x1	**5x5**	**13x13**	**997x997**	
1615	1615	600	1	1	1	**61**

Table 2. Decomposition that was obtained for the simulation of the reactive distillation column.

4.2. Process optimization

4.2.1. Network of heat exchangers

This case represents a heat-exchanger network, where the main stream is cooled. To achieve this, we used a network of heat exchangers, which is divided in two parallel parts. The objective function is the summation of the exchange areas, which must be minimized. The optimization variable for this problem is the split fraction of the main stream. A more detailed description of the model can be found in Domancich *et al.* (2004). The model that represents the plant consisted of 58 equations and 59 variables. Table 3 shows the results.

Nº Eq.	**Nº Var.**	**Simulation Blocks**	**Optimiz. Block**		**% Mod. Simul.**	**% Init. Var.**
		1x1	**Eq.**	**Var.**		
58	59	40	18	19	**69**	**32**

Table 3. Resulting decomposition for the optimization of the heat-exchanger network.

4.2.2. HDA production

In this case we considered the hydrodealkylation of toluene to produce benzene. The objective function is the economic potential obtained from the sale of the different products. The optimization variables are the reactor conversion and the H_2 fraction in the purge stream

of the process. A more detailed description of the model can be found in Douglas (1988). This model is made up of 634 and 695 variables. Table 4 shows the results.

Nº Eq.	Nº Var.	Simulation Blocks				Optimiz. Block		% Mod. Simul.	% Init. Var.
		2x2	3x3	4x4	6x6	Eq.	Var.		
634	695	184	2	1	1	435	496	**32**	**68**

Table 4. Resulting decomposition for the optimization of HDA production.

4.3. Discussion and Conclusions

The resulting decompositions depend on the coupling degree that exists between the variables of the model, as well as on the kind of problem to solve. In an optimization problem, we are trying to minimize an objective function by fixing the different optimization variables of the model, giving a high coupling degree as a result. By applying *MP4SO* to an optimization problem, we can detect which part of the model is uncoupled from the optimization variables and the objective function. This is the part that the program is able to decompose. Then, a final block is generated, which will be optimized with all the equations that contain the optimization variables and the objective function.

The execution times for all the examples for both models, the original one and the partitioned one, were between 10 and 120 seconds executed with a Pentium 4 (3.00 Ghz). Most of the blocks obtained in the examples have small size (dimension 1 or 2). Therefore, once these blocks are solved, the user can focus on the larger ones. The bounding and initialization of variables in these blocks usually require a trial and error process. By comparing any partitioned model with the original one, the following advantages arise:

- A partitioned model demands less computational requirements to be solved.
- The number of variables that should be initialized to solve the partitioned problem is remarkably lower. Therefore, the cumbersome procedure of initialization is simplified.
- *MP4SO* simplifies the debugging stage, which is necessary when resolution errors appear. The software individualizes the infeasible block, in such a way that the user only needs to carry out the trial and error initialization procedure only for this block.

References

A. Brooke, D. Kendrick, A. Meeraus, R. Raman, 2004, GAMS: A User´s guide.

COPS http://www-unix.mcs.anl.gov/~more/cops

A. Domancich, J. Ardenghi, G. Vazquez, N. Brignole, 2004, Desempeño de un Algoritmo Espectral para Optimización Rigurosa de Plantas de Procesos Industriales. Mec. Comput., XXIII, 3047-3057.

A. Domancich, N. Brignole, P. Hoch, 2009, Structural Analysis of Reactive Distillation Columns, Virtual Pro (http://www.revistavirtualpro.com) N°84, pp. 18. ISSN 1900-6241.

I. Duff, A. Eriisman, J. Reid, 1986, Direct Method for Space Matrices, Oxford University Press.

I. Duff, J. Reid, 1978, An Implementation of Tarjan's Algorithm for the Block Triangularization of a Matrix, ACM Transactions on Mathematical Software, 4, 137-147.

A. Dulmage, A. L.; Mendelsohn, N. S. Two Algorithms for Bipartite Graphs. J. Soc. Ind. Appl. Math. 1963, 11, 183-194.

J. Douglas, 1988, Conceptual Design of Chemical Processes, McGraw-Hill chemical processes.

F. Gustavson, 1976, Finding the Block Lower-Triangular Form of a Sparse Matrix. In Sparse Matrix Computations; Bunch, J. R., Rose, D. J., Eds.; Academic Press: London, U.K.

F. Madron, 1992, Process Plant Performance, Measurement and Data Processing for Optimization and Retrofits, Ellis Horwood Ltd.: Chichester, England.

J. Perkins, 1983, Equation-oriented Flowsheeting, Second Conference on Foundations of Computer Aided Process Desin. Snowmass, Colorado.

I. Ponzoni, M. Sanchez, N. Brignole, 2004, A Direct Method for Structural Observability Analysis, Industrial & Engineering Chemistry Research,43, 577-588.

A. Pothen, C. Fan, 1990, Computing the Block-Triangular Form of a Sparse Matrix, ACM Transactions on Mathematical Software. 16, 303-324.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

An Effective Technique for the Synthesis and Optimization of Steam System Networks

T. Price[a] and T. Majozi [a,b]

[a]*Department of Chemical Engineering, University of Pretoria, Lynnwood Road, Pretoria 0002, South Africa, thoko.majozi@up.co.za*
[b]*Department of Computer Science, University of Pannonia, Veszprém, Hungary.*

Abstract

The use of steam in heat exchanger networks (HENs) can be greatly reduced by the application of heat integration and optimization with the intention of debottlenecking the steam boiler and indirectly reducing the water requirement (Coetzee and Majozi, 2008). By reducing the steam flowrate the return condensate temperature to the boiler is compromised which adversely affects the operation of the boiler. A means of maintaining the efficient operation of the boiler is to reheat the return flow to the boiler to a sufficiently high temperature. A means of achieving this is utilizing the sensible heat from the superheated steam which would normally pass through a let down valve.

Steam systems typically employ turbines to generate shaft work from superheated high pressure steam. The outlet of these turbines is frequently used as a heating utility in the background process. Since turbines operate at various steam levels, a means for incorporating these steam levels into the HEN optimization framework is necessary.

Consequently this paper concerns the optimization and restructuring of all steam system heat exchangers using conceptual and mathematical analysis to create a series HEN with the aim of reducing the overall steam flowrate, whilst maintaining the boiler efficiency.

Keywords: multiple steam levels, heat exchanger network, boiler efficiency.

1. Introduction

Pinch analysis has found numerous applications in a wide range of process integration areas, most specifically mass and heat integration. Heat integration has the ultimate goal of reducing external utilities by maximizing process to process heat exchange but can also be used in the optimal placement of utilities (Linnhoff and Hindmarsh, 1983).

The work on steam network synthesis by Coetzee and Majozi (2008) encompasses a graphical targeting technique on a T/H diagram. This method is used to find the minimum steam flowrate for a particular HEN. Using this flowrate an appropriate HEN is then designed accordingly. This network design is done using an LP model to construct the network. The whole process can also be done using an MILP model.

By incorporating multiple steam pressure levels the graphical targeting technique becomes obsolete as it cannot cater for the extra dimension of various steam pressure levels. Thus the MILP model derived by Coetzee and Majozi (2008) is altered to accommodate multiple steam pressure levels and is used to target for a minimum flowrate as well as design the network.

However, the effects of minimising the steam flowrate on the entire steam system have not been considered. The efficient operation on the steam boiler is dependent on the condensate return flowrate and temperature. Reducing the steam flowrate reduces both of these operation parameters and as such affects the steam boiler.

2. Problem Statement

The problem addressed in this investigation can be formally stated as follows.

Given:

- a steam boiler with known efficiency,
- a set of heat exchangers linked to the boiler with limiting temperatures and fixed duties,
- turbines operating at multiple pressure levels with fixed power outputs which deliver saturated steam to the set of background processes through a let down valve,

Determine the minimum steam flowrate and corresponding heat exchanger network that can be achieved by integrating the steam turbine exhaust into the steam system whilst maintaining the boiler efficiency.

3. Paper Approach

Figure 1(a) shows a typical steam system in the context of this investigation. High pressure steam is produced inside the steam boiler. A portion of the steam is sent to a high pressure turbine. The remaining steam passes through a let down valve and proceeds to the process. The exhaust from this turbine is then used as process heat as well as to recover more shaft work in a medium pressure turbine. The exhaust from the medium pressure turbine can be used further as process heat.

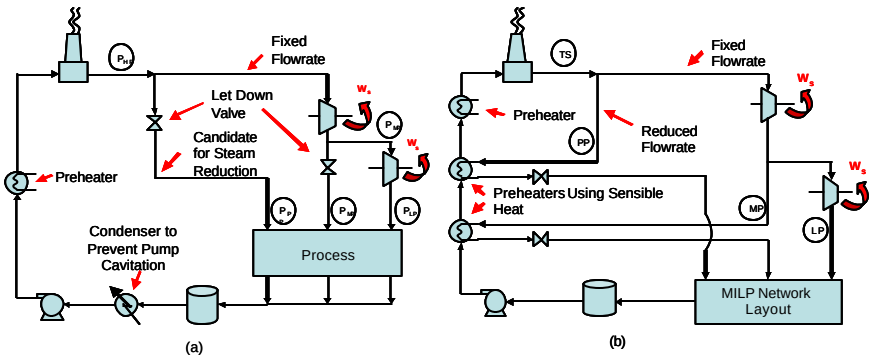


Figure 1: Steam system layouts before and after utilizing sensible heat.

3.1. Boiler Efficiency

Equation (1) relates boiler efficiency, $?_b$, to the effects of changing steam load, capacity and operating conditions as would be encountered in a realistic situation (Shang and Kokossis, 2004).

$$h_b = \frac{q\left(M/M^{\max}\right)}{\left(c_p \Delta T_{sat} + q\right)\left[(1+b)\left(M/M^{\max}\right)+a\right]} \tag{1}$$

In Equation (1), q is the heat load of the steam (i.e. the latent and superheated sensible heat), M is the steam load raised by the boiler and M^{max} is the capacity of the boiler. The parameters a and b are taken from a study by British Gas in work done by Pattison and Sharma (1980).

3.2. Methodology

The methodology consists of two parts. Firstly the minimum steam flowrate is determined and an associated heat exchanger network is found. The second part of the methodology deals with constraints concerning the boiler efficiency. An attempt is then made to reconstruct the HEN so as to maintain the initial boiler efficiency.

3.2.1. Steam Reduction and Initial Network Design

Figure 2 shows the superstructure that is used to derive the relevant mass and energy balances needed to determine the minimum steam flowrate for the particular HEN. These equations form an MILP model with the objective function being minimizing the steam flowrate. These equations are based on those of the MILP model of Coetzee and Majozi (2008) with considerations for the various steam pressure levels included.

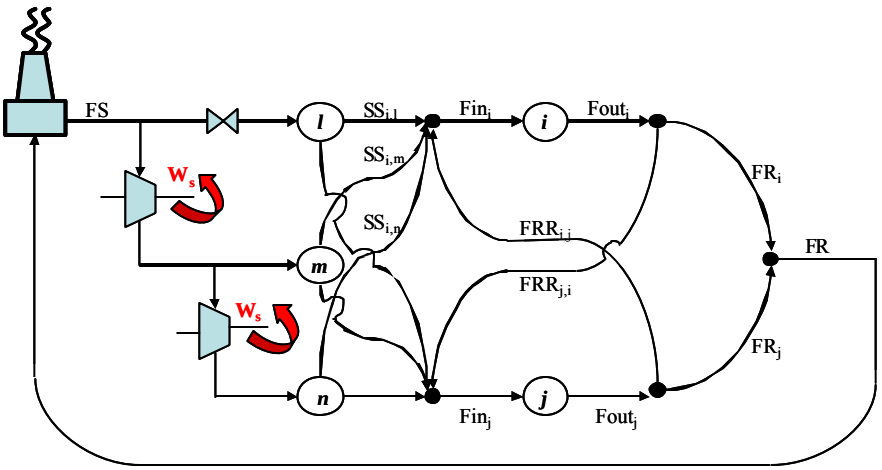


Figure 2: Superstructure used to derive mass and energy balance equations.

3.2.2. Boiler Efficiency Considerations and Altered HEN

To calculate the boiler efficiency as defined in Equation (1) several variables are required. The outlet temperature of the process must be known and can be calculated by Equation (2). The return flowrate to the boiler can take two forms, the first being saturated condensate and the other sub-cooled condensate, represented by $FRS_{i,l}$ and FRL_i respectively. Using these variables, the total return temperature to the boiler, T_{boil}, can be calculated with Equation (3). The efficiency can then be calculated using these variables in Equation (1), shown in Equation (4).

$$T_{proc} = \frac{\sum_{i \in I, l \in L} FRS_{i,l} Tsat_l + \sum_{i \in I} FRL_i Tout_i^L}{FS} \tag{2}$$

In Equation (2), T_{proc} is the process outlet temperature and FS is the total condensate flowrate.

$$T_{boil} = T_{proc} + \frac{Q_{preheat}}{FS \times c_p} \tag{3}$$

In Equation (3), T_{boil} is the return temperature to the boiler and $Q_{preheat}$ is the heat added to the return stream by the preheater.

$$h_b = \frac{q\left(FS/M^{\max}\right)}{\left(c_p\left(T_{sat} - T_{boil}\right) + q\right)\left[(1+b)\left(FS/M^{\max}\right) + a\right]} \tag{4}$$

Since the steam flowrate reduction causes a decrease in the return boiler temperature a means of reheating the boiler feed to its original return temperature must be found. It is suggested that the boiler return condensate be heated in heat exchangers which utilize the sensible heat of the superheated steam from the boiler and from the high pressure turbine exhaust. Thus using this method the energy can be reclaimed and used to maintain the boiler efficiency. Figure 1(b) shows the new steam system structure incorporating these changes. The new boiler return condensate temperature is then calculated using Equation (5).

$$T_{boil} = T_{proc} + \frac{(PP)\left(h_{\sup hp} - h_{sat,hp}\right)q + MP\left(h_{\sup mp} - h_{sat,mp}\right)j}{FS \times c_p} \tag{5}$$

In Equation (5), $h_{sup,hp}$ and $h_{sup,mp}$ are the enthalpies of the superheated high and medium pressure steam leaving the boiler and high pressure turbine respectively. $h_{sat,hp}$ and $h_{sat,mp}$ are the enthalpies of the saturated high and medium pressure steam at the boiler and high pressure turbine outlet conditions respectively. ? and f are the fractions of the sensible energy that can be used safely without the risk of condensation for the high and medium pressure steam respectively. These fractions are arbitrarily chosen with a large enough safety factor that is deemed appropriate so as to not compromise the saturated steam by condensation. For this investigation they were chosen as 80%.

Equations (2) to (5) are used to create the second part to the model. The MINLP model formed is solved using the technique of Quesada and Grossmann, (1994). Two cases are considered, each using the same basic equations but focusing on two objectives.

Case 1: Maintain boiler efficiency with slight compromise in minimum flowrate
Firstly the primary objective can be to maintain boiler efficiency. This can mean that the minimum steam flowrate may not be reached, if there is not enough sensible heat available to preheat the boiler feed stream. The minimum steam flowrate is first determined using the first part of the methodology. Then a slack variable is added to this steam flowrate and the equations relating to boiler efficiency are added. The boiler efficiency is then fixed while the slack variable is minimized, thus the total steam flowrate can be found with constant boiler efficiency.

Case 2: Maintain minimum flowrate with slight compromise in boiler efficiency
Secondly the objective could be to achieve the minimum steam flowrate with only a slight compromise in boiler efficiency. A slack variable is added to the efficiency

equation and minimized. If the slack variable is zero then the boiler efficiency can be maintained with the minimum steam flowrate. If it is not then the boiler efficiency will be compromised by minimizing the steam flowrate.

Either of these two cases can be explored depending on the circumstances of the steam system being examined. For example, arid regions may have a shortage of water and thus minimizing the steam flowrate could be the primary objective, or where water is in abundance maintaining the boiler efficiency can be used as the objective.

3.3. Case Study

The case study presented by Coetzee and Majozi, (2008) is used here to show how boiler efficiency is affected by a reduction in steam flowrate and how the formulations above can be used to maintain the original efficiency.

Using Equation (4) the boiler efficiency was calculated as 63.4%. After the first part of the methodology the steam flowrate was reduced from 10.73kg/s to 7.91kg/s, a 26.3% reduction. Figure 3(a) shows the network layout that corresponds to this flowrate. The HEN outlet temperature is consequently reduced to 63°C. This reduction in flowrate and temperature result in a new boiler efficiency of 58.2%, an 8.2% reduction.

Using the first premise of maintained boiler efficiency it was found that the boiler efficiency could be sustained without compromising the steam flowrate. This was successfully accomplished using 60.1% of the available sensible heat from the HP steam and 80% from the MP steam. With the reduction in steam flowrate, the return temperature had to be increased to 119.0°C.

Since the boiler efficiency could be maintained for the minimum steam flowrate it follows that the second premise would also yield the same answer.

3.4. Sensitivity Analysis

Both formulations gave the same result since there was enough sensible energy to heat the boiler feed to the point where the original efficiency could be maintained. By reducing the amount of sensible heat the models may show how they compromise either efficiency or minimum flowrate to maintain the other. Thus the amount of sensible heat available was reduced to 30% for both the HP and MP sensible heat sources.

To maintain boiler efficiency the minimum steam flowrate was indeed compromised. The new steam flowrate was found to be 8.33kg/s, 8.5% higher than the flowrate after the initial reduction but still yielded a 22.4% reduction from the original parallel arrangement. This new flowrate did require a new heat exchanger network to satisfy the duties of the various process streams. This network is seen in Figure 3(b).

Using the second objective to utilize the minimum steam flowrate the available sensible heat was once again reduced to 30% of the original amount. It was found that the minimum flowrate could be maintained but a small decrease in boiler efficiency was noted. The new boiler efficiency was calculated as 0.6140, a decrease of 3.2% from the original boiler efficiency. The process outlet temperature increased to 70.6°C, while the

boiler return temperature dropped to 96.8°C. The network required for this flowrate is the same as that of Figure 3(a), since there is no change in flowrate.

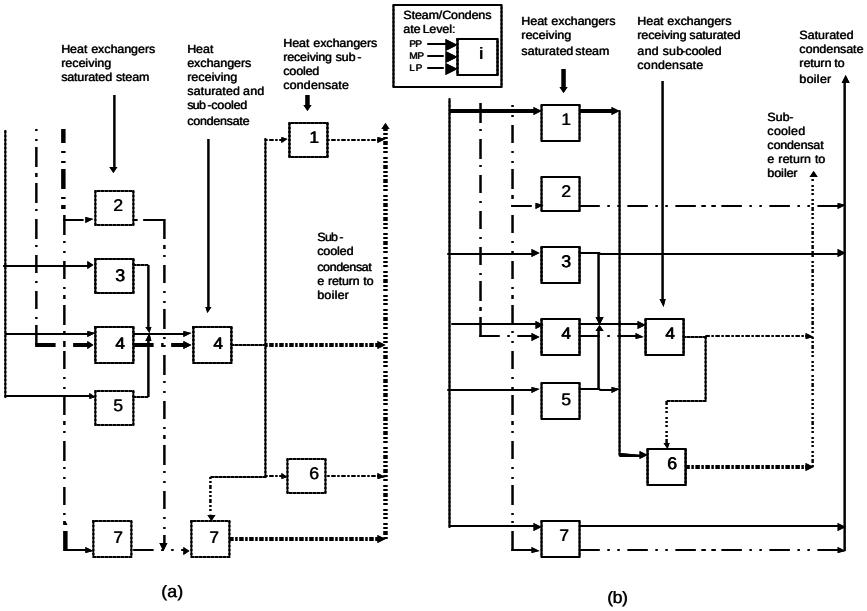


Figure 3 : HEN layouts before and after steps are taken to maintain boiler efficiency.

4. Conclusions

The following conclusions can be made with regard to incorporating multiple steam pressure levels into maintaining boiler efficiency while reducing steam flowrate:

- Preheating the return flow to the boiler to a slightly higher temperature will maintain the boiler efficiency for a reduced flowrate.
- Some of the excess exhaust from the turbines can be used in the heat exchanger network to further reduce the steam flowrate.
- In the event of insufficient sensible heat from the HP and MP steam a slight compromise in either the minimum flowrate or the boiler efficiency is necessary.

References

1. Coetzee, W.A. and Majozi, T. (2008) Steam System Network Design Using Process Integration, Ind. Eng. Chem. Res 2008, 47, 4405-4413.
2. Shang, Z. and Kokossis, A. (2004), A transshipment model for the optimisation of steam levels of total site utility system for multiperiod operation, Computers and Chemical Engineering 28, pages 1673-1688.
3. Quesada, I. and Grossmann, I. E. (1994) Global optimization of bilinear process networks with multi component flows, Computers and Chemical Engineering 19, No 12, pages 1219-1242 (1995).
4. Linnhoff, B. and Hindmarsh, E. (1983) The pinch design method for heat exchanger networks, Chemical Engineering Science 38, No. 5, pages 745-763.
5. Pattison, J. R. and Sharma, V. (1980) Selection of boiler plant and overall system efficiency. Studies in energy efficieny in buildings, British Gas.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Kinetic models for the homogeneous alkaline and acid catalysis in biodiesel production

José F.O.Granjo,[a] Belmiro P.D. Duarte,[a,b] Nuno M.C. Oliveira[a]

[a]*GEPSI–PSE Group, Department of Chemical Engineering, University of Coimbra, Rua Sílvio Lima – Pólo II, 3030-790 Coimbra, Portugal*
[b]*Department of Chemical and Biological Engineering, ISEC – Polytechnic Institute of Coimbra, Rua Pedro Nunes, 3030-199 Coimbra, Portugal*

Abstract

In this work, kinetic models were obtained from experimental data in the open literature, for both the alkaline (NaOH) and acid (H_2SO_4) homogeneous catalysis of the transesterification reaction, used to produce biodiesel from vegetable oils. Two approaches, designated as statistical and empirical, were adopted to obtain these models. For both catalysis types, the kinetic models obtained show average absolute errors of approximately 4%, using both empirical and statistical approaches, well within the precision of the experimental procedures. In the alkaline catalysis, kinetic models fitted better a pseudo 2nd-order reaction, while for the acid catalysis pseudo 1st-order reaction provided a best fit.

Keywords: Biofuels, Biodiesel, Reaction Engineering, Kinetic models, Model validation.

1. Introduction and Motivation

In EU-27 and Portugal, road transport was responsible in 2005 for about 60% and 58% of the oil demand, and approximately 25 % and 33% of the final energy consumption, respectively (European Communities, 2008). These figures represent a significant ecological footprint, with great impact on the countries' GDPs and constitute a geostrategic weakness. To mitigate this scenario, and also to comply with the Kyoto's goals and guidelines, the European Directive 2003/30/EC, transposed to Portuguese legislation in 2006 has established a series of measures to promote a more intense use of biofuels, in order to achieve a share of 5.75 % in the transport sector, by 2010. Biodiesel (a mixture of fatty acid methyl esters), a substitute of fossil diesel, is one possible and attractive alternative, because of its simple and mature technology of production, and well known benefits to the engine performance and environment.

Biodiesel can be produced from distinctive renewable sources (edible and non-edible), e.g., greases, vegetable oils or animal fats, through a transesterification reaction with methanol, leading to glycerol as a main by-product, also with some interesting industry applications. Many process alternatives are available for biodiesel production, which can be classified by the type of catalytic system used in transesterification reactions: homogeneous – alkaline or acid; heterogeneous – alkaline, acid or enzymatic; non-catalytic – supercritical alcohol. Several studies regarding the kinetics of transesterification and other aspects of the biodiesel production technology have been made. Ma et al. (1999), Vasudevan and Briggs (2008) and Demirbas (2009), among others, provide insightful critical reviews. Despite the considerable amount of work done with the aim of characterizing the kinetics of this system, no systematic study

integrating experimental data from different authors is currently available in the open literature. This drawback creates difficulties in the comparison of the performance of diverse catalysts at different reaction operation conditions, and in the comparison of alternative design configurations and setup conditions for this process, based on systematic process optimization methodologies.
To overcome these issues, the objective of this study is therefore to provide results relative to kinetic models based on experimental data available in the open literature, for the transesterification reaction of vegetable oils, through homogeneous alkaline and acid systems (the most common technologies used in industry).

2. Methodology

2.1. Assumptions and considerations
Earlier works (Noureddini and Zhu, 1997; Jeong and Park, 2006; Zheng *et al.*, 2006) studied the mass transfer limitations in the reaction kinetics and concluded that the heterogeneous effect is more pronounced for low intensity mixing rates and at the beginning of the reaction, leading to the formation of a lag phase and lower reaction rates. However, they also noted that with an intense mixing intensity (N_{Re}>6000) and the formation of biodiesel (an emulsion agent), the reaction becomes chemically controlled, and therefore the mass transfer resistance between the oil and alcohol phases is no longer important. Therefore, in this situation the reactant mixture can be considered as pseudo-homogeneous.
As described in Ma *et al.* (1999) and Marchetti (2007), the alcoholysis reaction consists in a group of parallel and reversible reactions. To prevent low yields of biodiesel and to avoid the formation of intermediary chemical components, an excess of alcohol is commonly used to push the equilibrium towards the products. Normally, a minimum molar ratio methanol/oil of 6:1 is chosen, making the reaction behave as practically irreversible, with a pseudo single step mechanism (Freedman *et al.*, 1984).
Freedman *et al.* (1984), Ma *et al.* (1999), and Canacki and van Gerpen (1999) reported that the free fatty acid (FFA) and water contents may have a significant inhibition effect on the reaction, for both catalysis types. In the alkaline catalysis, it was stressed that the FFA and water contents in oil should be below 0.5 % (w/w) and 0.3 % (w/w), respectively, to avoid soap formation that consumes catalyst and reduces the yield. With the acid catalysis, it was concluded that water and FFA strongly inhibit the reaction for concentrations above 0.5 % (w/w) and 5% (w/w), respectively. Water promotes the hydrolysis of triglycerides, fatty acids or esters, which is undesirable because it reduces the conversion efficiency of biodiesel. Hence, in industry, pretreatments are usually included to eliminate water and FFA in the transesterification reactors, preventing this negative effect (Zhang *et al.*, 2003).
To obtain kinetic expressions for the acid and alkaline catalysis as functions of the operating variables, the following major assumptions were considered: intense mixing rate; excess of methanol; negligible water and free fatty acid contents. Literature data was selected to fulfill the former conditions; consequently, the kinetic models obtained are only applicable to those situations.

2.2. Problem formulation
The procedure used to obtain the kinetic models for the acid and alkaline homogenous catalysis is now described. Considering the conversion of the limiting reactant (A), the mass balance for a batch system with a *nth* order of reaction is $dX/dt = kC_{Ao}^{n-1}(1-X)^n$.

At constant temperature, for the 1st and 2nd orders, if the initial time and conversion are both zero, k can be evaluated using values of conversion X_i and time t_i from:

$$k_i^{'} = \frac{-\ln(1-X_i)}{t_i} \quad [\text{min}^{-1}], \qquad k_i^{''} = \frac{X_i}{(1-X_i)C_{Ao,i}t_i} \quad [\text{min}^{-1} \cdot \text{w(oil+methanol)/w oil}] \tag{1}$$

These expressions were used with available experimental data and fixed operating variables, such as temperature (T) in Kelvin, catalyst concentration (xa) in % (w/w oil) and dimensionless molar ratio methanol/oil (xm). In this case, oil is the limiting reagent, since an excess of methanol is used.

Two approaches, here designated as *statistical* and *empirical*, were used to fit the experimental data to the mathematical models, for each catalysis type. In the statistical approach, each experimental data set was used to fit the dependent variables to a set of possible predictors, using linear regression. The dependent variables considered were the kinetic constants for first ($k^{'}$) and second ($k^{''}$) order reactions, and their respective logarithmic transformations. As predictors, the main operating variables (T , xa , xm), and their inverse, logarithmic and square roots transformations were considered. The quadratic terms of these predictors and their cross products were also included. For each dependent variable, a multivariate linear regression model was computed, using two different procedures to select the best set of predictors. Both a forward stepwise regression procedure (Al-Subaihi, 2002) and an optimization procedure based on the direct minimization of the Bayesian Information Criterion (BIC) (Duarte *et al.*, 2008) were implemented. During this procedure, data points that presented a standard residual greater than 2σ were considered outliers, and were consequently eliminated.

In the empirical approach, equation rates of the form $r_j = \alpha_j k(T,xa,xm) C_{oil0}^{\beta_A}(1-X)^{\beta_A}$, with

$$k(T,xa,xm) = A \cdot e^{-\frac{Ea}{R \cdot T}} \cdot \prod_{i=1}^{n} \gamma_i^{\beta_i} = f(T) \times f(xa) \times f(xm) \tag{2}$$

were considered. A first estimate of $k(T,xa,xm)$ was obtained through evaluation of the functions $f(T)$, $f(xa)$ and $f(xm)$ from careful chosen subsets of experimental data, changing only one variable at a time. For instance, to obtain $k(T,xa,xm)$ with a first order reaction in the acid catalysis, $f(xm)$ can be determined by choosing data with $xa = 3$ and $T = 333.15$ K. The partial expressions obtained were sequentially used to estimate the remaining terms. These tests also provided extremely valuable information relative to the required shape of each term $f(\bullet)$, for a good global fit. A first guess of the function $k(T,xa,xm)$ could therefore be obtained for both types of catalysis, considering both first and second reaction orders. Afterwards, the initial parameter values were used as initial guesses for a more rigorous nonlinear regression, considering all data points simultaneously. This task was performed by formulating a nonlinear least square optimization problem, employing a Least Squares criterion with respect to the residuals of the conversion. To prevent convergence to local optima, the numerical solvers CONOPT, OQNLP and BARON were used sequentially. At the end, a statistical analysis was completed for each empirical model.

3. Results

The alkaline catalyst studied was sodium hydroxide and the experimental data was retrieved from Freedman *et al.* (1984), Noureddini and Zhu (1997), Jeong *et al.* (2004) and Leung and Guo (2006). Most kinetic studies with alkaline catalysis use xm =6 (methanol/oil=6:1); therefore, in this case $f(xm) = 1$ was considered for all experiments (since no information relative to this factor could be extracted from the data), and the kinetic constant obtained is only a function of the temperature and catalyst concentration ($k(T, xa)$). In the acid catalysis, all the studies considered used sulfuric acid; experimental values from Freedman *et al.,* (1984), Canacki and van Gerpen (1999), Goff *et al.* (2004), and Zheng *et al.* (2006) were considered. Table 1 presents the correlation factors and the mean absolute relative errors for each catalysis type, modeling approach and model order.

Table 1. Mean absolute relative error of the conversions and correlation factors for alkaline and acid catalysis, using the statistical and empirical strategies, for different model orders.

Alkaline catalysis	**1st order**		**2nd Order**	
	Empirical	**Statistical**	**Empirical**	**Statistical**
Mean error / %	8.05	8.28	4.14	4.03
R^2	0.9919	0.9327	0.9975	0.9732
Acid catalysis	**1st order**		**2nd Order**	
	Empirical	**Statistical**	**Empirical**	**Statistical**
Mean error / %	4.64	4.21	8.30	4.10
R^2	0.9970	0.9900	0.9957	0.9792

For the alkaline catalysis, the range of experimental values considered was 32–70 ºC for T and 0.2–1% for xa. With the acid catalysis, the range of experimental values used 60-100ºC for T, 1-46.5% for xa, and 3.3–245 methanol/oil molar ratio (xm). Therefore, the kinetic models should only be used for reliable predictions within these boundaries. In the empirical approach and for both kinetic orders, the numerical solvers CONOPT, OQNLP and BARON were used sequentially, to ensure that the parameter values obtained correspond to global optima.

Table 1 shows that data from literature fits better to a 2nd order kinetic model for the alkaline catalysis, and a 1st order kinetic model with acid catalysis, since the mean error for both strategies are lower in these cases and the correlation factors higher. For the alkaline catalysis, the 2nd order kinetic models expressions obtained in the empirical and statistical approaches are given in Eqs. (3) and (4). The corresponding parameter values are presented in Table 2.

$$k(T, xa) = e^{\beta_0 - \frac{\beta_1}{T}} (-1 + \beta_2 xa + \beta_3 xa^2 + \beta_4 xa^3) \tag{3}$$

$$k(T, xa) = e^{\beta_0 + \beta_1 \ln(T / T_0) xa + \beta_2 (T_0 / T) \ln(xa) + \beta_3 \ln(xa) + \beta_4 2 \ln(T / T_0)}, \; T_0 = 300 \text{ K} \tag{4}$$

Table 2. Parameter values for Eqs. (3-4) considering a 2nd order reaction, using data from Freedman *et al.*(1984), Noureddini and Zhu (1997), Leung and Guo (2006), and Jeong *et al.* (2004).

	Empirical		**Statistical**	
	Parameter values	**Standard error**	**Parameter values**	**Standard error**
β_0	1.385E1	1.759E0	-1.2522E0	1.788E-1
β_1	5.362E3	4.568E2	2.372E1	4.002E0
β_2	2.160E1	5.973E0	1.198E1	2.171E0
β_3	-4.898E1	1.342E1	-1.052E1	2.116E0
β_4	4.401E1	1.400E1	-5.725E1	2.589E1

Table 3 displays the best parameter values obtained for the acid catalysis. The corresponding 1st order kinetic model expressions obtained with the empirical and statistical approaches are presented in Eqs. (5) and (6), respectively. Here, the complexity of the model obtained with the statistical approach advises against its use, due to the limited number of experimental data points used. In this case, Eq. (5) corresponds to a more parsimonious form, and should be used for the acid catalysis kinetic rate.

Table 3. Parameter values for Eqs. (5-6) considering a 1st order reaction, using data from Freedman *et al.* (1984), Canacki and Gerpen (1999), Goff *et al.* (2004) and Zheng *et al.* (2006).

	Empirical		**Statistical**	
	Parameter values	**Standard error**	**Parameter values**	**Standard error**
β_0	2.389E1	1.293E0	-9.494E0	2.047E-1
β_1	9.913E3	4.424E2	5.682E0	1.221E0
β_2	4.004E-3	3.785E-3	-8.946E-2	1.347E-2
β_3	1.865E-1	3.173E-2	2.163E1	3.805E0
β_4	1.379E0	1.172E0	1.265E-1	2.685E-2
β_5	5.739E-1	3.640E-1	-6.391E0	1.683E0
β_6	———	———	8.545E-2	2.277E-2
β_7	———	———	-5.858E-2	1.690E-2
β_8	———	———	-1.296E-1	3.853E-2

$$k(T, xa, xm) = e^{\beta_0 - \frac{\beta_1}{T}}(1 - e^{\beta_2 - \beta_3 xa})(1 - e^{\beta_4 - \beta_5 xm}) \tag{5}$$

$$k(T, xa, xm) = \exp(\beta_0 + \beta_1 \ln(T / T_0)\ln(xm) + \beta_2(1/xa)xm + \beta_3(T_0 / T)\ln(T / T_0) + \\ + \beta_4(1/xa)^2 + \beta_5(1/xa)(1/xm) + \beta_6(T_0 / T)xm + \beta_7(T / T_0)xm + \beta_8 \ln(xa)\ln(xm)) \tag{6}$$

4. Conclusions

The results obtained show that a good agreement between the available kinetic data and models proposed was reached; for both catalyst types, absolute relative errors of approximately 4% could be achieved, with limited model complexity. The best results were obtained considering a 1st order model kinetics for the acid catalysis, and a 2nd order kinetics in the alkaline case. The models obtained can either be used for preliminary process design using systematic methodologies or for delimiting operating regions where additional kinetic studies should be concentrated to reach models with better predictive performance.

References

A.A. Al-Subaihi, 2002, Variable Selection in Multivariable Regression Using SAS/IML, Journal of Statistical Software, 7, 12, 1–20.

M. Canakci, J.V. Gerpen, 1999, Biodiesel production via acid catalysis, Transactions of the ASAE 42, 5, 1203–1210.

A. Demirbas, 2009, Progress and recent trends in biodiesel fuels, Energy Conversion and Management, 50, 1, 14–34.

B.P.M. Duarte, M.J. Moura, F.J.M. Neves, N.M.C. Oliveira, 2008, A mathematical programming framework for optimal model selection/validation of process data, in: B. Braunschweig, X. Joulia (Eds.), 18th European Symposium on Computer Aided Process Engineering – ESCAPE 18, 343–348.

European Communities, 2008, EU Energy and transport in figures, Directorate-General for energy and transport.

B. Freedman, E.H. Pryde, T.L. Mounts, 1984, Variables affecting the yields of fatty esters from transesterified vegetable oils, Journal of the American Oil Chemists' Society, 61, 10, 1638-1648.

M. Goff, N. Bauer, S. Lopes, W. Sutterlin, G. Suppes, 2004, Acid-catalyzed alcoholysis of soybean oil, Journal of Americal Oil Chemists' Society, 81, 4, 415–420.

G.-T. Jeong, D.-H. Park, 2006, Batch (One- and Two-Stage) Production of Biodiesel Fuel From Rapeseed Oil, in J.D. McMillan (Ed.), Twenty-Seventh Symposium on Biotechnology for Fuels and Chemicals, 668–679.

G.-T. Jeong, D.-H. Park, C.-H. Kang, W.-T. Lee, C.-S. Sunwoo, C.-H. Yoon, B.-C. Choi, H.-S. Kim, S.-W. Kim, U.-T. Lee, 2004, Production of biodiesel fuel by transesterification of rapeseed oil, Applied Biochemistry and Biotechnology, 114, 1, 747–758.

D. Leung, Y. Guo, 2006, Transesterification of neat and used frying oil: Optimization for biodiesel production, Fuel Processing Technology, 87, 10, 883–890.

F. Ma, L.D. Clements, M.A. Hanna, 1999, The effect of mixing on transesterification of beef tallow, Bioresource Technology 69, 3, 289–293.

J. Marchetti, V. Miguel, A. Errazu, 2007, Possible methods for biodiesel production, Renewable and Sustainable Energy Reviews, 11, 6, 1300–1311.

H. Noureddini, D. Zhu, 1997, Kinetics of transesterification of soybean oil, Journal of Americal Oil Chemists' Society, 74, 11, 1457–1463.

P. Vasudevan, M. Briggs, 2008, Biodiesel production – current state of the art and challenges, Journal of Industrial Microbiology and Biotechnology, 35, 5, 421–430.

Y. Zhang, M.A. Dubé, D.D. McLean, M. Kates, 2003, Biodiesel production from waste cooking oil: 1. Process design and technological assessment, Bioresource Technology, 89, 1, 1–16.

S. Zheng, M. Kates, M. Dubé, D.D. McLean, 2006, Acid-catalyzed production of biodiesel from waste frying oil, Biomass and Bioenergy, 30, 3, 267–272.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Development of a Software to Analysis and Prepare Mass Balance Data to Reconciliation in Oil Refineries

Camila L. R. Santos [a], Douglas C. Mariani [a], Alexandre A. Campos [a,] Marcelo Malavazi Alves [a]

[a] *Soteica do Brasil, R. Dr. Cid de Castro Prado, 108, Planalto Paulista, CEP 04064-040, São Paulo – SP, Brazil*
E-mail address: camilaramos.soteica@petrobras.com.br

Abstract

The data utilized on daily mass balance in oil refineries are generated from the records of products movements and flow and quality measurements existing in the refining processes. This data usually shows low quality and inconsistencies and it should be examined and treated. It was used a technique of data reconciliation to assess those inconsistencies and generate a set of data that allow reliable analysis, which support decisions about process. The data generated by data reconciliation are going to feed the accounting systems and support decisions of the operation and analysis teams. In order to improve data reconciliation system management data, it became more agile and practical, it is necessary to treat raw data that come from site and were recorded in the refinery information systems. The software developed has a function of capture raw data record in the system of registration and control of operations and link these data with measurements in the refining processes. This pooling of information brings a gain of quantity and quality to the data reconciliation system, reducing mathematical problems of non-solubility drive by configuration of refining processes and allocation of measurements. The software also evaluates the closing of balance sheet of each tank and carries a previous distribution of differences, which come from low precision of measurements. This functionality of system intended that data were as close to reality of the process as possible and also attend accounting requirements, increasingly rigid, imposed to oil refining companies. To achieve this goal, it was made modifications on a consecrated numerical technique, in order to adapt this to the problem and to the singular objective function that mathematically describes the problem. All these analyses are generated by software algorithms and occur automatically in hours previously scheduled. With the data processed in the case of mathematical problems and evaluated in order to attend mass balance constraints, reports are generated and will be used by analysis accounting team of refinery. This software is deployed in the largest Brazilian oil refinery and the data generated through it as used as the basis for the accounting closure of this site.

Keywords: mass balance, software development, data analysis.

1. Introduction

For an efficient management of the process industries, especially in oil refining companies, it is necessary reliable process information that supports each decision.

This kind of information is only possible to achieve when reliable balances exist in each process unit and in all industrial plant. These balances generate reliable information, without gross errors arising from measurement problems, which were removed by appropriate treatment of data.
This information is fundamental for decision support. These data will feed planning and scheduling production systems, supervision teams, inventory and tax control systems and will compose costs and other features.
Compiling these information, there will be defined what and where to produce, buy or sell a given product volume, revamps, local market, and others.
To generate this kind of information one of the techniques available is data reconciliation that consists in a statistic technique that minimizes difference between data and gives a better estimate to mass balance.
These techniques are being used at Petrobras oil refineries through utilization of software called Sigmafine, which was developed by OSI Software, sold and operated in Brazil by Soteica team.
Use of data reconciliation has varied applications when it is needed to treat the inaccuracies of measuring instruments on a statistical methodology which converge in a more reliable mass balance for a process or equipment.
The work of Weiss (1996) presents the use of this technique to a pyrolysis reactor, in this work the problem was developed around a simplified energy and mass balance that was solved by successive linearizations. The results obtained were used in survey of the equipment operational parameters.
On this basis has the work of Placido (1998), which applies an algorithm developed for reconciliation, in language Fortran 77 to ammonia's plant of FAFEN/BA. This study showed the importance of knowledge of the plant, the need to have sufficient redundancy and the importance of addressing the non-solubility that was found.
Bonfim (2004) shows the importance of data reconciliation for the Petrobras. This importance is linked to the need to obtain reliable process data that can be used to take decisions.
In applying this technique in the refineries of Petrobras, the main problem is the presence of non-solubility alignments that result in occurrence of errors or coarse spreading of errors in the refineries models.
To combat this problem Soteica's team has developed the software that uses data from the refining units to estimate the logistic operations, reducing the problems on non-solubility, improving data quality and reducing reconciliation time.
On this work, it will be described the development and results achieved by Soteica's team with this tool on the implementation on Paulinia's Refinery. The largest oil refinery in Brazil and Petrobras.

2. Problem Description

The main problem found on reconciliation process was the logistic operation originated in tanks that are receiving some products and sending to other places at the same time. These logistic operations are classified as simultaneous.
Thus the movement quantity of these logistic operations cannot be determined by the tanks' gap, because more than one operation influences this gap.
So, the logistic operations recorded in the Petrobras database, BDEMQ, have inconsistent values, because BDEMQ use the tanks' gap to do the calculation. The example shown in Figure 1 shows the operations on the August 10th of 2006 which

involve the tank 4728 of Paulinia's Refinery. Note that the times of reception and transmission line, so the tank has simultaneously operations.

Órgão: 270 Período: 10/08/06 a 10/08/06

Para ver detalhes, clique na referência desejada.

Referência		E/R	Orig./Dest.		Viagem	Produto		Início	Término	Estoq. Op.	Movimento	M/T
			Tipo	Identificador		Mov.	Estq.	Data - Hora	Data - Hora	Ini. Amb. (m3)	Vol. Amb. (m3)	
◉	TQ 4728	R	UP	200		25P	25P	03/08/06 11:05		12.851	30.142	D
○	TQ 4728	R	UP	200A		25P	25P	05/08/06 17:32		9.257	15.257	D
○	TQ 4728	E	UP	MLD		64X	25P	09/08/06 06:36	10/08/06 13:16	12.909	5.725	D
○	TQ 4728	E	UP	MLD		64X	25P	10/08/06 13:30	10/08/06 19:10	13.814	912	D
○	TQ 4728	E	UP	MLD		64X	25P	10/08/06 19:12	11/08/06 06:08	14.072	1.680	D

1 de 1

bomb.externo | item bomb. | qualidade | comparar quantidades | consultar proprietários | voltar

This makes the estimated values for these operations be inconsistent and often even negative, which complicate the work of teams of accounting and mass balance.
The software developed came to reduce this problem using the measured data of the refining units to make a preliminary estimate of the operations of the entire plant and provide this estimate to the accounting team and for the mass balance.

3. Software Structure

This software was developed by Soteica's team in Visual Studio 2005 with the Visual Basic and uses a database build in SQL Server.
Data is captured directly from Petrobras' database using an ODBC connection and the connection with the Plant Information System is done through the library of the OSI communication with Visual Basic.

4. Software Algorithm

The Sigmafine's model was based on nodes, which consist in points of closing balance. In these models the nodes have meters of the refining units coupled and they are source or destination of logistical operations.
The software uses data measurement and density of Plant Information System and tank stock that is recorded on BDEMQ to close mass balance in each node.
The record of logistic operations is done by transfer and storage team when they occur.
To perform this procedure the software uses a table called "Nos", which contains the information listed below:

- Process unit that the node refers;
- Node's code;
- Meters related to each node;
- Measure unit of each meter;
- Conversion factor of each node;
- Density of each meter;
- Density measure unit of each meter;
- Level of confidence in each measure, expressed in %;
- Logic of calculation of each node;

For each node, is held the following algorithm for calculation:

- Select logistic operations related to each node;
- Logistic operations divided into periods of time;
- Calculates the amount of movement for a period;
- Calculates the density and the percentage of confidence of each logistic operation;
- Record this operation in the database with the estimated value;

The procedure described above gives an estimative for the balance of each tank, but is not able to close their balance due to measurement inaccuracies.
Once chosen a tank and the operations involved with the tank, with the estimated data and the tank's gap recorded on BDEMQ.
With details of operations and tank's gap, the difference between the calculated sum of operations related to the tank and the tank's gap, which is the objective function to minimize. So, it is wanted that the difference reaches the value zero.
In the minimization is used a modification of Newton's Method, and the search is performed within a space defined by the percentage of confidence of logistic operations. That is, the estimated value of each logistic operation is the center of search space and its limits is the estimated value less the value of percentage of confidence and the estimated value plus the value of percentage of confidence, as it can be seen in equation (1).

Search space = (estimated value – value of percentage of confidence … estimated value + value of percentage of confidence) (1)

The construction of search was based on the Theory of Constraints introduced by Goldratt (1990).
To search the space is used a modification of Newton's Method, where the function to minimize is given by equation (2).

Objective Function = ∑ input logistic operations - ∑ output logistic operations – tank's gap (2)

The modification of the Newton's Method is on the derivate, which now is given by equation (3).

Value of Logistic Operation = Value of Logistic Operation – (Function / (Value of Logistic Operation * Value of Logistic Operation Estimated)) (3)

The minimization procedure is iterative and stops when the value zero of function is reached or until 100.000 iterations are reached.
If the value of zero the function will not be achieved with 100.000 iterations, it means that the algorithm not converged, and then it will begin the process of distribution of differences.
The distribution of differences procedure is the analysis of the logistic operations involved with the tank and the identification of which has a greater value of percentage of confidence. The logistic operation with the biggest value of percentage of confidence receives the difference.

The procedure for closing the balance of the tank is iterative and it is repeated for each tank that operated on the daily analysis.

5. Results

This software is operating in Paulinia's Refinery since November 2008 and it has been showing good results as it can be seen below in the reports presented in Figures 2 and 3. These reports are submitted to the accounting team in the first hours of the morning and sets values of the tanks in which mass balance is closed.

Relatório de Movimentações - Área B

Data Início	Data Fim	Origem	Destino	Densidade	BDEMQ (m3)	Quantidade (m3)	Quantidade (te)	Diferença (m3)
TQ-4211								
22-09-2008	22-09-2008 23:59:59	U200ANL	TQ-4211	0.76	2,664.02	**891.92**	680.63	-1,772.10
22-09-2008	22-09-2008 23:59:59	U260TPO	TQ-4211	0.71	4,341.13	**1,627.64**	1,148.47	-2,713.49
22-09-2008	22-09-2008 23:59:59	U200NL	TQ-4211	0.71	-2,530.76	**2,413.90**	1,703.25	4,944.66
TQ-4232								
22-09-2008	22-09-2008 23:59:59	U980NLP	TQ-4232	0.74	-764.77	**874.80**	647.36	1,639.57
22-09-2008	22-09-2008 23:59:59	TQ-4232	U220NFT	0.74	-136.33	**1,624.40**	1,202.06	1,760.73
22-09-2008	22-09-2008 23:59:59	U980ANLP	TQ-4232	0.74	766.05	**1,020.86**	755.44	254.81
TQ-4626								
22-09-2008	22-09-2008 23:59:59	TQ-4626	OP10	0.72	4,920.76	**4,920.76**	3,560.66	0.00
TQ-4627								
22-09-2008	22-09-2008 23:59:59	U220NC	TQ-4627	0.73	8,981.93	**4,452.26**	3,228.33	-4,529.67
22-09-2008	22-09-2008 23:59:59	U220ANC	TQ-4627	0.73	-49.59	**4,479.57**	3,248.09	4,529.16
TQ-4628								
22-09-2008 00:02:00	22-09-2008 19:05:00	TQ-4628	EMEDFT02	0.73	1,701.24	**1,701.24**	1,237.31	0.00
TQ-4634								
22-09-2008 01:18:00	22-09-2008 22:03:00	TQ-4211	TQ-4634	0.72	9,284.82	**4,810.43**	3,473.61	-4,474.39
TQ-4703								
22-09-2008 11:07:00	22-09-2008 23:59:59	TQ-4703	EMEDQAV	0.81	3,246.61	**3,246.61**	2,620.67	0.00
TQ-4704								
22-09-2008	22-09-2008 22:34:00	U340QT	TQ-4704	0.81	1,089.97	**1,089.97**	882.11	0.00
TQ-6305								
22-09-2008	22-09-2008 23:59:59	TQ-6305	U631QITG	0.81	-0.07	**-0.07**	-0.06	0.00

Data de Impressão: 23-09-2008 **Página:** 2

Figure 2: Reports gasoline and LPG accounting area

Relatório de Movimentações - Área A

Data Início	Data Fim	Origem	Destino	Densidade	BDEMQ (m3)	Quantidade (m3)	Quantidade (te)	Diferença (m3)
TQ-4311								
23-09-2008	23-09-2008 23:59:59	U220OLR	TQ-4311	0.97	0.00	**0.00**	0.00	0.00
23-09-2008	23-09-2008 23:59:59	TQ-4311	U210DIL	0.97	-262.49	**280.53**	271.83	543.02
23-09-2008	23-09-2008 23:59:59	U980ADIL	TQ-4311	0.97	164.69	**0.00**	0.00	-164.69
23-09-2008	23-09-2008 23:59:59	U220AOLR	TQ-4311	0.97	-95.46	**149.16**	144.54	244.62
23-09-2008	23-09-2008 23:59:59	U980DIL	TQ-4311	0.97	-164.27	**0.00**	0.00	164.27
TQ-4403								
23-09-2008	23-09-2008 23:59:59	U980GOCR	TQ-4403	0.94	-783.77	**526.78**	496.44	1,310.55
23-09-2008	23-09-2008 23:59:59	TQ-4403	U220INJ	0.94	467.02	**1,939.29**	1,827.58	1,472.27
23-09-2008	23-09-2008 23:59:59	U980AGOC	TQ-4403	0.94	788.06	**549.80**	518.13	-238.26
TQ-4405								
23-09-2008	23-09-2008 23:59:59	TQ-4405	U220ACRG	0.94	-7,913.43	**8,186.13**	7,699.87	16,099.56
23-09-2008	23-09-2008 23:59:59	U210GO	TQ-4405	0.94	-7,900.79	**6,875.02**	6,466.64	14,775.81
TQ-4407								
23-09-2008	23-09-2008 23:59:59	TQ-4407	TQ-4405	0.94	2,357.50	**2,357.50**	2,222.18	0.00
TQ-4408								
23-09-2008	23-09-2008 23:59:59	U210AGO	TQ-4408	0.94	2,764.09	**6,401.07**	6,020.20	3,636.98
23-09-2008	23-09-2008 23:59:59	TQ-4408	U220CRG	0.94	9.60	**3,984.05**	3,747.00	3,974.45
TQ-4410								
23-09-2008	23-09-2008 23:59:59	TQ-4410	U980REC	1.00	28.29	**32.67**	32.67	4.38
23-09-2008	23-09-2008 23:59:59	TQ-4410	U980AREC	1.00	16.12	**61.06**	61.06	44.94
TQ-4411								
23-09-2008	23-09-2008 23:59:59	U980RSUN	TQ-4411	0.92	1.78	**1.78**	1.65	0.00
TQ-4412								
23-09-2008	23-09-2008 23:59:59	TQ-4412	U980ARV	1.02	-2,820.49	**2,849.64**	2,906.63	5,670.13
23-09-2008	23-09-2008 23:59:59	U210OC	TQ-4412	1.02	-108.48	**1,634.32**	1,667.01	1,742.80

Data de Impressão: 24-09-2008 **Página:** 1

Figure 3: Reports of fuel oil accounting area

It is important to notice in these reports that the value estimated by BDEMQ to the tanks that operate simultaneously most often is inconsistent, even showing negative values, as it can be seen on column BDEMQ (m^3), in the reports demonstrated in Figures 2 and 3. What generates an enormous amount of work for the accounting team that lost a long time to make estimates with less precision.
These data also feed data reconciliation system, Sigmafine, and improve quality of initial data, reducing gross errors and time spent in the reconciliation procedure. You can also notice an improvement in quality scores of reconciliation.
One interesting thing that can be noted with the use of these data is that when for some reason the process data are not available or have very huge errors, the algorithm presents values that make clear that there are problems in the area.
As the algorithm works with each tank individually, the errors are confined to the tank and not spread in the model as a whole.

6. Conclusion

This work shows that the algorithm developed and implemented in the software was efficient and able to prepare the data for the data reconciliation system and for the accounting team, that has now consistent estimates for operations simultaneously.
Due to good results, this work should be replicated in others refining units of Petrobras by Soteica. The next to receive it will be Landulpho Alves Refinery, in Bahia.

References

Bomfim, C. H. M.. Reconciliação de dados: porque e quando precisamos. Documento de circulaçao interna da Petrobras, Refinaria Gabriel Passos - REGAP. 2005.
Goldratt, E. M., Cox, J.. What is this thing called Theory of Constraints and how should it be implemented. North River Press, 161. 1990.
Placido, J., Loureiro, L. V.. Industrial application of data reconciliation. Computers and Chemical Engineering, v. 22, p 1035-1038. 1998.
Weiss, G. H, Romagnoli, J. A., Islam, K. A.. Data reconciliation – An industrial case study. Computers and Chemical Engineering, v. 20, p 1441-1449. 1996.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Genetic Algorithm for Tailored Production of Polymer Resins

Karen V. Pontes[a], Marcelo Embiruçu[a], Rubens Maciel[b]

[a]*Industrial Engineering Programme – Bahia Federal University (UFBA), Rua Prof. Aristides Novis, 2, Federação, Salvador – BA – Brasil.*
[b]*LOPCA/DPQ - Department of Chemical Engineering – Campinas State University (Unicamp), Cidade Universitária Zeferino Vaz, Campinas – SP – Brasil*

Abstract

This study presents an optimization model based on genetic algorithm (GA) for the production of target polymer resins. Unlike most contributions in open literature, this study takes into account an economic optimization criterion while satisfying the desired polymer properties through constraints at the end of the reactor. The case study is the ethylene coordination polymerization, which takes place in a CSTR (continuous stirred tank reactor). Due to the high non-linearity of the process model, a stochastic optimization method based on GA is used. Several simulations were carried out and the results show that the GA is able to determine the optimal operating conditions satisfactorily. Binary codification proved to be more robust than real codification.

Keywords: Genetic algorithm, optimization, polymerization

1. Introduction

In last years much attention has been given to the design of tailored polymer resins through optimization models. Computer-based tools constitute an alternative to the expensive pilot plant or industrial scale experiments. The majority of contributions focuses on batch polymerization and then minimizes the batch time while specifying the polymer properties through constraints of the optimization (Sundaram et al., 2005). Another very common approach is the minimization of a single scalar objective function which corresponds to the quadratic deviation of the observed and desired polymer properties values, i.e., a sort of multiobjective optimization (Hanai et al., 2003; Lemoine-Nava, 2006). This approach however, requires the definition of weighting factors, what may not be trivial since some properties in the objective function can present conflicting behaviors.

Genetic algorithm (GA) is a powerful alternative for complex and highly non linear problems, such as the optimization of polymerization reactors. GA does not require the computation of derivatives nor an initial estimate to find an optimum, but it can still be improved if some problem knowledge is provided to the initial population (Pontes, 2008). Because of these characteristics, GA has gained acceptance in almost all areas of science and engineering, and the polymer reaction engineering is not an exception (Kasat et al., 2003). Many efforts have gone on the multiobjective optimization, which account for the polymer properties and, in some studies, for the batch time, obtaining Pareto sets (Agrawal et al., 2006) through a non-dominated sorting genetic algorithm. However, when using such approach, the decision of the best operational condition still requires process knowledge and/or intuition.

Bearing in mind the mentioned drawbacks of multiobjective optimization and that several operational conditions may yield the same polymer grade, a more robust approach should account for operational profit as objective, while satisfying the desired polymer properties through constraints of the optimization. Within this scope, the present study proposes an optimization problem for targeting polyethylene resins by means of Genetic Algorithm. An industrial scale ethylene polymerization in solution with Ziegler-Natta catalyst is taken as case study due to the scarceness on modeling and optimization of such systems in open literature. To the authors' knowledge, solely Brandolin et al. (1991), Asteasuain et al. (2001) and Agrawal et al. (2006) address the optimal operating policies problem to the ethylene polymerization. However, these works consider processes taking place at high pressure plug flow reactors through a free radical mechanism, i.e., for the production of low density polyethylene (LDPE).

The paper is structured as follows. Firstly, the process is described. In the next section the optimization problem formulation using genetic algorithm is discussed. Then, some results are shown in order to illustrate the potential of the proposed approach. Finally, some conclusions are presented.

2. Optimization Problem Formulation

The polymerization process takes place in a series of tubular and stirred reactors that can be combined in different flowsheets in order to produce a wide range of polymer grades. The focus of this study is the stirred configuration, which is illustrated in Figure 1. Monomer, hydrogen and a catalyst mixture, dissolved in a solvent, are fed to the non-ideal CSTR (continuous stirred tank reactor). The main feed is split between the top and the bottom of the reactor, enhancing the mixing effect of the agitator. The reaction is almost completed when entering the PFR (plug flow reactor).

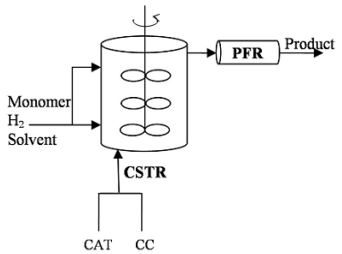


Figure 1: Typical reactor configuration: tubular (left) and stirred (right).

Pontes et al. (2008a) determined the most important variables regarding their effects on the objective function and on the polymer properties by means of a design of experiments approach. The degrees of freedom of the optimization were then determined: monomer (M), hydrogen (H_0) and catalyst (CAT) inlet concentrations, total (W_t) and side feed (W_s) flow rates and inlet pressure (P_{in}).

The polymer quality is verified through melt index (MI), stress exponent (SE) and density at the outlet reactor, i.e., PFR_b. The MI is a measure of the polymer average molecular weight: the higher the MI, the lower is the molecular weight. On the other hand, SE measures the polydispersity of the molecular weight distribution (MWD): a more heterogeneous polymer corresponds to a higher SE value. The density is greatly influenced by the comonomer content, but the present study focuses on the

homopolymerization process. Therefore, mainly MI and SE are used to specify the polymer quality. As mentioned previously, the polymer properties are attained through constraints at the outlet of the PFR. In addition to that, operational constraints on outlet variables may be accounted for in order to ensure safe operation. For example, the outlet temperature (T_{out}) and outlet pressure (P_{out}) may be constrained in order to avoid phase separation. Hence, a vector with constrains (equality or inequality) at the reactor outlet is set:

$$\mathbf{h} = \begin{bmatrix} MI & SE & T_{out} & P_{out} \end{bmatrix}^T \tag{1}$$

This reactor was modeled in previous studies (Embiruçu et al., 2000; Pontes et al. 2008b) and the parameters were exhaustively validated with actual plant data (Embiruçu et al. 2008). The numerical strategy was to use the previously developed dynamic model, taking the stationary state as response to be evaluated by the optimization model. Such approach is justified because of the difficulties encountered when initializing a set of highly non linear algebraic equations for the non-ideal CSTR. Considering the index $i = \{s, t\}$, where s indicates the CSTR and t the PFR, the optimization problem can be formulated according to:

$$\max_{\mathbf{u}} \Phi \tag{2}$$

s.t.

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{p}, z), \quad \left\{ i = t : \; z \in [z_0 \; z_f], \; i = t : \; z = 0 \right\} \tag{2a}$$

$$0 = \mathbf{g}(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mathbf{p}), \tag{2b}$$

$$0 = \mathbf{x}(z_0) - \mathbf{x}_0, \tag{2c}$$

$$\mathbf{u}_{LB} \le \mathbf{u} \le \mathbf{u}_i, \quad z = z_0, \tag{2d}$$

$$\mathbf{h}_{LB} \le \mathbf{h} \le \mathbf{h}_i, \quad z = z_f, \tag{2e}$$

where Φ is profit (difference between incomes with polyethylene sales and costs with raw materials), $\mathbf{x}$ are differential state variables, $\mathbf{y}$ the algebraic state variables, $\mathbf{u}$ the decision variables, $\mathbf{p}$ the invariant parameters, z the axial coordinate, the subscripts 0 and f refer to the reactor inlet and outlet, $\mathbf{u}$ and $\mathbf{h}$ are vectors with decision variables and constraints and the sub indexes LB and UB are lower and upper bounds respectively. The initial conditions of equation (2a) are the inlet conditions to the CSTR; they are given by equation (2c) or by the decision variables. The algebraic equation (2) comprehends additionally some output process variables such as conversion and polymer production rate as well as correlations for the polymer properties.

The optimization is carried out by the genetic algorithm (GA), a stochastic method that mimics the mechanism of natural selection, i.e., the fittest individuals survive in a competitive environment. The probability of an individual to survive is given by its fitness, which is related to the objective function. Each individual or chromosome has to be coded, e.g. into binary or real representation. If binary representation is used, the chromosome is composed of the so-called bits, which can assume only 0 or 1 values. This representation attempts to represent real variables with a limited number of possibilities. The variable precision is then given by the chromosome length. Real representation overcomes this by directly using the real values of the variable. Basically, three operators are responsible for the GA search potential: reproduction, crossover and

mutation. The selection of the best individuals is based on the tournament selection, which randomly selects two individuals, compares them, and selects the one with the best fitness to the next generation. When handling constraints, the potential of the tournament selection is explored (Deb, 2000): (a) when two feasible solutions are compared, the one with the best objective function is selected; (b) when a feasible solution is compared with an infeasible solution, the former is selected; (c) when two infeasible solutions are compared, the one that less violates the constraint is adopted.

In order to use the GA a number of parameters have to be set. In previous work (Pontes, 2008) a systematic procedure based on design of experiments is used to ascertain the optimum set of parameters: seed number for generation of random initial population (0.4), mutation (0.05) and crossover (0.7), population size (100) and chromosome length for binary codification (8).

3. Results and Discussion

The optimization model was implemented in Fortran code and several optimization studies were carried out in order to investigate the performance of the genetic algorithm using both binary and real codifications when targeting polyethylene resins.

According to the optimization model, GA must maximize profit in order to produce polymer resins with pre-specified properties, namely melt index (MI) and stress exponent (SE). A sequence of optimizations was carried out, and the respective MI and SE constraint limits are given in Figure 2, together with the optimum profit for each SE specification. The optimal operational conditions for the inlet concentrations of reactants are illustrated in Figure 3. Results are normalized due to confidentiality reasons.

GA shows an overall good performance at identifying optimum operating conditions for designing target polymer resins. GA with binary codification performs better than real GA. When comparing both results, it is worth observing that the real GA reproduce the random characteristic of the method since it finds most of the time just a local optimum. These local optima may be associated to the particular crossover and mutation operators used for real codification (Deb, 2000). A deeper investigation on their parameters and on other operators might improve real GA performance.

	SE_{LB}	SE_{UB}
A	0.250	0.260
B	0.260	0.270
C	0.270	0.280
D	0.280	0.290
E	0.290	0.300
F	0.300	0.310

0.282 = MI = 0.294

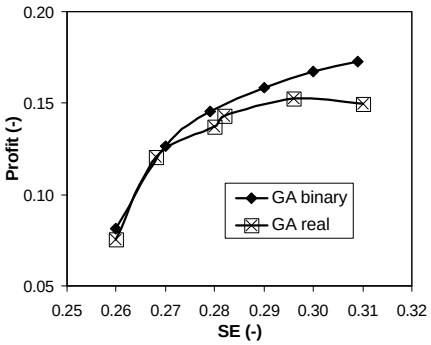


Figure 2: Profit versus SE.

Catalyst inlet concentration and side feed show to be the most important variables to control the stress exponent. A tradeoff between both variables on the polymer properties allows the desired polymer specification while a tradeoff between catalyst and monomer inlet concentrations is responsible for the increasing profit. For more details on the phenomena taking place inside the reactor the reader should address previous works (Pontes et al., 2008b; Pontes, 2008).

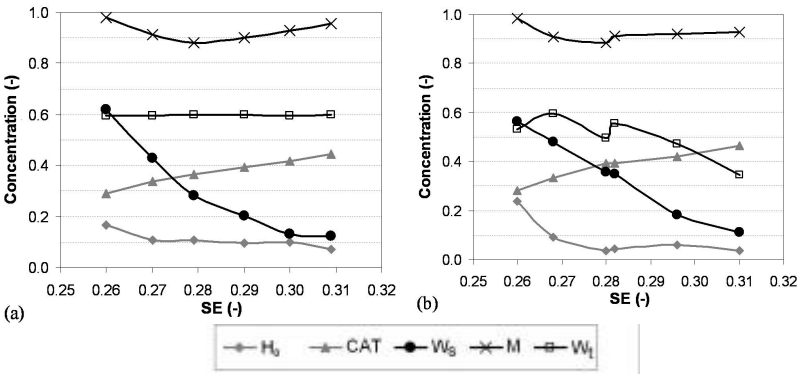


Figure 3: Inlet concentrations *versus* SE: (a) GA binary; (b) GA real.

Attempting to further investigate and compare the performances of binary and real codifications, the evolution of objective and penalty functions values over the 100 generations for example F above are illustrated in Figure 4. Real codification finds the first feasible solution around the 50th generation, whereas binary GA around the 10th generation. The decrease in objective function in the first generations is associated to the constraint handling method: unfeasible solutions with lower penalty functions are selected regardless their objective function values.

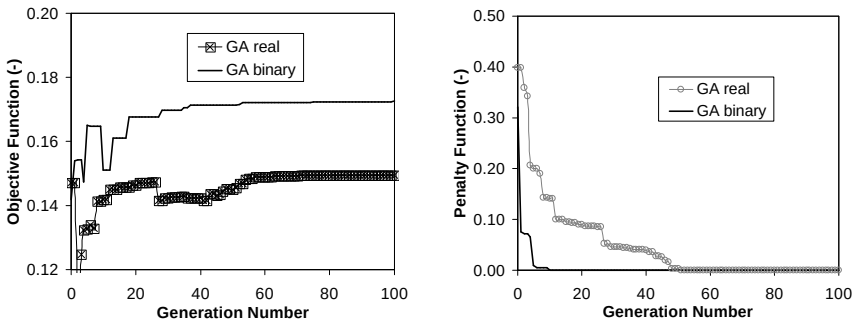


Figure 4: Comparison between binary and real codifications for example F: objective (left) and penalty (right) functions over generations.

Unlike SQP based optimization methods, the GA does not require an initial guess to find an optimum since the initial population is randomly created. However, adding valuable information to the GA initial population may guide the search for the global optimum. In this scope, a hybridization method is considered here, which incorporates into the GA initial population an initial guess based on process knowledge (GA-I). Example F is again taken as case study. No significant improvement is observed with

binary codification but with real GA, on the other hand, this approach led to an objective function 13.22% higher (Figure 5). The improvement is associated with the first feasible individual detected around generation 20th, much earlier than pure GA (Figure 5, right). However, the new optimum still is a local optimum compared to binary GA: the latter is 1.87% higher than the former (GA-I real).

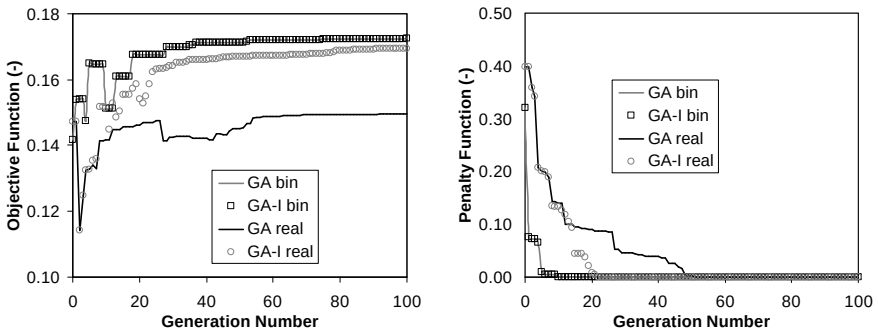


Figure 5: Hybrid GA: objective (left) and penalty (right) functions over generations.

4. Conclusions

The paper presents the application of genetic algorithm for targeting polyethylene grades. Economic objectives are considered, while satisfying the quality requirements through constraints. The method can satisfactorily predict the optimal operating conditions in order to produce a polymer with desired quality. The results illustrate that the performance of the method can be highly influenced by the codification, i.e., real or binary. The binary GA presented a more robust performance than the real one, which fails to locate the global optimum for some optimizations since a better optimum is found when using the binary codification.

It is widely known that the GA does not require an initial guess to find an optimum, but valuable process knowledge can guide the search towards the global optimum, which was observed by the hybrid approach. When designing similar polymer grades or when optimizing current industrial operating conditions, the hybrid approach can therefore be particularly useful in the search for a global optimum.

References

N. Agrawal, GP Rangaiah, AK Ray, SK Gupta, 2006, Ind. Eng. Chem. Res. v. 45, p. 3182-3199.
M Asteasuain, PE Ugrin, MH Lacunza, A Brandolin, 2001, Polym. React. Eng. v. 9, p. 163-182.
A Brandolin, EM Valles, JN Farber, 1991, Polym. Eng. Sci. v. 31, p. 381-390.
K Deb, 2000, Comput. Methods Appl. Mech. Engrg., v. 186, p. 311-338.
M Embiruçu, EL Lima, JC Pinto, 2000, J. App. Pol. Sci., v. 77, p. 1574-1590.
M Embiruçu, DM Prata, EL Lima, JC Pinto, 2008, Macrom. Reac. Eng., v. 2, p.142-160.
T Hanai, T Ohki, H Honda, T Kobayashi, 2003, Comp. Chem. Eng., v. 27, p. 1011-1019.
RB Kasat, AK Ray, SK Gupta, 2003, v. 18, p. 523-532.
R Lemoine-Nava, A Flores-Tlacuahauc, 2006, Ind. Eng. Chem. Res. v. 45, p. 4637-4652.
KV Pontes, 2008, PhD Thesis, Universidade Estadual de Campinas, São Paulo.
KV Pontes, R Maciel, M Embiruçu, 2008a, In: PPS-23, Salvador, Brazil.
KV Pontes, R Maciel, M Embiruçu, A Hartwich, W Marquardt, 2008b, AICHE J., v. 54, p. 2346.
BS Sundaram, SR Upreti, A Lohi, 2005, Macrom. Theory Sim. v. 14, p. 374-386.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Mathematical modeling and optimal operation of industrial tubular reactor for naphtha cracking

G.-Y. Gao[a], M. Wang[b], C. C. Pantelides[c], X.-G. Li[a], H. Yeung[b]

[a]*School of Chemical Engineering, TianJin University, 300072, P.R. China*
[b]*Process Systems Engineering Group, School of Engineering, Cranfield University, MK43 0AL, UK*
[c]*Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, SW7 2AZ, UK*

Abstract

The tubular reactor in a naphtha cracking furnace is modelled rigorously in this paper. The mathematical model can be used to predict product yields, coking buildup inside the tube wall, run length (i.e. the time between two consecutive decoking operations), residence time and pressure drop. A powerful modelling, simulation and optimisation tool gPROMS was chosen to implement the proposed work. This model provides detailed understanding of the naphtha cracking process. Steady-state optimisation was then applied to the operation of this industrial tubular reactor. The operating profit is maximised when the process gas temperature profile along the reactor and the inlet steam to naphtha ratio vary within certain ranges. The effects of coking on heat transfer, on reduction of manufacturing time and the decoking cost have been considered in the optimisation. Process simulation and optimisation based on this detailed model will give process engineers in the ethylene industry some insights on ethylene furnace design and operation.

Keywords: Mathematical modeling, tubular reactor, ethylene furnace, case study

1. Introduction

Ethylene is the one of most important building blocks used in the petrochemical industry. Thermal cracking furnace is the heart of the whole ethylene manufacturing process. A typical unit generally includes the convection section, the cross-over section and the radiation section (Sundaram *et al.* 2001). Naphtha is preheated in the convection section and then mixed with steam. The mixture of naphtha and steam is introduced into the tube reactors in the radiation section. The cracking reactions take place inside these long tubular reactors. Steam is a diluent and it is used to improve olefin selectivity and reduce coking rate.

Naphtha cracking process involves many free-radical reactions. Due to computational difficulties, models based on molecular reactions have been widely used (Sundaram *et al.* 2001). Kumar and Kunzru (1985**a**) proposed a molecular reaction mechanism for naphtha steam cracking based on their experimental research. In the meantime, Kumar and Kunzru (1985**b**) developed a kinetics model for predicting coking rate.

Shahrokhi and Nejati (2002) carried out an optimal operation study based on one-dimensional (1D) steady-state model for propane thermal cracking. Operating profit was used as objective function and the optimal temperature profile was found. Decoking

cost was counted in this study. Hu *et al.* (2004) modelled and simulated an industrial tubular reactor for naphtha cracking. This mathematic model was developed based on the molecular reaction scheme proposed by Kumar and Kunzru (1985**a**) and the coking reaction scheme proposed in Kumar and Kunzru (1985**b**). Masoumi *et al.* (2006) developed a 1D steady-state model for tubular reactors in naphtha cracking. A free-radical reaction scheme including 90 species and 543 reactions was used. An optimisation study was performed with the aim to maximize the operating profit. Han *et al.* (2006) reported a study of numerical simulation on fuel combustion and naphtha pyrolysis reactions in another naphtha cracking furnace.

Mathematical model for naphtha thermal cracking in ethylene manufacturing was developed in this paper. A real industrial thermal cracking furnace presented in Hu *et al.* (2004) was selected as benchmark. The modified molecular reaction scheme was used. The novel contributions of this paper include (a) the modeling of coking buildup and the evaluation of its impact; (b) the accounting of the effects of coking on heat transfer and reduction of manufacturing time, and the decoking cost in the optimal operation case study.

2. Mathematic modeling

The tube reactor for naphtha cracking was modeled under the following assumptions:

- One dimensional flow (i.e. negligible radial concentration gradients)
- Plug flow for process gas & ideal gas behaviour

Due to the fast dynamics of the process gas inside the tube, the mass and energy balances were expressed in steady-state form. The generic equations for modelling a typical plug flow reactor can be referred to Pantelides (2003). The coke layer was modelled specifically. The coking buildup inside the tube was modeled in pseudo steady-state (updated hourly).

3. Model validation

The simulation is based on the industrial cracking tube reactors described in Hu *et al.* (2004). The reactor parameters and operation conditions are listed in Table 1. The molecular reaction scheme proposed by Kumar and Kunzru (1985**a**) was adopted. Due to different naphtha as feedstock, the stoichiometric coefficients for the primary reaction were adjusted (Hu *et al.* 2004).

Table 1 Tube reactor geometries and operating conditions

Reactor tube		**Operation variables**	
Tube length, *m*	24.159	Naphtha feed，*kg/h*	6,666
Diameter of inlet tube, *mm*	Φ72×6.5	Steam/naphtha ratio, *kg/kg*	0.5
Diameter of outlet tube, *mm*	Φ98×7.5	Tube feed temperature, *K*	900
Number of reactor tubes	32	Tube outlet temperature, *K*	1,100
Highest temperature allowed for tube, *K*	1,373	Tube outlet pressure, *kPa*	216

Table 2 Primary reaction for naphtha cracking used in this study

Reaction equations	**K_0, s^{-1}**	**E, cal/mol**
$C_{6.5}H_{14}$ => $0.406H_2 + 0.910CH_4 + 1.06C_2H_4 + 0.176C_2H_6 + 0.504C_3H_6 + 0.001C_3H_8 + 0.001C_4H_{10} + 0.219C_4H_8 + 0.123C_4H_6 + 0.002C_4$'s	6.565E+11	52,580

The model for calculating coking rate proposed by Kumar and Kunzru (1985**b)** can be described as follows:

$$r_c = 1.95 * 10^{15} * EXP(-212,260 / RT) * C_{arom}^{1.97} \qquad (1)$$

where r_c denotes the coking rate in kg/(m^2 h) and C_{arom} denotes the total aromatics concentration in kmol/m^3.

The process gas temperature profile applied on the tube reactor is shown in Figure 1, which is the same as in Hu *et al.* (2004). In Table 3, the main product yields from the base case simulation in gPROMS were compared with the industrial data published in Hu *et al.* (2004). As can be seen, simulation results are generally in good agreement with industrial values. When the tubular reactor is operated under the same conditions as listed in Table 1 and shown in Figure 1, the model predicts that the production time before a decoking operation is 38 days and the coking thickness is 7.73 mm at the coil outlet. This is very close to real life production time 39 days. These two aspects indicate that the developed model can reflect the real naphtha cracking process.

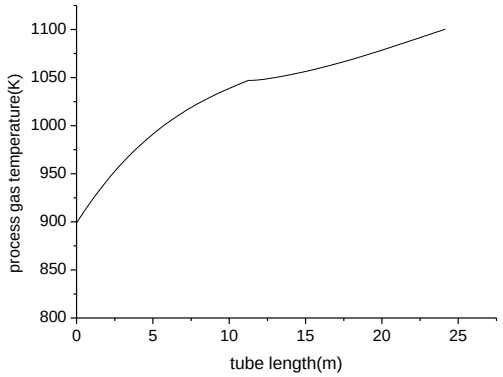


Figure 1 Process gas temperature along the tube reactor

Table 3 Comparison of product yields between simulation and industrial measurements

Components	**Yields, % (industry)**	**Yields, % (simulation)**
C_2H_4	29.16	30.36
C_3H_6	14.94	14.86
CH_4	14.94	15.28
H_2	0.96	0.97
C_2H_6	3.80	3.66
C_4H_6	5.14	4.45
C_4H_8	5.08	4.86
Process gas outlet temperature, K	1,109	1,103
Process gas outlet pressure, kPa	171	170.3

4. Naphtha cracking process analysis

This section is to analyse the naphtha cracking in the whole manufacturing cycle.

4.1. Tube outlet wall temperature

Figure 2 indicates the change of the tube external wall temperature along tube length with production time. It can be seen that the highest temperature is at about the middle of whole tube reactor at the beginning of production. With the production time going on, outer tube temperature is growing gradually and the position of highest temperature moves to the outlet of tube. When the outer tube temperature reaches 1373K which is the allowed highest temperature, the cracking furnace has to be shut down for decoking. From Figure 2, the run length is 38 days, which matches the industrial data 39 days very well.

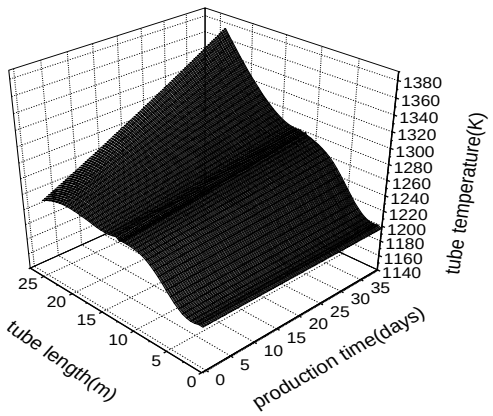
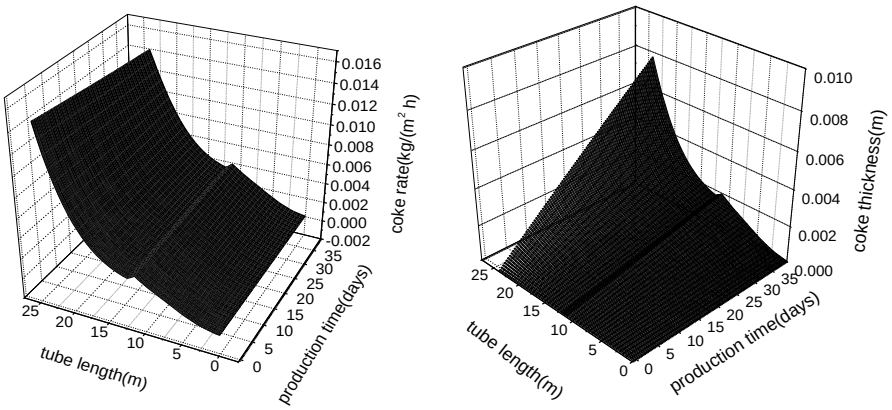


Figure 2 Outer tube temperature changes along tube length with production time

4.2. Coking rate and coking buildup



(a) Coking rate along tube reactor length with production time

(b) Coke thickness growth along tube with production time

Figure 3 Coking process inside reactor tube

Figure 3 shows coking rate and coke thickness vary with tube length and production time. Form Equation (1), it can be seen that coking rate depends upon both process gas

temperature and aromatics concentration in process gas. Although aromatics concentration drops along tube reactor, coking rate increase along the two pass of the tube reactor due to process gas temperature going up. But there was a decrease of coking rate along the joint bend tube due to aromatics concentration drop sharply. Figure 3(b) shows coke thickness growth with tube length and production time (days). The reactor outlet has thickest coke built-up obviously at the end of the production cycle.

5. Optimal operation case study

The following assumptions have been made for formulating this optimisation problem: (a) No downstream product separation costs are counted; (b) The tubular reactor operates in steady-state.

5.1. Mathematical formulation

The operating profit was used as the objective function. It was defined as the income from desired products minus various costs. Costs include raw material cost (naphtha here), cost for steam, cost for radiant heat plus the decoking cost. A fixed value (denoted by DCC) was used for decoking cost (per time). The objective function was calculated on a yearly basis.

$$f_p = Income - Cost \tag{2}$$

$$Income = (t_p - n_d t_d) * \sum F_i c_i \tag{3}$$

$$Cost = (t_p - n_d t_d) * (F_0 c_0 + F_H c_H + Q c_Q) + n_d * DCC \tag{4}$$

Where f_p represents the operating profit, F_i the desirable product flow rate, F_0 the naphtha flow rate; F_H the steam flow rate; n_d denotes the decoking times per year. The meaning for other parameters can be referred to Table 4.

The decoking frequency is calculated via Equation (5).

$$n_d = \frac{t_p}{t_d + t_e} \tag{5}$$

Where t_e denotes the production time between consecutive decking processes (unit in hours). In this study, t_e is predicted by the mathematical model with the criteria that the outlet metal temperature is greater than and equal to 1373 K.

Table 4 gave all the parameters assumed for this case study. Most of the price factors were taken from *ICIS Chemical Business* (2007).

Table 4 Parameters for optimisation function

Physical meaning	Parameters	Values
The yearly production time	t_p, hour	8,160
decoking time	t_d, hour	48
ethylene price factor	C_1, \$/t	1,200
Propylene price factor	C_2, \$/t	1,180
naphtha price factor	C_0, \$/t	541
Steam price factor	c_H, \$/kg	0.0129

Heat price factor	c_Q, *$/kJ*	1.26E-05
Decoking cost	*DCC* ($/time)	66,600

5.2. Results and discussions

After performing steady-state optimisation in gPROMS, optimisation results were summarised and then compared with the operating conditions described in Hu *et al.* (2004) in Table 5. From Table 5, it can be seen that the operating profit was improved obviously after optimisation.

Table 5 Steady-state optimisation results

Items	Original value	Optimal value
Coil outlet temperature (K)	1,100	1,118
Steam to naphtha (kg/kg) ratio	0.500	0.670
C_2H_4 yield (wt%)	0.315	0.341
C_3H_6 yield (wt%)	0.148	0.138
heat flow(*kJ/h)*	624,858.2	759,304.0
Objective function, $/year	70,054	85,291
Production time(days)	39	27

6. Conclusions

In this paper, mathematical model for tubular reactors in the industrial thermal cracking furnace for ethylene manufacturing was developed first. The coking layer inside the tube wall has been modelled, the coking impact on heat transfer and pressure drop has been considered. Process analysis based on this model provides detailed understanding of the naphtha cracking process. Steady-state optimisation was applied to the operation of this industrial tube reactor for maximum operating profit. The decoking cost has been considered. This study will provide some insights to process engineers in the ethylene industry on ethylene furnace design and operation.

References

Han Y.-L., Xiao, R. and Zhang, M.-Y. (2006), Combustion and pyrolysis reactions in a naphtha cracking furnace, *Chemical Engineering Technology*, Vol. 29, p112-120.

Hu**,** Y.-F, Xu Y.-M, He X.-O. (2004), Modeling and simulation of naphtha pyrolysis process, *Acta Petrolei Sinica (Petroleum Processing Section)*, Vol. 20, No. 6, p51-57.

ICIS Chemical Business (2007), Vol. 2, Issue 48, p41, 08 Jan. 2007.

Kumar, P. and Kunzru, D. (1985**a**), Modelling of Naphtha Pyrolysis, *Ind. Eng. Chem. Process Des. Dev.*, Vol. 24, p774-782.

Kumar, P. and Kunzru, D. (1985**b**), Kinetics of coke deposition in naphtha pyrolysis, *Canadian Journal of Chemical Engineering*, Vol. 63, No. 4, p598-604.

Masoumi, M.E., Sadrameli, S.M., Towfighi, J. and Niaei. A. (2006), Simulation, optimisation and control of a thermal cracking furnace, *Energy*, Vol. 31, p516–527.

Pantelides, C.C. (2003), *The mathematical modelling of the dynamic behaviour of process systems*, Centre for Process Systems Engineering, Imperial College London.

Shahrokhi, M. and Nejati, A. (2002), Optimal Temperature Control of a Propane Thermal Cracking Reactor. *Ind. Eng. Chem. Res.*, *41,* p6572-6578.

Sundaram, K.M., Shreehan, M.M. and Olszewski, E.F. (2001), *Ethylene*, in *the Kirk-Othmer Encyclopedia of Chemical Technology* (5th Ed), John Wiley & Sons, Inc.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling of ammonia removal in RBCs: an industrial case

Marco A. Müller[a], Luís G. S. Longhi[b], Argimiro R. Secchi[c] and Fabio Brião[d]

[a]*GIMSCOP, Department of Chemical Engineering - UFRGS, Rua Eng. Luiz Englert, s/nº, CEP 90040-040, Porto Alegre-RS, Brazil. E-mail: marcoa.muller@yahoo.com.br*
[b,d]*Alberto Pasqualini REFAP S.A, Av. Getúlio Vargas, 11001, CEP 92420-221, Canoas-RS, Brazil. E-mail: {longhi, fabio.briao}@petrobras.com.br*
[c]*COPPE/PEQ/LMSCP - UFRJ, Av. Horácio Macedo, 2030 - Centro de Tecnologia. CEP 21941-914, Rio de Janeiro - RJ, Brazil. E-mail: arge@peq.coppe.ufrj.br*

Abstract

This paper describes the mathematical modeling of a set of rotating biological contactors (RBC) used for ammonia removal by nitrification of the wastewater stream that comes from a petroleum refinery. The entire biological wastewater treatment unit is composed by an aerated facultative lagoon, four sets of RBCs, and a pair of clarification decanters. In this paper we are concerned with the nitrification step. It has two sequential biochemical pathways (nitritation and nitratation) for the nitrification process and the existence of two distinct biological agents, related to *Nitrobacter* and *Nitrosomonas* autotrophic bacteria. We discuss some characteristics about the kinetics and important points that should be concerned for such a study. We also propose an alternative way to model the spatial distribution on the rotating bioreactor, with necessary simplifications and the discussion on their limitations. Available data sets from on-line analyzers for ammonium, composed by 3-hours sampled input and output, were used for parameter estimation and model validation. The process modeling, dynamic simulation, and parameter estimation were completely accomplished in the software EMSO. The results show the good prediction capability of the proposed model. It is expected to use the developed model for monitoring the wastewater treatment and for operator training. Meanwhile, the acquired knowledge obtained during the model development has improved the efficiency and the operational procedures of the industrial unit.

Keywords: nitrification, modeling, wastewater, RBC.

1. Introduction

Most of the industrial wastewaters have ammonia to be removed, and the most common way to fulfill this task is by a biological process known as nitrification, on which ammonia is oxidized to nitrite and nitrate, combined with a further denitrification stage. In this last stage, nitrite and nitrate act as electron acceptors for heterotrophic bacteria, under anoxic condition, converting them to nitrogen gas [2].

The nitrification part is usually divided on two most important steps, known as nitritation, on which ammonium is oxidized to nitrite by the autotrophic bacteria of the genus *Nitrobacter*, and nitratation, on which the nitrite formed previously is oxidized to nitrate by the autotrophic bacteria of the genus *Nitrosomonas*. These two steps occur under aerobic condition and are very sensitive to several factors, such as dissolved

oxygen, pH, temperature, presence of inhibitory compounds, and even the concentration of the own nitrification products [4]. Any attempt to analyze a nitrification process should concern all these variables.

These biological processes can be run on several manners, if their conditions are fulfilled. Older processes use suspended biomass that requires higher residence time and their efficiencies will depend on the contact between the biological agents and the nitrogenous compounds, on the mixture conditions, on the quality of the biomass flake formed and so on. Those processes are simple to design, but are generally less efficient than processes with immobilized biomass that grows on the surface of the equipment (or over the filling) and the affluent flows through it. In the later, the biomass reaches much larger concentration of bacteria, comparing to disperse flakes, resulting on a more effective removal of ammonia [1]. One example of such equipments is the Rotating Biological Contactors (RBC), which is composed by several disks of polymeric material, good for attaching the biomass, connected by a central axis, and with very large available surface for biomass growing. These disks spin around the axis by a mechanical system in a frequency between 1 and 2 rpm, and the RBC is usually designed to operate about 40% of the disks submerged in the affluent, providing good aeration for the immobilized biomass. The aeration of the affluent is also due to the turbulence generated by the disks movement.

On the next sections, it will be discussed the kinetics of the nitrification process, and how it should be treated for a mathematical modeling procedure. Further, it will be presented an industrial case and practical results to illustrate the model application.

2. Nitrification kinetics

Nitrification is held by autotrophic bacteria, which have smaller growth rates than heterotrophic bacteria and are more sensitive to adverse conditions. Usually it is the most critical process of the biological treatment in a wastewater unit and, therefore, its mathematical modeling has to be treated carefully. From now on, it is described the procedure to model this process, the parameters meaning, and some relations among them.

2.1. Modeling the growth rate

The traditional equation to describe the relation between the substrate (ammonium, in this case) and the specific growth rate of the microorganism (μ, units of h^{-1}) is the Monod equation, Eq. 1, which has the advantage of simplicity (just two parameters to estimate) but cannot represent the process on high loads of ammonium, because of the phenomenon of auto-inhibition, when the growth rate starts to decline as the concentration of ammonium increases [6]. For concentrations of 50 g $N\text{-}NH^{4+}/m^3$ or higher, alternative equations must be considered. Carrera et al. [1] compared different equations and concluded that, for nitrification with immobilized biomass, the equation that best fits experimental data is the Aiba's model, Equation 2.

$$\mu = \mu_{\max} \frac{S}{k + S} \tag{1}$$

$$\mu = \mu_{\max} \frac{S}{k + S} \exp\left(- S \middle/ k_i \right) \tag{2}$$

In these equations, μ_{max} is the maximum specific growth rate, which stands for the case when there are no limitations of the substrate for the microorganism. It is a characteristic of the microorganism and, according to Ferreira [2], the maximum specific growth rate is larger for the nitrifying bacteria of the type *Nitrobacter* than for the type *Nitrosomonas*. The parameter k is called half-saturation coefficient and

measures the affinity between the microorganism and the substrate (the less, more affinity) and the parameter k_i measures the auto-inhibition relation.
The actual growth rate of the biomass is the product between the specific growth rate and the microorganism concentration. Other influences can be added to this equation to consider the effects of other operating conditions.

2.2. Operating conditions influence
Ferreira [2] presented the relations of nitrification under several operating conditions and their applications are briefly discussed below.

2.2.1. Dissolved oxygen
This is the most important condition to control a nitrification process, due to its high influence on the biochemical pathways and its very low solubility in water. Experimental results and experience show that nitrification efficiency is limited significantly when the dissolved oxygen reaches 2 mg/L, under normal operating conditions. If it decays even more, it can accumulate nitrite instead of nitrate, since *Nitrobacter* bacteria are more sensitive to lack of oxygen [9]. This could be problematical because nitrite will inhibit the microorganisms to grow on smaller concentrations than nitrate.
To take this influence into account, a similar formulation to the Monod equation can be used, and multiplied by the main growth rate equation. As an example:

$$\mu = \mu_{\max} \frac{S}{k+S}\left(\frac{OD}{k_{OD}+OD}\right) \tag{3}$$

The parameter k_{OD} measures how much the microorganism is sensitive to the lack of oxygen, so it can be assumed that its value for the *Nitrobacter* type of bacteria is greater than for the *Nitrosomonas* type.

2.2.2. pH and alkalinity
The control of the pH can be as important as the dissolved oxygen control, if the affluent has high variations on the nitrogen concentrations. Its influence can be added by an efficiency term based on experimental results [2], or by considering the ammonium equilibrium [3]. Free ammonia is very toxic for the biological agents and should be avoided.
Alkalinity control, on the other hand, is necessary for nitrification, but usually is not so crucial. A simple stoichiometric balance should define the minimum amount of alkalinity.

2.2.3. Temperature
Temperature influence on biological processes is modeled with the Arrhenius equation until the maximum growth rate, followed by a steep decrease, due to cellular lyses. For the nitrification, the maximum growth rate is around 30 ºC.

2.2.4. Inhibitor compounds
Other compounds that are not related in the biochemical pathway can affect its efficiency. Wastewaters can have several different compounds that can inhibit nitrification such as phenol, ethanol, acetone, free ammonia, and nitrite [2,4,6]. There are several equations to model inhibition effects in literature.

3. Study of an industrial case

The wastewater treatment facility from a petroleum refinery located at the south of Brazil is composed of a complete primary wastewater treatment, that is, a fully removal of free and emulsified oil and particulates by physicochemical processes, followed by an aerated lagoon that promote the biological removal of organic carbon and by a newer

plant of rotating biological contactors (RBC), with 4 sets of RBCs: the first two are used for nitrification, the third one (with submerged disks) for denitrification, and the last one for polishing (excess organic carbon remover).
A previous study on the biological processes revealed that the facultative lagoon cannot promote nitrification, but it removes all the biodegradable organic carbon and also acts as an equalization tank, to relief intense disturbances on the wastewater composition, that are typical of this kind of industry.

4. Modeling the nitrification process

Usually, the data available from an industrial plant is limited to the necessary for the process control. Therefore, we developed a model simple enough to be evaluated from readily available data, which needs to contemplate the most important phenomena on the process. Lin [7] has published a more complex model that could be of interest if there were more available data.
Our model is applied to the second set of RBCs, since they have instrumentation that provides on-line data for inlet ammonia concentration and outlet ammonium concentration, dissolved oxygen and pH. The analyzers provide measurements with 3-hours sampling time and the data were selected from time windows when the process was considered to be running without major disturbances or abnormal operation.

4.1. Mass balance

The basis of the model derives from the mass balance on the volume V of liquid around the disks and a mass balance on the volume occupied by the attached biomass (called the biofilm). The first basic assumptions are that the volume of liquid is well mixed and the biological activity of the suspended biomass in the liquid is negligible ($X_{liq} << X$). Therefore all the biological transformations occur on the biofilm. Also neglecting the mass transfer resistance of ammonium from the liquid to the biofilm, we get the following preliminary form of the mass balance for the biomass and ammonium:

$$\frac{d}{dt}\left(V_{biofilm}X\right) = \int\left(\mu - b_{metabolism}\right)XdV_{biofilm} - \int b_{shear}XdA_{biofilm} \tag{4}$$

$$\frac{d}{dt}\left(V\ NH_4^+\right) = F\left(\ NH_{4\ in}^+ - NH_{4\ out}^+\right) - Y_{NH_4^+|X}\int \mu XdV_{biofilm} \tag{5}$$

To model the biofilm growth we use the hypothesis that the biofilm is homogeneous, based on the idea that the *Nitrobacter* type of bacteria should grow near to the *Nitrosomonas* type (and the available nitrite), in agreement with the work of Carrera et al. [1], when comparing experimental data from immobilized and suspended nitrification biomass. In this way, the biofilm has variable volume and constant biomass concentration. Although, if both biological agents had independent dynamics, it could be necessary to consider a biofilm composed of two components with distinct dynamics and growth.
Considering that the biofilm is a thin layer, with the occupied area $A_{biofilm}$ much larger than its depth L, we can rewrite the left hand side of the Eq. 4 as:

$$\frac{d}{dt}\left(V_{biofilm}X\right) = XL\frac{dA_{biofilm}}{dt} \tag{6}$$

Defining ε as the fraction of occupied area on the RBC by the biofilm, $A_{biofilm}/A_{total}$, and combining both metabolism and shear loss parameters in one, b, Equation 4 can be rewritten as:

$$\frac{d\varepsilon}{dt} = \left(\mu - b\right)\varepsilon \tag{7}$$

In order to consider limitations on the growth and loss terms, we multiplied them by factors of ε, Eq. 8, implying that the growth should decrease as the available area gets fully occupied and the loss (considering it mostly by shear forces) should decrease as the occupied area also decreases. The new parameter β should be chosen (or estimated) to fit these limitations in the best way.

$$\frac{d\varepsilon}{dt} = \left[\mu\left(1-\varepsilon\right)^{\beta} - b\varepsilon^{\beta} \right]\varepsilon \tag{8}$$

The final form for the ammonium balance is the following::

$$\frac{d}{dt}\left(V\ NH_4^+\right) = F\left(\ NH_{4\,in}^+ - NH_{4\,out}^+\right) - Y_{NH_4^+}\mu\varepsilon \tag{9}$$

4.2. Operating conditions influence

As stated before, pH, temperature, and alkalinity are well controlled on this case, so their influence would not be appreciable on the model. Also, due to the long contact time between the biofilm and the air and turbulence around the disks, the RBCs are very effective with the aeration. The dissolved oxygen analyzers, located after the set of RBCs, indicate that it is always above 4 mg/L, which is safe to neglect its influence.

4.3. Nitrification kinetics

Since the affluent of the studied set of RBCs already passed through nitrification from the previous set, concentrations greater than 20 mg/L have not been achieved. Therefore, it is acceptable to use Monod equation, keeping the number of parameters to estimate as low as possible. And since there is no deficiency of dissolved oxygen, the growth of bacteria of the type *Nitrobacter* is always faster (the limiting dynamics is with the type *Nitrosomonas*), justifying the use of the homogeneous biofilm model.

5. Model implementation and results

The whole model was implemented, simulated, and fitted to the process data in the software EMSO [8] (Environment for Modeling, Simulation, and Optimization), maintained by a consortium of Brazilian universities and petrochemical industries. EMSO is a generic equation-oriented simulator for dynamic systems and its main applications have been to model equipments and processes of the chemical and petrochemical industry. Due to its dynamic principle, it can be used to simulate units on start-up and shutdown events, as well as changes on the operating point.

Since the set of RBCs is composed by 8 units, being 2 parallel series of 4 units, the final model is the series association of 4 instances of the model described before, with the equivalent volume of two real RBCs.

As we already stated, the first set of RBCs already nitrificates and so the affluent to this set of RBCs has considerable amount of nitrite and nitrate with no available data. Then, we can only compare the ammonium removal, as seen on Fig. 1.

The major difficulty on this implementation is with initial conditions for the RBC, as it is a continuous process with real dynamic conditions and the variable ε cannot be measured. Therefore all the initial values of the state variables need to be estimated, beside the kinetic parameters.

5.1. Results analysis

The model parameters were estimated using the maxim verisimilitude procedure, solved by IPOPT optimizer, a primal-dual interior point method, which is implemented on EMSO. The first set of data of Fig. 1 and validated with the second set, and their values are: μ_{max} = 0.0342 h^{-1}, b = 0.005 h^{-1}, k = 35.26 mg/L, $Y_{NH_4^+}$ = 1.78 x 10^{-6} g^{-1} and β = 1.3.

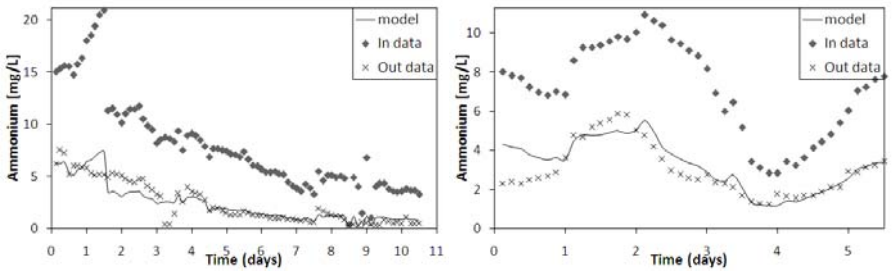


Fig. 1: Ammonium curves before and after the set of RBCs and the model prediction, on two different occasions, using the same kinetic parameters.

The model can predict well the system behavior, and the estimated parameters obey to what is expected from their meanings. The obtained values of ε are smaller than 0.5, which is considered a representative result, based on the experience with the process.

6. Conclusion

The proposed model for ammonia removal in RBCs allowed a better understanding of this process and can result on several improvements on the real process. In the studied industrial case, the analysis of the nitrification kinetics, from the process point of view, revealed the best way to operate it and explained why some actions did not work as expected before. The simplified model has shown that is capable to describe the process fairly well. Maybe a more detailed model would be desired, but it would need more data to estimate the additional parameters, which could be obtained from laboratorial analyses or pilot-scale experiments. A wider approach of the model is been developed and it will be used for operators training with the simulation of how to proceed in adverse conditions. Further on, it could be used as a virtual analyzer, for advanced control of the process.

References

[1] Carrera, J., Jubany, I., Carvalho, L., Chamy, R., Lafuente, J., 2004, Kinetic models for nitrification inhibition by ammonium and nitrite in a suspended and an immobilised biomass system, Process Biochemistry, 39, p. 1159-1165.

[2] Ferreira, E.S., 2000, Chemical kynetics and fundaments of nitrification and denitrification biological processes, XXVII Interam. Congress of Sanitary and Environmental eng., I-121.

[3] Béline, F., Boursier, H., Daumer, M.L., Guiziou, F., Paul, E., 2007, Modeling of biological processes during aerobic treatment of piggery wastewater aiming at process optimisation, Bioresource Technology 98, p. 3298-3308.

[4] Vania, M., J., S., 2000, Nitrification and denitrification on the hydrous effluent from REFAP: Pilot and industrial scale study, Technical report 01/2000 from the environment sector from CENPES.

[5] Vasiliadou, I.A., Tziotziosa, G., Vayenas, D.V., 2008, A kinetic study of combined aerobic phenol and nitrate removal in batch suspended growth cultures, Inernational Biodetereioration & Biodegradation, 61 (3), p. 261-271.

[6] Kim, Y.M., Park, D., Lee, D.S., Park, J.M.. 2008, Inhibitory effects of toxic compounds on nitri cation process for cokes wastewater treatment, J. Hazard. Mater., 152 (3), p. 915-921.

[7] Lin, Y.-H., 2008, Kinetics of nitrogen and carbom removal in a moving-fixed bed biofilm reactor, Applied Mathematical Modeling 32 (11), p. 2360-2377.

[8] Soares, R.P., Secchi, A.R., 2003, EMSO: A new environment for modeling, simulation and optimization, In Proceedings of ESCAPE 13. Elsevier Science Publishers, p. 947-952.

[9] Ruiz, G., Jeison, D., 2003, Nitrification with high nitrite accumulation for the treatment of wastewater with high ammonia concentration, Water Research 37, p. 1371-1377.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Heuristic Particle Swarm Optimization for Integration Problem of Batching and Scheduling in Chemical Industries

Lixin Tang,[a,*] Qi Tang,[a,b] Ping Yan[a]

[a] *Liaoning Key Laboratory of Manufacturing System and Logistics, The Logistics Institute, Northeastern University, No.11, Lane 3, WenHua Road, HePing District, Shenyang, 110004, China*
[b] *The Department of Logistics Management, Shenyang University of Technology, Shenyang, 110023, China*

Abstract

This paper considers the integration problem in the batch production mode of chemical industries, which contains a series of decisions such as the size and amount of batches, assignment of the batches on the units, sequence of the batches on each unit, making a timetable of batches production. For such integration problem, we develop a new solution approach, Heuristic Particle Swarm Optimization (HPSO) which combines some heuristics with PSO in order to obtain feasible solution in a simple method. Then two tactics are presented to speed up the algorithm. The performance of the new solution procedures for the integration problem is demonstrated by solving several instances of a case study from process industries. In contrast to the methods in the literature, HPSO can solve problem with sufficient accuracy in faster speed and principle simple. Moreover, HPSO can solve much larger practical problems.

Keywords: integration, HPSO, chemical industries.

1. Introduction

Integration of batching and scheduling in chemical industries has been the subject of extensive research since the early 1990s. Comparing with separate batching and scheduling, the integration can shorten the production time. The range of application areas for the theory goes beyond the chemical industries to include medicine, foodstuff, etc. The main focus is on the appropriate decision of batch size on units and efficient allocation of one or more resources to batches over time.

On the production environment, integration of batching and scheduling in chemical industries can be classified into two main groups such as singe stage and multiple stages. As for single-stage, Erdirik-Dogan and Grossmann (2007) present a multi-period mixed integer linear programming (MILP) formulation to deal with parallel units, which is effective for small size. They propose bi-level composition algorithm to decompose the problem into batching and scheduling sub-problems. The sub-problems were resulted corporately. Fleilschmann and Meyr (1997) present a local search algorithm to deal with one machine. As for multiple stages, there are three classes such as flowshop (FS), hybrid flowshop (HFS) and job shop (JS). Liu and Karimi (2005) present a MILP formulation which can deal with hybrid flowshop using a continuous-time representation with synchronous slots and a new idea of several balances such as time,

mass, resource. Janak and Floudas (2007) address a MILP formulation for job shop based on unit-specific event continuous-time approaches.

In the literature, most are based on MIP formulation. However, the proposed MIP models often have complex structures and large integrality gaps and are inefficient for solving practical problems. This paper employs another approach to do research, which is based on the mechanisms of problems. This approach has received very limited research attention and few papers can be found. Fleilschmann and Meyr (1997) has addressed on the problem of single stage. This paper studies integration of batching and scheduling in HFS, which is very useful as many practical production systems can be considered as HFS.

The remaining contents are organized as follow: Section 2 is problem description. Section 3 is Heuristic Particle Swarm Optimization (HPSO) algorithm. Section 4 is computational results. Finally, section 5 is conclusions.

2. Problem Description

In this paper, we consider integration problem of batching and scheduling a product in HFS with nonidentical machines and finite intermediate storage. Fig.1 shows the production process of the problem. The HFS has m stages with parallel machines. There are multiple products. The process of a product in a stage is defined as a task which is fulfilled in the batch mode. The relationship of the processing time and batch size is given. The objective is to minimize makespan.

The problem has most of the features in HFS, which is described as follows:

(1) Process consists of several stages. Raw material is transformed into intermediate and final products through some stages.

(2) There are several parallel units which are identical and non-identical in the each stage.

(3) Batches are produced on the units. Relationship of time and batch size is provided.

(4) Non-preemptive. Once a batch begins, batch can't stop until the whole batch is over. The batch is carried out without interruption.

(5) Finite intermediate storage should be considered between stages.

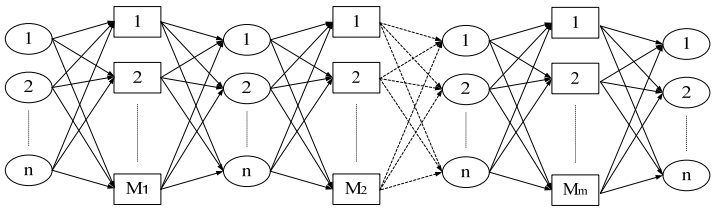


Fig.1 production Process of hybrid flowshop

3. HPSO Algorithm

PSO is a population based evolutionary meta-algorithm which is first presented by Kennedy and Eberhart. The swarm population is made of particles. A particle is correlated to a solution. The particle is composed of position and velocity. The positions are placed in the search space of some problem or function. And the velocity determines position's movement by combining information from its own and swarm. All particles improve their position and velocity in the mode of iteration. The swarm move close to an optimum of the fitness, like a flock of birds collectively foraging for food.

Integration of batching and scheduling is a complex problem. PSO is difficult to obtain feasible solution. Therefore, we propose hybrid PSO (HPSO) which combines some heuristics with PSO in order to obtain feasible solution in a simple method. Then we present two tactics to speed up the algorithm.

3.1. PSO

We use position value from the iterative formulation of standard PSO to denote the production quantity of units, which deals with the material balance between production levels.

$$y_{ijt} = \omega y_{ijt-1} + c_1 r_1 \left(xbest_{ij} - x_{ijt-1} \right) + c_2 r_2 \left(gbest_j - x_{ijt-1} \right) \quad (1) \qquad x_{ijt} = x_{ijt-1} + y_{ijt} \quad (2)$$

3.2. Some Heuristics

Based on the production quantity of units from PSO, we give some heuristics to obtain feasible solutions.

3.2.1. Batching

Based on the analysis of the problem, we found out that the less batches, the more possible to improve the results of minimized makespan. We develop heuristic algorithm to do batching as follow:

$$a = \frac{\text{Production quantity of unit}}{\text{maximal Limit of batches on the unit}} \quad (3) \qquad \text{Minimal batch quantity} = \lceil a \rceil \quad (4)$$

$$batchsize_i = \text{maximal Limit of batches on the unit} \qquad i = 1, \cdots, \lfloor a \rfloor \quad (5)$$

$$batchsize_i = \text{Production quantity of unit} - \sum_{j=1}^{\text{Minimal batch quantity}} batchsize_j \quad i = \lceil a \rceil \quad (6)$$

3.2.2. Precedent Relationship

Precedent relationship implies time relations between batches. Constrains such as material balance, storage limitation can be represented by precedent relationship. Problems with precedent relationship are similar to single machine problem and will be result in a simpler method. We provide Batch Splitting Mechanism (BSM) and Precedence Search Scheme (PSS) to describe precedent relationship with the ideas of RCPSP.

(1) BSM

This paper describes the problem as nonstandard RCPSP. In the literature, different algorithms have been suggested based on Schedule Generation Scheme (SGS) [1] for standard RCPSP. However there is no general scheme to result nonstandard RCPSP. We provide Batch Splitting Mechanism (BSM) to change nonstandard RCPSP to standard RCPSP.

1) Standard RCPSP

Precedence concept can be extended by so-called minimal time lags between the start times of two batches, implying $s_i + d_{ij}^{\min} \le s_j$.

2) Nonstandard RCPSP

Considering constrains of the finite intermediate storage between two stages. Some precedence relations between the start times of batches from the two stages have to be presented by the maximal time lags, implying $s_i + d_{ij}^{\max} \geq s_j$.

3) Change into Standard RCPSP

In order to make problems easy, we propose Batch Splitting Mechanism (BSM) to change nonstandard RCPSP into standard RCPSP. BSM is to split one batch into head and end activities in the project. The head activity implies the start processing the batch and the end activity implies off from processing machine. The limited intermediate storage constrains can be described by the minimal time lags between the start time of j in the later stage and the finish time of i in the former stage, implying $s_j + d_{ij}^{\min} \leq f_i$.

(2) PSS

We use the PSS to set up the network of precedent relationship which consists of node and line. The node implies the start and finish time points of batches and the line imply the span of time. After analysis we found out four kinds of relations in this problem that are shown as follow:

1) Relation in the same batch. The start and finish time points of the same batch should be linked.

2) Relation in the same task. There are two cases. One is that batches in a task have been assigned on the units. The other is that batches haven't been assigned. As for the first case, we should link finish time point and start point between the two batches. And as for the second case, we should link the start time points of the batches.

3) Relation to meet the supply of materials. To meet the supply of materials means that quantity of production in the former step should satisfy consumption in the next step. And among the point satisfying the conditions we choose the first.

4) Relation to meet storage constrain. With finite intermediate storage, some batch in the former step can't end until consumption in the next step free storage room.

3.2.3. Sequencing

Sequencing implies the sequence to schedule activities in the network of precedent relationship, which can guarantee our solutions feasible. The activities will be chosen one by one by the method as follow:

(1) Find out available set. A implies set that all its precedent activities have scheduled. B implies scheduled activities. C implies filtrated set. Filtrated set means once head activity of a batch happens, others in the same machine won't start until the end of the batch happen. $D = A - B - C$ Means set in which we can choose one to put into sequence this time.

(2) Sort activities. We choose an activity with large priory in D every time. The priory is given out by heuristic algorithm as follow: the number of the time points is given before we schedule the batches. Then the number is the priory of the time points.

(3)At last we gain a sequence satisfying precedence restrictions.

3.2.4. Timetable

Timetable implies to give a timetable to schedule the batches. The idea is to locate time for the activities in the sequence as early as possible. And at meantime, we complete machine assignment. We describe the method as follow:

(1) We use different way to calculate the start and finish time of batches. As for start time, it is the maximum of the machine released time and precedent point time. As for the finish time, it is the maximum of the machine released time, precedent point time and start time of the batch plus processing time.

(2) We propose heuristic algorithm to assign batches on parallel machine. Here we only consider the start time. And finish time should be assigned on the same machine as the start time.

The heuristic algorithm is as follow:

1) Calculate the time that the point is possible to start on the parallel machines.

2) Comparing these value and choose the minimum as the start time.

3.3. Two tactics

In order to speed up to find feasible solutions, we propose two tactics as follow.

3.3.1. No-wait

The tactic is to deal with no-wait relationship between the batches. In the network, it is possible that two activities are precedent one of each other, which implies the no-wait relation. To simple the problem, we only consider one side in the former. Then here, we deal with the other side, which is described as follow:

(1) Based on the former process, we can gain a time table.

(2) Mark the two point of no-wait.

(3) If the time of two points is not equal, move the front point to the time of back point.

3.3.2. Processing time

The tactic is to make the difference between the start and finish time of the same batch equal to the processing time. The method is described as follow:

(1) Check the start time point according to the sequence obtained above.

(2) Find out the first one that the difference between the start time and its corresponding finish time is not equal to the processing time.

(3) Move the start time point until the difference is equal to the processing time.

(4) Reschedule back points in the sequence.

(5) Repeat the above to make all the time spans of batches are equal to the processing time.

```
Initial HPSO parameters
for i=1 to n //each particle
    Initial HPSO variables
    Obtain production quantity of unit
    Obtain batch size
    Obtain precedent relationship
    Sequence the batches
    Compute fitness value
        Unit assignment and time allocation
        Simultaneity of no-wait relationship
        Make span of a batch equal to processing time
    Obtain the best solution of the particle
Obtain the best solution of the all particles
for j=1 to t//each iteration
    for i=1 to n //each particle
        Particle flies to new place
        Update production quantity of unit
        Update batch size
        Update precedent relationship
        Update sequence
        Compute fitness value
            Unit assignment and time allocation
            Simultaneity of no-wait relationship
            Make span of a batch equal to processing time
        Update the best solution of the particle
    Update the best solution of all particles
Output the optimization solution
```

Fig 2. Procedure of HPSO

4. Computational Studies

There are obvious difference between MIP formulation and HPSO. The HPSO is implemented with pure neighborhood search. To evaluate our algorithm, we compare it with that of S&K (Sundaramoorthy and Karimi, 2005) [2] using the same example. There are two demand values in the example such as 2000mu and 4000mu. The objective is to minimize the makespan. It is usual to measure the effectiveness of an algorithm by relative deviation of value of solution that is from the best know solution value. In the literature, the results of MIP formulation are 29.772h and 56.432h. Running the HPSO, results are 28.8h and 53.3h. HPSO improved 3.26% and 5.36%. From the computation analysis, we prove HPSO can obtain effective solution in limited time.

For give an indication of the consistency of HPSO in the larger size problem, we extend the demand value from 2000mu to 10000mu in the example. When evaluating HPSO, there is clearly a balance between the consumption of computation time and the quality of solution. To compare fairly, different methods should run in the same computation time. Otherwise it is difficult to draw firm conclusions. Then the run time is limited in 180 seconds. Table1 shows the results belong to S&K and HPSO for each instance. The results show that HPSO is efficient to deal with larger problems.

Table 1 Solution from S&K and Ours

	2000mu	3000mu	4000mu	5000mu	6000mu	7000mu	8000mu	9000mu	10000mu
S&K(h)	29.772	42.769	56.432	70.101	81.417	97.718	109.62	124.61	137.02
Our(h)	28.668	41.328	53.4644	66.6565	79.0458	90.7359	103.678	115.981	131.162

5. Conclusion

Integration problem of batching and scheduling multistage plants with nonidentical parallel machines and finite intermediate storage in chemical industries is a complicate problem. Methods in the literature are mostly based on MIP. In this paper, we develop HPSO which is based on the mechanisms of problems. HPSO combines some heuristics with PSO in order to obtain feasible solution in a simple method. Then two tactics are presented to speed up the algorithm. From computational studies, we can bring out conclusions as follow:

1) The proposed MIP formulations in the literature often have complex structures and large integrality gaps. Our HPSO is implemented with pure neighborhood search, which is simpler in finding a near-optimal solution. From the computation analysis, HPSO can obtain effective solution in limited time.

2) MIP formulations are often inefficient for solving practical problems. HPSO is a promising method to deal with very large and difficulty problem in the real world.

3) In the future, we should develop approaches to improve the qualities and decrease the computing time. Besides, more constrains such as arrival time and deadline should be considered.

References

S. Hartmann, 1999, Project Scheduling under Limited Resources, Springer, Berlin, Germany, 61-65

A. Sundaramoorthy, I.A. Karimi, 2005, A simpler better slot-based continuous-time formulation for short-term scheduling in multipurpose batch plants, Chemical Engineering Science, 60: 2679–2702.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Dynamic modeling of comminution using a general microscale breakage model

Rodrigo M. de Carvalho[a], Luís Marcelo Tavares[a]

[a]*Department of Metallurgical and Materials Engineering, COPPE – Universidade Federal do Rio de Janeiro, Rio de Janeiro – RJ, Brazil.*

Abstract

The traditional population balance model has been successfully used in the past to describe size reduction in industrial crushers and grinding mills, but presented limitations in scale-up from information obtained in the laboratory. The paper presents a general model that describes comminution of multi-component feeds in full-scale machines using a combination of information from breakage mechanisms, fundamental material characteristics, and the mechanical environment, allowing to overcome the limitations of traditional population balance model formulations. Since it is based on a detailed description of each stressing event, this model can be used to describe particle size reduction in different types of crushers and mills using the same fundamental material characteristics. In order to demonstrate its potential, the model has been applied to describe grinding in a ball mill, where the mechanical environment was predicted using the Discrete Element Method and breakage of particles was described using a model based on continuum damage mechanics. A good agreement was observed between measured and predicted results from a batch grinding test for a single-component feed. Simulations of the dynamics of grinding in a continuous mill demonstrated the effect of change in feed composition in both product size and fracture strength of particles discharged from the mill.

Keywords*:* microscale, breakage, comminution, grinding, modeling.

1. Introduction

There has been great success in the application of the population balance model (PBM) in the modelling of comminution machines (Herbst, 1979). By decoupling the contributions of selection, breakage and transport, size reduction in different types of crushers and mills has been successfully described quantitatively. While excellent in describing size reduction, the traditional PBM proved to be relatively weak as a predictive model. This is particularly critical, given the great demand for reliable tools for scale-up to industrial-sized machines from information obtained at a small scale in the laboratory. One of the reasons has been the dependence of both material and process on the functions used to describe breakage in the traditional PBM framework, which is obviously undesirable. It results in an additional limitation of the traditional PBM, in which functions determined for one particular type of equipment cannot be used to describe size reduction in another machine.

These and other limitations to the traditional PBM can be overcome with the development of a general microscale model of comminution. By combining a detailed description of the mechanical environment in the size reduction device, fundamental breakage characterization tests, which describe the response of the material to the

different breakage modes (body and surface) and stressing intensities, to a detailed description of the material transport, the proposed model allows to describe size reduction in different devices using a unified mathematical framework. While the mechanical environment and the transport in crushers and grinding mills may be precisely described using techniques such as the Discrete Element Method (DEM) and computational fluid dynamics (Mishra and Rajamani, 1992; Litcher et al., 2008; Cleary et al., 2006), experimental techniques and models are available which allow characterizing in detail the response of material to the different breakage models found in size reduction machines (Tavares, 2007).

The present paper proposes a general microscale model, which considers that the material contained in the feed is not homogeneous, but constituted by a number of components, each containing particles having a distribution of strengths (actually fracture energies). The model also considers that when particles suffer stresses that are insufficient to cause catastrophic (body) fracture, they will undergo surface breakage (abrasion/chipping) and may become progressively weaker. The application of the model has been demonstrated to batch grinding in a laboratory ball mill and also to investigate the dynamics of a continuously-operated ball mill, subject to disturbances in feed rate and composition of the feed material.

2. Model development

2.1. Generalized model

The generalized population balance for particles contained in size class i and component class l is

$$\frac{d\left[w_{il}(t)M(t)\right]}{dt} = w_{il}^{in}(t)W^{in}(t) + \omega\left[-D_{il,b} - D_{il,s} + A_{il,b} + A_{il,s}\right] - w_{il}^{out}(t)W^{out}(t) \qquad (1)$$

where $M(t)$ the mass of material at time t inside the mill, also called hold-up. $w_{il}(t)$, $w_{il}^{in}(t)$ and $w_{il}^{out}(t)$ are the mass fractions of particles contained in class i l in the mill, in the feed and in the discharge streams, respectively. W_{in} and W_{out} are the feed and discharge mass flowrates, respectively. ω is the frequency of stressing events in the comminution machine. Functions A and D represent the rate of appearance and disappearance of material in class i l due to fracture, being defined in Table 1, whereas subscripts b and s stand for body and surface breakage mechanisms.

Table 1. Definition of terms in Equation (1)

Breakage mode	Rate of appearance (A)	Rate of disappearance (D)
Body (b)	$\sum_{j=1}^{i-1} w_{jl}(t)\int_0^\infty m_j(E)p(E)\int_0^1 b_{ijl}(eE,t)F_{jl}(eE,t)p(e)de\,dE$	$w_{il}(t)\int_0^\infty m_i(E)p(E)\int_0^1\left[1-b_{iil}(eE,t)\right]F_{il}(eE,t)p(e)de\,dE$
Surface (s)	$\sum_{j=1}^{i-1} w_{jl}(t)\kappa_{jl}\int_0^\infty m_j(E)p(E)\int_0^1 a_{ijl}\left[1-F_{jl}(eE,t)\right]p(e)de\,dE$	$w_{il}(t)\kappa_{il}\int_0^\infty m_i(E)p(E)\int_0^1\left[1-F_{il}(eE,t)\right]p(e)de\,dE$

In the equations listed in Table 1, $p(E)$ is the distribution of stressing energies E in the comminution machine, m_j is the mass of particles contained in size class j captured in each stressing event and $p(e)$ is the energy split among these particles. a_{ij} and b_{ij} are the breakage functions in density form, corresponding to the mechanisms of surface and

body breakage, respectively, the later depending on stressing energy (Tavares and Carvalho, 2009). κ is the surface breakage rate.

Equation (1) should be solved simultaneously with the equations describing the overall mass transfer in the equipment and the other describing how the fracture probability distribution of each component varies with time, which may be calculated by

$$F_{il}(E,t+\Delta t)=\frac{G_{il}F^{*}_{il}(E,t+\Delta t)+(H_{il}+Q_{il})F_{il}(E,0)+I_{il}F_{il}(E,t)}{G_{il}+I_{il}+H_{il}+Q_{il}} \tag{2}$$

where the fracture probability distribution of the original material $F_{il}(E, 0)$ is well described using the upper-truncated lognormal distribution (Tavares and King, 2002), given by

$$F_{il}(E,0)=\frac{1}{2}\left[1+\operatorname{erf}\left(\frac{\ln E^{*}-\ln E_{50,il}}{\sqrt{2\sigma_{il}^{2}}}\right)\right] \tag{3}$$

where

$$E^{*}=\frac{E_{\max,il}\,E}{E_{\max,il}-E}\ \text{ and }\ E_{50,il}=E_{\infty}\left[1+\left(\frac{d_{ol}}{d_i}\right)^{\phi}\right]\overline{m}_{p,il} \tag{4}$$

where $E_{50,il}$ is the median particle fracture energy, σ_{il}^{2} is the variance, $E_{\max,il}$ is the upper truncation of the distribution, E_{∞}, d_{ol} and ϕ are material parameters and d_i and $\overline{m}_{p,i}$ are the mean size and weight of particles in class i, respectively.

$F_{il}(E, t)$ is the distribution of fracture energies of the material that did not suffer any impact event during the time interval and $F_{il}^{*}(E,t+\Delta t)$ is the distribution of fracture energies of the particles that suffered impact, but did not fracture, being given by

$$F_{il}^{*}(E,t+\Delta t)=\frac{\int_{0}^{E_i^{*}}p(E_k)\int_{0}^{1}\left[\frac{F_{il}\left[E/(1-D),t\right]-F_{il}(eE_k,t)}{1-F_{il}(eE_k,t)}\right]p(e)\,de\,dE_k}{\int_{0}^{E_i^{*}}p(E_k)\,dE_k} \tag{5}$$

and

$$D=\left[\frac{2\gamma(1-D)}{(2\gamma-5D+5)}\frac{eE_k}{E}\right]^{\frac{2\gamma}{5}} \tag{6}$$

where E_{il}^{*} is the maximum fracture energy of particles contained in class $i\ l$, which is equal to $E_{\max,il}$ in the first time interval and $F_{il}(E_{il}^{*})=1$ as comminution progresses (Tavares and Carvalho, 2009). D is the damage accumulation parameter, given by a model based on continuum damage mechanics (Tavares and King, 2002; Tavares, 2009), where γ is the only material-specific constant. The model is able to account for the fact that particles that are stressed but that do not fracture in a stressing event, may become progressively weaker. The various terms in Equation (2) are given by

$$G_{il}(t)=\frac{\omega\Delta t}{M(t)}w_{il}(t)(1-\kappa_{il})\int_{0}^{\infty}m_i(E)\,p(E)\int_{0}^{1}\left[1-F_{il}(eE,t)\right]p(e)de\,dE \tag{7}$$

which is the fraction of material in the class that has been damaged but remained in the original size range, and

$$H_{il} = \frac{\omega \Delta t}{M(t)} \left[\sum_{j=1}^{i} w_{jl}(t) \int_0^{\infty} m_j(E) p(E) \int_0^1 b_{ijl}(eE,t) F_{jl}(eE,t) p(e) de\, dE + A_{il,s} \right] \tag{8}$$

which is the fraction of material that appeared due to body and surface breakage, and

$$I_{il}(t) = \frac{w_{il}(t)}{M(t)} \left(M(t) - \omega \Delta t \int_0^{\infty} m_i(E) p(E) dE \right) \tag{9}$$

which is the fraction of material that was not captured in the time interval, and

$$Q_{il}(t) = \frac{w_{il}^{in}(t) W^{in}(t) \Delta t}{M(t)} \tag{10}$$

which is the mass fraction of material that enters the comminution equipment.

2.2. Batch grinding equation

In the case of a ball mill operating in batch mode, Equation (1) may be written as

$$\frac{dw_{il}(t)}{dt} = \frac{\omega}{M} \left[- D_{il,b}(t) - D_{il,s}(t) + A_{il,b}(t) + A_{il,s}(t) \right] \tag{11}$$

as the mill hold-up M remains constant and no material enters or leaves the mill.

Unfortunately, no generally well-accepted functions are available to calculate $m_i(E)$ and $p(e)$. The mass of material captured $m_i(E)$ in the active breakage zone in each collision event is probably influenced by contact geometry, size distribution of particles inside the mill and the friction coefficient between balls and particles (Schönert, 1979). This mass may be calculated by $m_i = N_{cap,i} \overline{m}_{p,i}$, where $N_{cap,i}$ is the number of particles captured in a stressing event, which may be estimated by considering that the finer the size of particles in comparison to balls, the larger their number. A relationship of this type is

$$N_{cap,i} = 1 + a (d_b / d_p)^b \tag{12}$$

where parameters a and b should be fitted to experimental data (Tavares and Carvalho, 2009). $p(e)$ allows to consider that particles involved in each collision inside a mill absorb different proportions of the stressing energy, according to their radial position. A simplifying assumption found to be reasonably valid (Tavares and Carvalho, 2009) is that the stressing energy is split equally among particles positioned within the active breakage zone in the mill. In this case the energy distribution function may be given by

$$p(e) = \delta \left(e - 1 / N_{cap,i} \right) \tag{13}$$

where δ is the Dirac delta function.

Parameters describing the material response to stressing have been determined using fundamental characterization tests, described elsewhere (Tavares, 2007).

3. Results and Discussion

The microscale model was verified by comparing batch grinding results to model predictions calculated using Equation (11). The complete list of material characterization parameters and the DEM collision energy distribution used in the simulations is given elsewhere (Tavares and Carvalho, 2009). Figure 1 shows that the model predicted fairly well the variation of the size distributions with time for grinding

of two different materials and initial sizes (copper ore -1.68 mm and granulite -4.75 mm), specially considering that it required no fitting to batch grinding data. Table 2 shows the parameters used in these simulations.

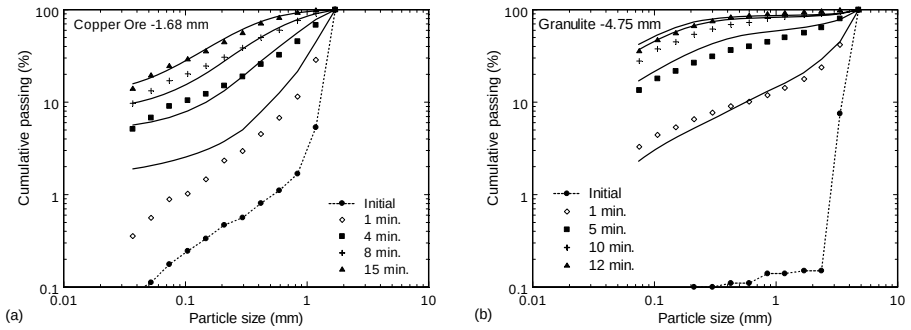


Figure 1. Comparison of measured and simulated results of batch grinding of -1.68 mm copper ore (a) and -4.75 mm granulite (b) in a 30x30 cm ball mill equipped with four lifters and running at 67% of critical speed at 30% mill filling and 100% voids filling.

The full dynamic version of the microscale model has been used to simulate a continuous laboratory ball mill operating with no internal classification. In this simulation, the mill is modelled as a perfect mixer and the relationship between the hold-up and the discharge rate is given by (Carvalho and Tavares, 2006)

$$W_{out}(t) = 0.015\left[M(t)/M*\right]^{1.5} \tag{14}$$

where $M*$ is a reference weight (1.0 kg) and W_{out} is given in kg/min.

The simulations consider that the feed is made up of two components, where component 2 has 50% lower fracture energies than component 1 and particles contained in a single size range (1.68-1.18 mm) are fed to the mill. After being initially fed exclusively by particles of component 1, the feed then becomes equally composed of particles of both components after 5 minutes of grinding. As expected, Figure 2 shows that the product size distribution, characterized by the 80% passing size increases continuously, given the lower average fracture strength of the material entering the mill. Figure 2 also shows that mean fracture energy of particles contained in the third size class increase as a result of the change in feed composition.

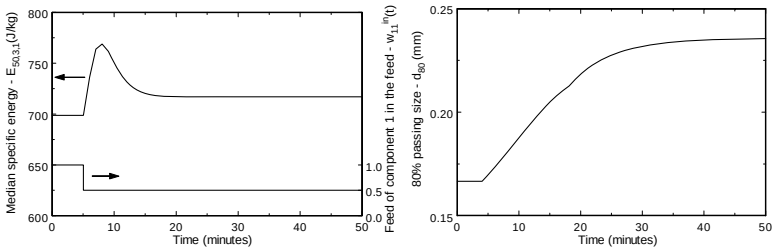


Figure 2. Variation of d_{80} and E_{50} as a function of time for a change in feed composition after 5 minutes of grinding (W_{in} = 0.18 kg/min, M = 5.24 kg)

Table 2. Model parameters used to simulate batch grinding data displayed in Figure 1a and 1b.

Material	Batch mill parameter	Material parameters						
	ω (impacts/s)	ρ (g/cm^3)	E_∞ (J/kg)	d_o (mm)	ϕ (-)	σ (-)	γ (-)	κ (%/s)
Granulite	230000	2.80	123.0	0.91	1.68	0.76	5.0	0.004
Copper ore		2.93	206.7	20.30	0.68	0.74	4.5	0.011

4. Conclusions

A general microscale breakage model that can be used to describe size reduction in different types of crushers and grinding mills was proposed. The model was applied to describe size reduction in ball mills operating both batch and continuously. The validity of the model has been demonstrated by the good correspondence between measured and simulated size distributions from grinding a copper ore in a batch laboratory ball mill.

Although the validation of the full dynamic approach of the proposed model is still a work in progress, which demands intensive experimental tests, this model formulation has been used to simulate a small laboratory ball mill subject to a step change in feed composition (fracture energy), and the results demonstrate that the mill product responds to this variation, in spite of a constant feed rate.

Provided that appropriate information exists that allows describing the milling environment under, for instance, different conditions, such as mill diameter, mill filling, ball size distribution and liner design, the model developed should enable predicting full-scale grinding exclusively using characterization data from small-scale breakage tests.

Acknowledgements

The authors would like to thank the financial support from the Brazilian research council CNPq.

References

P.W. Cleary, M. Sinnott, R. Morrison, 2006. Prediction of slurry transport in SAG mills using SPH fluid flow in a dynamic DEM based porous media, Miner. Eng. 19(15) 1517-1527.

R.M. Carvalho, L.M. Tavares, 2006. Simulação Dinâmica da Moagem A Seco em Moinho de Bolas, Proc. IX Encontro de Modelagem Computacional, Belo Horizonte, pp. 1-10.

J.A. Herbst, 1979. Rate processes in multiparticle metallurgical systems. In: Rate Processes in Extractive Metallurgy (H. Sohn and M.E. Wadsworth, eds.), Plenum Press, pp. 53-112.

J. Lichter, K. Lim, A. Potapov, D. Kaja, 2008. New developments in cone crusher performance optimization, Proceedings, Comminution'2008, Falmouth, pp. 1-9.

B.K. Mishra, R.K. Rajamani, 1992. The discrete element method for the simulation of ball mills, Appl. Math. Model. 16, 598-604.

L.M. Tavares, 2007. Breakage of Single Particles: Quasi-static. In: Handbook of Particle Breakage (A.D. Salman, M. Ghadiri, M.J. Hounslow, eds.), Elsevier, vol. 12, pp. 3-68.

L.M. Tavares, 2009. Analysis of particle fracture by repeated stressing as damage accumulation, Powder Technol., in press.

L.M. Tavares, R.M. Carvalho, 2009. Modeling breakage rates of coarse particles in ball mills, Miner. Eng., in press.

L.M. Tavares, R.P. King, 2002. Modeling of particle fracture by repeated impacts using continuum damage mechanics, Powder Technol. 123, 138-143.

K. Schönert, 1979. Aspects of physics of breakage relevant to comminution, Proc. 4th Tewksbury Symposium, Melbourne, Australia, pp. 3.1-3.9.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Development of Process Inverse Neural Network Model to Determine the Required Alum Dosage at Segama Water Treatment Plant Sabah, Malaysia

A. Robenson[a], S.R. Abd. Shukor[b] and N. Aziz[b*]

[a]*Sabah Water Supply Department, Ministry of Infrastructure Development, Locked Bag No. 1, 91009 Tawau, Sabah, Malaysia*
[b]*School of Chemical Engineering, Engineering Campus, Universiti Sains Malaysia, Seri Ampangan, 14300 Nibong Tebal, Pulau Pinang Malaysia*
**Phone: +604 -5996457, Fax: +604 -5941013, Email: chnaziz@eng.usm.my*

Abstract

The determination of the optimal coagulant dosage in the coagulation process of a water treatment plant (WTP) is very essential to produce satisfactory treated water quality and to maintain economic plant operation such as reducing manpower and expensive chemical costs. Failing to do this will also reduce the efficiency in sedimentation and filtration process in the treatment plant. Traditionally, jar test is used to determine the optimum coagulant dosage. However, this method is expensive, time-consuming and does not enable responses to changes in raw water quality in real time. Modeling such as neural network can be used to overcome these limitations. In this work, an inverse neural network model is developed to predict the optimum coagulant dosage in Segama WTP in Lahad Datu, Sabah, Malaysia. Real data from the WTP was obtained along with extensive data analysis and preparation, significant input-output selection and consideration of important raw and treated water lag parameters were carried out. The modeling results shown that the prediction capabilities are improving with the consideration of appropriate input parameters. Neural network models with different network architectures, including single and two hidden layers were developed and the optimum network architecture obtained was [11-27-9-1]. This model performed very well over the range of data used for training, with r-value of 0.95, mean square error (MSE) of 0.0019 and mean absolute error (MAE) of 0.024 mg/l when applied on the testing data set. Hence, the proposed techniques can significantly improve and have a great potential of replacing the conventional method of jar test due to its advantages; quick responsive tools, economical operating cost and its ability to be applied in real time process.

Keywords: Water treatment plant; Coagulant dosage; Artificial neural network; Inverse model.

1. Introduction

The water industry is working very hard to produce high quality of drinking water at a lower cost in order to meet the mandatory drinking water quality standard. The treatment of water involves physical, chemical and biological changes which will transform raw water into safe drinking water.

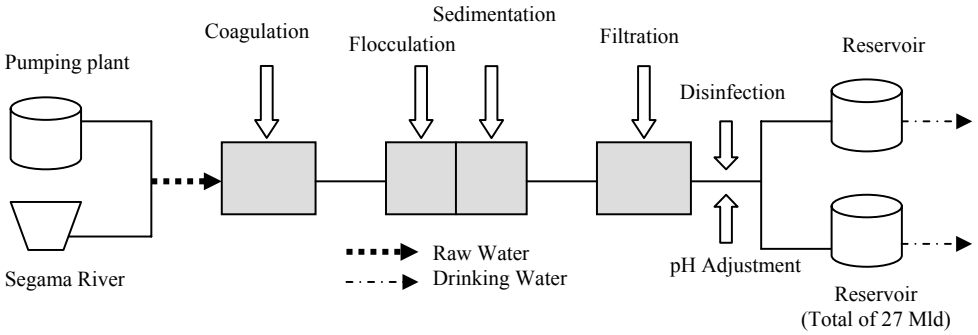


Figure 1: Segama WTP unit processes

A conventional water treatment process normally consists of coagulation, flocculation, sedimentation, filtration, pH adjustment and disinfection processes. The sequence of the treatment processes is shown in Figure 1. Among all processes involved in the treatment plant, coagulation process can be considered as the most important stages as it allows the removal of dirt and colloidal particles. Good coagulation control is very essential to produce satisfactory treated water qualities and to maintain economic plant operation. On the other hand, poor control of the same will cause wastage of chemicals, low water quality and failure in sedimentation and filtration processes.

Traditional method of controlling coagulant dosage relies very much on manual method called jar test (Baxter *et al.,* 1999; Joo *et al.,* 2000; Yu *et al.,* 2000). However, disadvantages associated with jar test are the necessity to rely on manual intervention and lack of adaptation to abrupt changes of raw water characteristics (Valentin *et al.,* 1999). Furthermore, conducting too often jar testing will consume a lot of chemicals for testing, contribute to higher electricity bills and also require an experienced manpower to obtain good results in determining the required coagulant dosage.

Since the coagulation process involve complex chemical and physical phenomena, therefore conventional process control used in chemical engineering is difficult to be applied (Mirsepassi *et al.,* 1995). The limitations of jar test for determining optimum alum dosages can be overcome by using artificial intelligence models such as Artificial Neural Network (ANN) (Yu *et al.,* 2000; Maier *et al.,* 2004).

2. Overview of Segama Water Treatment Plant

Operating data of Segama WTP, a plant which is managed by Lahad Datu Water Supply (LDWS) were used in this analysis. The treatment plant has the capacity of 27 million liters and operating on a 24 hour basis in order to serve a population of about 130,000 in the Lahad Datu District. The raw water comes from Sungai Segama; about 1 km away from the treatment plant area. A large proportion of the upstream watershed is uninhabited mountains and forest, palm oil plantations and mills, marbles factory and with little municipal or industrial development. The major source of upstream contamination is run-off from forested and agricultural lands.

3. Methodology

In order to predict the coagulant dosage required for any characteristic of incoming raw water into the treatment plant, thousands of potential data sets for the year 2005 was obtained. The available data sets were analyzed statistically using Microsoft Excel and

MINITAB 14 software to remove any major outliers and to select the appropriate input-outputs variables of the proposed model.

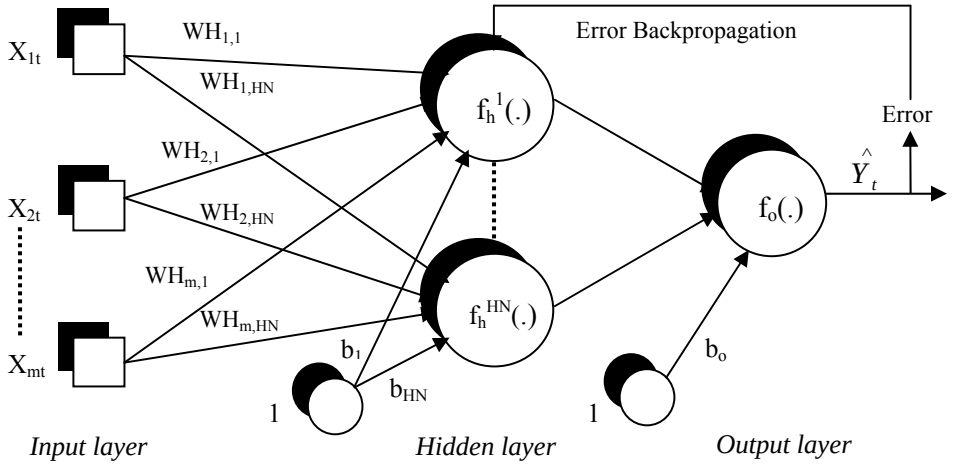


Figure 2: Architecture of multilayer feed forward neural network

All of the data were normalized linearly between 0.0 and 1.0 in order to overcome the significant minimum and maximum values used in the training process and to be commensurate with the limits of the log-sigmoid activation function in the hidden and output of neural network layers. Later, those preprocessed data were divided into training, validation and testing sets for modeling purposes. The neural network architecture which was used in this study is demonstrated in Figure 2, which consists of three layers; input, hidden and output layers. However, network with two hidden layers were also investigated. The hidden layer(s) also consist of PEs and carry out several calculations. The output, $\hat{Y}_t$ of the above neural network model, with one hidden layer and one output neural network is given by the following:-

$$\hat{Y}_t = f_o\left[\sum_{j=1}^{HN} WO_j \times f_h\left(\sum^{m} WH_{ij} \bullet X_{it} + b_j\right) + b_o\right] \tag{2.1}$$

where WH_{ij} is the weight of the link between the i^{th} input and the j^{th} hidden neuron, m is the number of input neurons, WO_j is the weights of the link between the j^{th} hidden neuron and the output neuron, f_h is the hidden neuron activation function, f_o is the output neuron activation function, b_j is the bias of the j^{th} hidden neurons, b_o is the bias of the output neuron, and HN is the number of hidden neurons. The equation of log-sigmoid function, which was used as the transfer function for both hidden and output layers;

$$f(x) = \frac{1}{1+e^{-x}} \tag{2.2}$$

The Matlab® R2006a package was used to develop the ANN process inverse model in a desktop computer with specification of Intel® Core(TM)2 CPU 1.86GHz and 1 GB of RAM. The ANN architecture used is supervised multilayer of feedforward network, and trained using Levenberg-Marquardt training function. In model training, the objective is

to find the appropriate weights and biases which will be further used in the model validation prior to testing phase. The best model architecture was found by cross validation technique, based on its error performances in term of mean square error (MSE), mean absolute error (MAE) and correlation coefficient of r-values between target and network output, all of which targeting the predetermined network goals. To reduce the network complexity, some of the less significant input variables were removed via backward elimination of stepwise regression analysis. Performance of the neural network model was compared among the model developed prior and after the inputs elimination.

4. Results and Discussion

4.1. Data Analysis and Preparation

Ranges of the available raw and treated water parameters as well as the coagulant dosage which covers and represent various operating conditions for the whole year of 2005 are shown in Table 1.

Table 1: Ranges of available data sets, 2005 (Robenson *et al.*, 2006)

	Raw Water					Treated Water						Coagulant Dosage
	pH	Turb (NTU)	Col (HU)	TDS	Alk (mg/l)	pH	Turb (NTU)	Col (HU)	TDS	Alk (mg/l)	Al res (mg/l)	Alum (mg/l)
Min	6.6	11	36	60	50	6.5	0.19	0	90	20	0.01	20
Max	7.9	3405	656	160	170	7.5	5.24	6	170	100	0.2	170

Notation: Turb = Turbidity, NTU = Nephelometric Turbidity Unit, Col = Color, HU = Hazen Unit, TDS = Total Dissolved Solids, Alk = Alkalinity, Al res= Aluminum residue

The quality of the raw water parameters is very uncertain due to natural perturbations and very much affected by the upstream activities of Segama River like palm oil plantations and its mills, marbles factory as well as logging activities. Three main parameters which give great effects in the determination of coagulant dosage i.e. pH, colour and turbidity (Zhang and Stanley, 1997) have shown unpredictable values and very uncertain throughout the calendar year.

The colour and turbidity values keep fluctuating over the year with values ranges from 11 to 3405 NTU and 36 to 656 HU respectively. Even though the plant operation manage to treat the water and producing acceptable drinking water qualities, fluctuations of raw water qualities give difficulties to the plant operators as they will only rely on conventional method of jar tests to determine the best coagulant dosage. Selection of significant input-output parameters was made based on the available literatures, priori knowledge and analytical studies of correlation coefficient.

Among the initial 18 inputs which were used to develope the model, 7 less significant input variables were eliminated based on the stepwise regression of forward and backward elimination. The remaining 11 significant inputs which were used in the model development are RW pH (t) and (t-1), RW Turbidity (t), (t-1) and (t-2), RW Colour (t) and (t-1), TW Turbidity, TW Colour, Alum Dosage (t-1) and (t-2).

4.2. Results of Process Inverse Model

The best network architecture found was [11-27-9-1] with MSE value of 0.0010. When applied to each of the data sets, the corresponding results were consistent with r-value of 0.95 for both the validation sets as well as the testing set. Similarly, the MAE for all data sets was quite small and ranged from 0.015 mg/l to 0.026 mg/l. This suggests that the model architecture is reliable and can be further applied in the real-time plant operation. The trained network was tested by comparing the simulated predictions to the measured values of coagulant dosage using the testing set and the results are presented graphically in Figure 3. The model follows the trends in the actual coagulant data quite well and accurate.

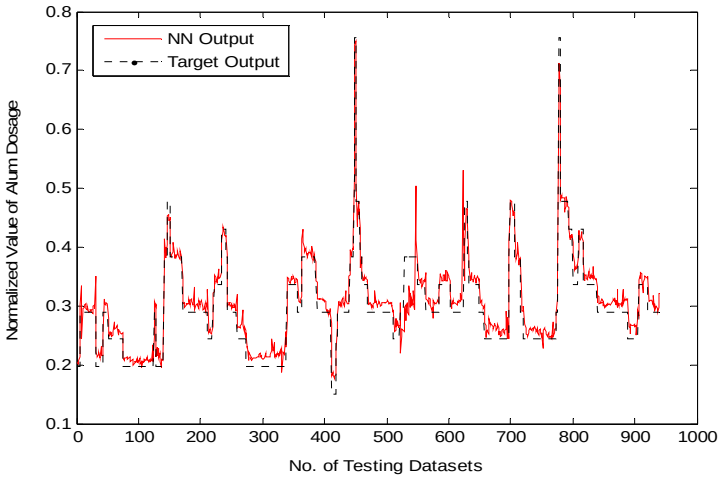


Figure 3: Comparison of actual & predicted coagulant dosage on testing set [11-27-9-1]

In order to ensure that there are no obvious trends in the model residuals, a plot of the residuals across all of the data patterns in the testing set is presented in Figure 4.

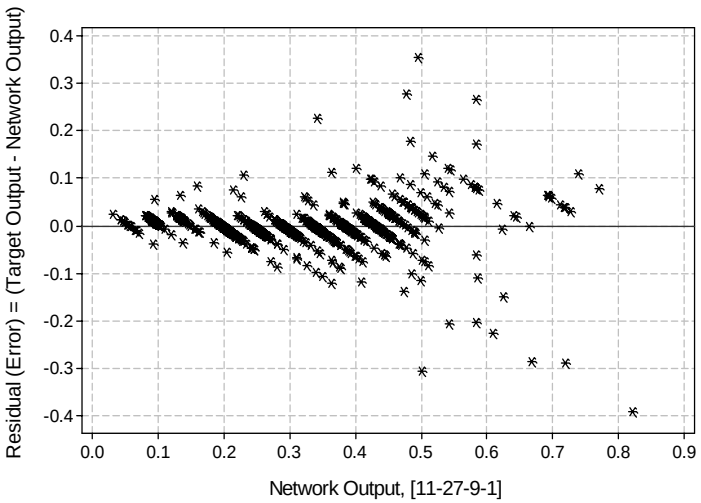


Figure 4: Residual plot of network output on [11-27-9-1]

The prediction capability is confirmed since almost all samples do not diverge from the range [-0.1 mg/l, 0.1 mg/l] of the residual plot. This result is better than the residual range obtained by Lamrini *et al.* (2005), which was [-0.3 mg/l, 0.3 mg/l]. The network with 11 inputs is performed better compared to the models with 18 inputs because those eliminated input parameters may have so many noises or outliers which make them less significant and contribute to poor generalization of the network. Furthermore, those noises will limit the accuracy of generalization that can be achieved no matter how extensive the training set is (Sarle, 2002).

5. Conclusion

Thousands of raw and treated water data sets from Segama Water Treatment Plant, Lahad Datu, Sabah have been used in the development of inverse neural network models, to predict the optimum coagulant dosage in the plant. The results obtained from the models indicate that it is capable of producing accurate predictions over the range of the data used in model training. With the application of such models in water treatment plant, difficulties associated with conventional method, jar test can be overcome due to its advantages; quick responsive tools, economical operating cost and its ability to be applied in real time process.

Acknowledgements

The authors acknowledge the assistance provided by Sabah Water Supply Department and LDWS for the data collection from Segama Water Treatment Plant in Lahad Datu, Sabah. The authors also appreciate the Sabah State Government and Universiti Sains Malaysia in providing financial support throughout the study.

References

C.W. Baxter, S.J. Stanley and Q. Zhang, 1999, Development of a fullscale artificial neural network model for the removal of natural organic matter by enhanced coagulation. *Journal of Water Science Research and Technology-Aqua,* 48(4): 129–136.

D.S. Joo, D.J. Choi and H. Park, 2000, The effects of data preprocessing in the determination of coagulant dosing rate. *Water Research,* 34(13): 3295-3302.

B. Lamrini, A. Benhammou, M.V. Le Lann and A. Karama, 2005, A Neural Software Sensor for Online Prediction of Coagulant Dosage in a Drinking Water Treatment Plant. *Transaction of the Institute of Measurement and Control*, 27(3): 195-213.

H.R. Maier, N. Morgan and C.W.K Chow, 2004, Use of artificial networks for predicting optimal alum doses and treated water quality parameters. *Environmental Modelling and Software*, 19: 919-928.

A. Mirsepassi, B. Cathers and H.B. Dharmappa, 1995, Application of artificial neural networks to the real time operation of water treatment plants. *Proceedings International Joint Conference on Neural Networks,* 1: 516-521.

A. Robenson, N. Aziz and S.R. Abd. Shukor,2006, Water Quality Analysis at Segama River, Sabah, Malaysia. International Conference on Environment, 13-15 November, 2006, Penang, Malaysia.

N. Valentin, T. Denoeux and F. Fotoohi, 1999, A hybrid neural network based system for optimization of coagulant dosing in a water treatment plant. *Proceedings of the International Joint Conference on Neural Networks, Washington D.C, IEEE*: 23-39.

R.F. Yu, S.F. Kang, S.L. Liaw, M.C. Chen, 2000, Application of artificial neural network to control the coagulant dosing in water treatment plant. *Water Science and Technology*, 42(3-4): 403-408.

Q. Zhang and S.J. Stanley, 1997, Forecasting raw-water quality parameters for the North Saskatchewan River by neural network modeling. *Water Research*, 31(9): 2340-2350.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

MIMO Neural Network Model for Pilot Plant Distillation Column

Z. Abdullah[a], Z. Ahmad[b] and N. Aziz[b]*

[a]*Faculty of Chemical Engineering, Universiti Teknologi MARA 40450 Shah Alam, Selangor, Malaysia*
[b]*School of Chemical Engineering, Engineering Campus, Universiti Sains Malaysia, Seri Ampangan, 14300 Nibong Tebal, Pulau Pinang Malaysia*
** Phone: +604 -5996457, Fax: +604 -5941013, Email: chnaziz@eng.usm.my*

Abstract

A distillation column is a complex multivariable system and exhibits nonlinear dynamic behavior due to the nonlinear vapor-liquid equilibrium relationships, the complexity processing configurations and high product purities. In order to gain better product quality and lower the energy consumption of the distillation column, an effective nonlinear model based control system is needed in order to allow the process to be run over a large operating range. The availability of a suitable nonlinear model is crucial important in the development of a nonlinear model based control. Neural networks framework have been used extensively in nonlinear process model development. In this paper, the multiple-input multiple-output (MIMO) neural network model to predict the top and bottom product compositions of a methanol-water pilot plant distillation column was developed. The neural network approach was applied by previous researchers in the development of nonlinear model for continuous distillation column based on simulation data (Yu, 2003; Singh *et al*., 2005 and Singh *et al*., 2007). However, only a handful of works have been carried out to validate the model with real plant data. The validation of the model with real data was important in order to ensure the ability of the model to represent the real processes being considered. Therefore, experimental works was carried out to separate the methanol water mixture in a continuous pilot plant distillation column. The error between the model and the actual pilot plant data was observed.

Keywords: Pilot Plant Distillation column, Modeling, MIMO.

1. Introduction

Distillation columns are fairly complex units and exhibit nonlinear dynamic behavior due to their nonlinear vapor liquid equilibrium relationships, the complexity processing configurations (e.g., prefractionators, sidestreams and multiple feeds) and high product purities (Luyben, 1987).). The dynamics of the distillation column that is variations in time constants with the size and direction of an input change are caused by a mixture of very fast vapor flowrate changes, moderately fast liquid flowrate changes, slow temperature changes and very slow composition changes (Luyben, 2002).

Advanced control strategies such as model based control have been revealed as the better control systems to improve product yield, reduce energy consumption, increase capacity, improve product quality and consistency, increase responsiveness and improve process safety of distillation columns. Focusing on the distillation control problem, development of several advanced control schemes utilizing the neural network

model has been reported. Neural network technique is one of the most useful data driven model that can be utilized in nonlinear model based control system. The neural network provide powerful analysis properties such as complex processing of large input-output information arrays, representing complicated nonlinear associations among data and the ability to generalize.

The main aim of this paper is to develop a nonlinear model by using a neural network technique which is able to represent the continuous distillation column. This is based on the distinct capability of the model to capture the complex dynamic and static interactions of the input-output pattern of the distillation column. In this paper, the multiple-input multiple-output (MIMO) model is developed. The MIMO model consisted of 2 outputs; top and bottom product compositions. Reboiler heat duty, reflux flowrate and tray temperatures were selected as the inputs for the neural network model. The development of this MIMO model is focused on the selection of the input-output of the neural network and development of neural network models based on past and current input which are useful in light of the dynamic systems of the continuous distillation column. Validation with experimental data is carried out in order to determine the ability of the developed models to represent the distillation process.

2. Methodology

2.1. First principle model

The general first principle model for the distillation column which was developed consisted of a non-ideal column, a non-ideal mixture and a non-equimolal overflow (Luyben, 1989). The subroutine of the first principle model was developed in the Matlab® environment. The development of this general first principle model involved mass balance, component balance and enthalpy balance equations, the vapor-liquid phase equilibrium and liquid-hydraulic relationship resulting in 47 differential equations and sets of algebraic equations.

The nonideality of the distillation is performed using UNIFAC for activity coefficient and Peng-Robinson for fugacity coefficient. The system which was studied was a distillation column consisted of 15 trays, a reboiler and a condenser which separate methanol-water binary mixture.

2.2. Input-output selection

It is imperative to select variables that are relevant to the output so that possible model structures can be determined. Reflux rate, reboiler heat input and tray temperatures were selected as the inputs meanwhile top and bottom composition were selected as the output of the neural network model developed. These variables were chosen based on variables analysis of first principle model. Analysis results show that these variables play the prominent role in separation performance and have great effect on product quality.

2.3. Data collection

Different patterns of input variables were generated in order to provide sufficient information to identify the dynamic nonlinear behavior of the distillation column. Several reboiler heat input and reflux flowrate profiles were designated to excite the first principle model and the changes were made simultaneously for these two variables. The profiles considered were random, step change, staircase and pseudo random multilevel sequence. These profiles were selected because they provide a large range of

values. These characteristic is also important since strong input is needed for nonlinear system identification. Seven sets of data with different reflux flowrate and reboiler heat input profiles were generated from the simulation of the general first principle model. The reboiler heat input was varied from 11.25kW to 12.75kW and the reflux flowrate was varied from 0.54kmol/hr to 1.08kmol/hr. These sets of data were divided into three groups: training, validation and testing.

2.4. Neural network model development

Computational programs for designing the neural networks are implemented step by step in MATLAB® to find the suitable model for identification of distillation column. The toolbox, MATLAB®'s pre-defined functions is used to provide tools for design the neural networks model. One hidden layer feedforward neural network model was proposed in this study since two hidden layers rarely improve the model and may introduce a greater risk of converging to local minima. Sigmoid activation functions were used for the hidden and output layers.

The procedure to develop the optimum neural network model started with training process to select the structure of the neural network model. The Levenberg-Marquardt training algorithm is adopted to train the neural networks. The structure was determined by the number of neurons in each layer of the model which was determined by trial and error because this is the most widely used method for optimum topology selection (Miller *et al*., 2005) and the most efficient way has not be established yet. The structure was determined based on the performance of the model to predict behavior based on independent or validation data. The best topology was defined as the one that gives the largest correlation coefficient (R value) and the smallest sum squared error (SSE) value during the validation process. This method is known as the hold out method (Bishop, 1995).

2.5. Input-Output Scheme of the Neural Network Model

In this study, The inputs to the neural network models considered were the past value of the top and bottom compositions data, reflux rate (R) and reboiler heat input (QR) at two time instants, reboiler temperature (TB) and condenser temperature (TD) at three time instant and four tray temperatures; T2, T6, T10 and T14. The past values of the variables were included as model inputs. Using the current and one past (historical) sample of each variables permit for a limited degree of gradient information i.e. are given an idea of the recent rates of changes of each variable, will provide useful in light of dynamic systems, though it is not apparent that this particular choice is optimal (Kershenbaum and Magni, 2001).

The MIMO model can be presented as the following function respectively:

$$\left[xD(t), xB(t)\right] = f\begin{pmatrix} xD(t-1), xB(t-1), R(t), R(t-1), QR(t), QR(t-1), TD(t), TD(t-1), \\ TD(t-2), TB(t), TB(t-1), TB(t-2), T14(t), T10(t), T6(t), T2(t) \end{pmatrix}$$

3. Results and discussions

3.1. Training Performance

Table 1 tabulates the correlation coefficient, the R-value and the sum squared error, SSE between the validation data and the neural network output. These two characteristics

represent the performance of the model and were used in this paper for validation and testing of the MIMO model.

Due to the validation process which involved two sets of data, the performance of the network was determined based on the average of the R-value and the sum squared error for both sets of data. The numbers of hidden neurons considered were 10 to 21 because the neural network did not converge when the number of neurons was reduced to less than ten. Increasing the neurons to more than 21 did not improve learning; instead it increased the total training time.

From the results, it was found that the MIMO model with 18 neurons in the hidden layers was the optimum network architecture since it gave the highest average R-value of 0.983 and the smallest average SSE of 1.72×10^{-4}.

Table 1: R-value and SSE of MIMO model for different numbers of hidden neurons

Architecture	Average R-value	Average SSE
16-10-2	0.982	0.000181
16-11-2	0.945	0.000483
16-13-2	0.971	0.000525
16-14-2	0.912	0.000811
16-15-2	0.938	0.000657
16-16-2	0.982	0.000227
16-17-2	0.935	0.000836
16-18-2	0.983	0.000172
16-19-2	0.972	0.000888
16-20-2	0.975	0.0014
16-21-2	0.904	0.00131

3.2. Testing Results

In order to observe a better network generalization ability, the MIMO networks were tested with testing data which was independent of the training and validation data. This was carried out in order to confirm the performance of the network selected (Bishop, 1995).

The testing results for the MIMO models are summarized in Tables 2. The performance of the model during the testing process for the top and bottom product compositions estimation using the MIMO and models are illustrated in Figures 1. The figure shows the deviation of the current value of the top and bottom product compositions in the particular time instant from the steady state value for the testing data and the neural network output.

Table 2: Testing results for MIMO-1 model

Architecture	Architecture	R-value	SSE ($\times 10^{-4}$)
Top product composition	16-18-2	0.992	8.02
Bottom product composition		0.993	1.00

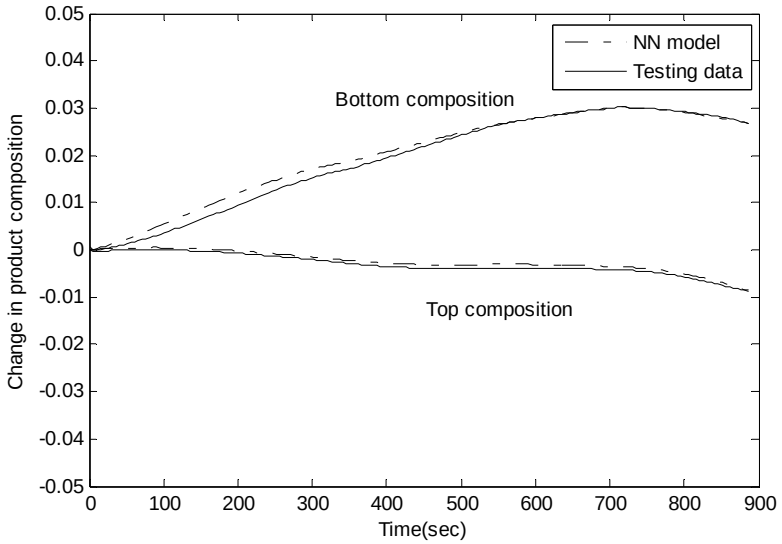


Figure 1 Testing performance of MIMO model

3.3. Comparison with Experimental Data

The comparison of the model with the experimental data was made by observation of changes in products compositions in changes of selected parameters within a certain period of time after the steady state condition as achieved. Figures 2 shows the deviation of the current value of the top and bottom product compositions in the particular time instant from the steady state value for the experimental data and the model output. Based on these figures, the trends in compositions obtained through the simulation of the neural network models were in good agreement with those found by experimental studies even though small deviations were observed. One possible reason for this mismatch is due to the choice of the number of historical data used.

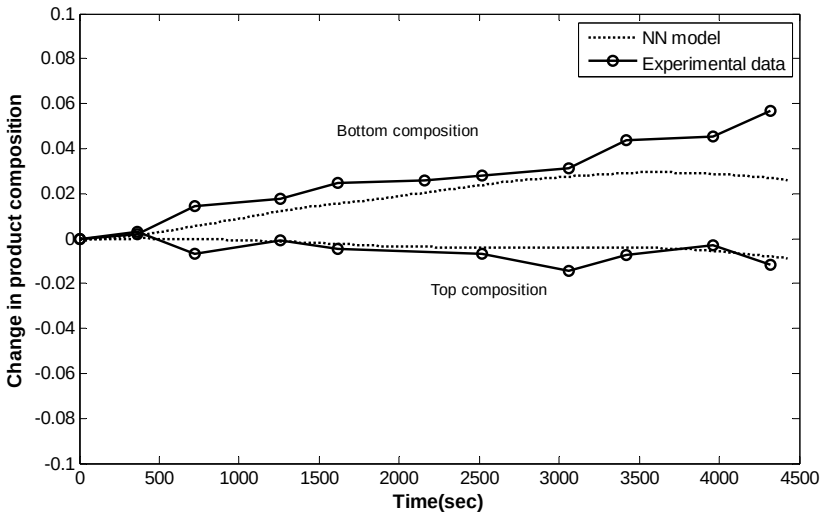


Figure 2 Comparison of the MIMO-1 model with experimental results

4. Conclusion

The MIMO model was found to be able to predict the top and bottom product composition precisely. From the results, it was observed that the model output closely followed the testing data trends. The R-value was 0.992 for the estimation of top product composition and the SSE was 8.02×10^{-4}. For the bottom product composition, the R-value and SSE were 0.993 and 1.00×10^{-4} respectively. The results showed that the MIMO models developed was in good agreement with the experimental data. The results obtained in this study proved that this neural network model could be used to represent the distillation process.

5. Acknowledgement

This research was supported by Ministry of Science, Technology and Innovation (MOSTI), Malaysia through IRPA-8 project No: 03-02-4279EA019.

References

C.M. Bishop, 1995, *Neural Networks for Pattern Recognition*. New York: Oxford University Press.

L.S. Kershenbaum and A.R. Magni, 2001, Training radial basis function networks for process identification with an emphasis on the Bayesian evidence approach. *In*: M. Mujtaba and M. A. Hussain, (Ed). *Application of Neural Network and Other Learning Technologies in Process Engineering*, London: Imperial College Press.

W.L. Luyben, 1987, Derivation of Transfer Functions for Highly Nonlinear Distillation Columns, *Industrial & engineering Chemistry Research.* 26, 2490-2495.

W.L. Luyben, 2002, *Plantwide Dynamic Simulators in Chemical Processing and Control*. New York: Marcel Dekker, Inc.

W.L. Luyben 1989 *Process Modeling, Simulation and Control for Chemical Engineers 2nd Edition*. New York: Mc Graw-Hill.

R.L. Miller, C.E. Del Castillo and B.A. McKee, 2005, *Remote Sensing of Coastal Aquatic Environments: Technologies, Techniques and Applications.* The Netherlands: Springer.

V. Singh, I. Gupta and H.O. Gupta 2005, ANN Based Estimator for Distillation-Inferential Control. *Chemical Engineering and Processing*. 44 (7), 785–795.

V. Singh, I. Gupta and H.O. Gupta 2007, ANN-Based Estimator for Distillation Using Levenberg–Marquardt Approach. *Engineering Applications of Artificial Intelligence*. **20** (2), 249-259.

X. Yu, 2003, A Neuromorphic Controller for a Distillation Column. *In: 4th IEEE International Conference on Control and Automation (ICCA), 2003, Montreal, Canada.*

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Selection of models for pollutant transport in river reaches using case based reasoning

Elisabeta C. Ani[a,b], Yuri Avramenko[a], Andrzej Kraslawski[a], Paul S. Agachi[b]

[a]*Department of Chemical Technology, Lappeenranta University of Technology, P.O. Box 20, FIN-53851, Lappeenranta, Finland. E-mail: Cristina.Ani@lut.fi*
[b]*Faculty of Chemistry and Chemical Engineering, "Babes-Bolyai" University Cluj-Napoca, No. 11, Arany Janos Street, 400028, Cluj-Napoca, Cluj, Romania*

Abstract

The paper presents a case based reasoning (CBR) tool designed for the identification of crucial information for the modelling of pollutant transport in rivers. The developed CBR tool is using a case base comprising published information. This information regards pollutant transport modelling work carried out for different rivers. The models from the case base are designed for a specific river reach. For an unstudied reach it is possible to find a model for pollutant transport based on the comparison between the characteristics of that river reach with the characteristics of the already studied reaches, stored in the case base. The reaches of Somes River were compared against the reaches in the case base, and information for pollutant transport modelling was identified. The developed tool is also applicable in the case of other rivers were no detailed concentration measurements are available and the mathematical models for pollutant transport are needed.

Keywords: pollutant transport modelling, water quality modelling, case-based reasoning, Somes River.

1. Introduction

This work is motivated by environmental and as well economic issues encountered in the everyday environmental management. In this field the water quality assessment is essential, and the experience in solving problems related to river pollution shows that mathematical models for the evaluation of pollutant concentration along river are very important. The specific models for pollutant transport are needed for every studied river. However, the development of such models requires a lot of time and resources.

Our aim is to use the existing knowledge in order to identify the most suitable pollutant transport models for the Romanian Somes River. To achieve this goal the case based reasoning (CBR) approach is being used. The method is based on the idea that the similar problems have the similar solutions. According to the present literature, CBR is being applied in multiple engineering fields in order to find solutions for new problems using the past experience (Avramenko and Kraslawski, 2008; Botar-Jid, 2007).

In the field of water quality modelling Chau (2007) presents a prototype knowledge management system (KMS) useful for the selection of models and parameters for coastal hydraulic and transport processes. This KMS employs the principles of CBR to extract information from a knowledge base comprising 11 scenarios related to the water quality (e.g. flood forecast, water pollution and eutrophication). In the presented paper, the case base is much bigger and the CBR is going to rely on the similarity between the characteristics of river reaches in order to find suitable models for pollutant transport in

Somes River. According to our literature survey no similar CBR tools have been developed yet.

The Somes River, one of the most important rivers in Romania, covers an area inhabited by approximately 2 millions people. The river has to be subjected to an increased attention, due to the accidents which took place, and found the authorities in charge unprepared (Moldovan, 2006). Another reason is the continuous pollution taking place along the river branches due to the increasing industrialization and urbanization (Moldovan, 2006; Serban and Rosu; 2005). Consequently, a good model is needed for the assessment of the water quality. The literature search on Somes River reveals the existence of works aimed at the assessment of the water quality (Moldovan et al., 2007; Laszlo, 2006; Serban and Rosu; 2005; Ponta et al., 2003), but there are no studies focused on the modelling of pollutant transport.

The models for pollutant transport are very complex due to the large number of factors affecting the transport, the variable geometry of the channel along the river, and the multiple phenomena occurring simultaneously: convective-diffusive mass transport, physical processes, chemical and biochemical transformations (Socolofsky and Jirka, 2005). The model has to rely on field data in order to obtain the best estimates of the parameters and a high accuracy of the concentration prediction (Lee and Seo; 2007; Socolofsky and Jirka, 2005). This kind of data is obtained through high frequency monitoring of the concentrations along large river distances. Such approach requires time and resources and sometimes is difficult to carry out. The existing data for the Somes River consists of: (1) pollutant concentration data collected on a monthly basis in just 12 monitoring points along 421km; and (2) measurements of the channel features. This kind of data could be used to test a model, but not to build a reliable one.

Usually, the mathematical models for pollutant transport in the rivers rely on variables capturing the river channel characteristics as well. As the geometry of the channel could be similar from one river reach to another one, a solution to overcome the lack of data is to use the already validated models. For the case of Somes River, or any other case where data are missing, it is possible to implement the old, commonly-accepted models. These models will be identified and selected using CBR.

2. Method description

The identification of the models for pollutant transport in Somes River requires several research steps including (1) the division of the river length into five reaches; (2) building a case base; (3) the development of the CBR tool and (4) the comparison of the stored cases with the reaches of the Somes River in order to (5) identify the most similar cases and (6) retrieve the model description and related information from the sources.

The case base is built using information from the published literature and from existing databases about rivers. There are rivers which have been the subject of water quality modelling work, and can be used as old, solved cases in the present research. The information related to the problem statement and to the assigned solution for each case is stored in the case base. Our interest is focused on the details related to: the mathematical models for pollutant transport; models for the calculation of the transport parameters (water velocity, dispersion coefficients) and models for the transformations suffered by the pollutant during the transport.

Each stored case is characterised by the following variables corresponding to the investigated river reach: channel width, river bed slope, water depth, reach's length, width to depth ratio, mean water velocity, and longitudinal dispersion coefficient. The Somes River reaches are also assigned with values corresponding to these variables.

Further on each river reach from the case base is compared against the reaches of the examined river (in this work – Somes River), in order to find the most similar stored cases for each of the reaches. The similarity value between two river reaches (equation 1) is calculated as a weighted sum of similarity measures corresponding to the parameters of the reaches (a_i, b_i for A and B respectively):

$$SIM(A,B) = \sum_{i=1}^{k} w_i \cdot sim(a_i, b_i) / \sum_{i=1}^{k} w_i \quad (1)$$

where w_i is the weight of importance of parameter i; sim_i is the similarity function for the parameters; k is the number of parameters.
The similarity function for the parameters is based on the difference between the solved case and the new one. The smaller is the difference the bigger the similarity is. The similarity *sim* for two parameters *a* and *b* is defined as:

$$sim(A,B) = 1 - \frac{|a-b|}{range} \quad (2)$$

where *range* is the difference between maximum and minimum values in the overall set of parameters' values.
More details about the CBR approach and the development of the CBR tools are provided in Avramenko and Kraslawski (2008), Avramenko and Kraslawski (2006), and Botar-Jid (2007).
The solutions assigned to the cases with the highest similarity degree are further used in order to formulate the mathematical model for pollutant transport in the Somes River. After solving the problem of pollutant transport in the Somes River the five investigated reaches are going to be added to the case base, and further used in studies related to the modelling of pollutant transport in other rivers.

3. Results and discussions

3.1. The Somes River reaches

The measurements show that Somes River has a non-uniform channel geometry, which is causing the longitudinal variability of characteristic parameters of the pollutant transport and makes the modelling task more difficult. The approach to use in these complex cases, firstly suggested by Fischer and colleagues (1979), is to split the river into reaches, according to the characteristics of each river fragment.

Table 1 Normalized parameters of the reaches of Somes River

Reach number	Reach length [km]	Width	Depth	Slope	Width/ Depth	Flow rate -average-	Flow rate –min–	Flow rate -max-
1	90	0.16	0.19	1.00	0.82	0.11	0.138	0.053
2	80	0.42	0.51	0.18	0.82	0.19	0.290	0.118
3	8	0.58	0.83	0.09	0.70	0.43	0.489	0.401
4	59	0.78	1.00	0.10	0.78	0.70	0.715	0.723
5	113	1.00	1.00	0.07	1.00	1.00	1.000	1.000

Somes was preliminarily divided into five reaches of unequal length, as presented in Table 1. The CBR approach is used to find the suitable models for each reach.

3.2. The case base

The case base contains 147 cases regarding studies on the convective-diffusive pollutant transport. One river could be represented as different cases if it consists of multiple reaches studied on different occasions.

Most of the stored cases (90 cases) provide information about the estimation of dispersion coefficients. This knowledge can be used to formulate the models for parameter estimation which is included in the pollutant transport models. Other 26 cases are related to travel time studies, the velocity depending on the river morphology or tracer experiments. The most important stored cases (31 cases) are presenting details about pollutant transport modelling, including employed equations, initial and boundary conditions, pollutant transformation models and/or parameter estimation models.

The case base is implemented as Microsoft Excel workbook.

3.3. The CBR tool

The CBR tool is a single computer application with supplemented data files (e.g. the case base, the literature sources).

The tool is able to perform the following tasks: (1) to import the data from Excel documents or set of text files with parameters of the river reaches; (2) to compare the input (or import) parameters of a specific river reach against those in the case base; (3) to rank the cases according to the value of similarity to the case under consideration; (4) present the most similar cases; and (4) retrieve information from the data source (usually a PDF file).

3.4. The most similar cases and utility for the Somes River reaches

The CBR tool ranked all stored cases according to the similarity to the Somes River reaches. For each reach the three cases assigned with the highest similarity value are listed in Table 2. The first column contains the name of the river reach (as it was stored in the case base), the second column contains a synthesis of the available information, then the level of similarity is listed, and in the last column is mentioned at least one of the used references.

The results reveal that the models developed for the Swale, Rhine and Ebro rivers are the most appropriate to use in the modelling of the pollutant transport in Somes River. By analysing the similarity level two approaches could be proposed for the models allocation to the Somes River reaches. In the first approach the existing model for Swale 3 is used for reach #1; Swale 4 for reach #2; Rhine 4 for reach #3 and Ebro 2 for reaches #4 and #5. In the second approach the model for Swale 3 is used for reach #1, Rhine 4 for reaches #2 and #3 and Ebro 2 for reaches #4 and #5.

Since the order of the Somes reaches is in relation to distance from the spring it could be stated that the model from case Swale 3 is suitable to model the beginning of Somes River (reach #1), model from case Rhine 4 could be applied for the upper middle part of the river (reaches #2 and #3), and for the downstream end (reaches #4 and #5) the model from case Ebro 2 can be used.

Table 2 The cases with the highest similarity degree corresponding to the Somes River reaches

Case name	Available information	Similarity value (%)	Example of references
Reach #1			
Swale 3	QUESTOR, LOIS, in-stream water quality modelling Intensive monitoring campaigns	94.63	House and Warwick (1998) Eatherall et. al. (1998)
Cowaselon Creek	tracer experiments dispersion coefficient estimation reaeration coefficients estimation	93.65	Socolofsky and Jirka (2005) Kilpatrick et al. (1989)
Narew 3	tracer experiments dispersion coefficient estimation Rhodamine WT transport model	93.24	Rowinski et al (2007)
Reach #2			
Rhine 4	model for pollutant transport alarm system for accidents	95.11	Dunnivant and Anders (2005) Broer (1991)
Waikato 1	nutrients transport dispersion coefficient surface heat transfer	93.04	Sanchez-Cabeza and Pujol (1999) Rutherford and Williams (1992) Alexander et. al. (2002)
Swale 4	QUESTOR, LOIS, in-stream water quality modelling Intensive monitoring campaigns	91.84	House and Warwick (1998) Eatherall et. al. (1998)
Reach #3			
Waikato 1	see above	95.65	see above
Rhine 4	see above	93.08	see above
Ebro 2	tracer experiments hydrodynamics study two pollutant transport models	91.56	Mestres et al. (2003) Sanchez-Cabeza and Pujol (1999) Pujol and Sanchez-Cabeza (2000)
Reach #4			
Ebro 2	see above	95.31	see above
Waikato 2	nutrients transport dispersion coefficient surface heat transfer	94.02	Sanchez-Cabeza and Pujol (1999) Rutherford and Williams (1992) Alexander et. al. (2002)
Swale 4	see above	84.04	see above
Reach #5			
Ebro 2	see above	95.31	see above
Waikato 2	see above	94.02	see above
Rhine 4	see above	89.37	see above

4. Conclusions

The paper presents the development and application of a CBR tool for the identification of pollutant transport models for the reaches of Somes River. The results reveal that the prior developed models for the reaches of the rivers Swale, Rhine and Ebro could be applied to the reaches of the Somes River in further research studies. The developed CBR tool has general applicability with respect to river reaches, as it could be used to propose pollutant transport models also for other rivers than Somes.

References

R.B. Alexander, A.H. Elliott, U. Shankar, G.B. Mcbride, 2002, Estimating the source and transport of nutrients in the Waikato River Basin, New Zealand. Water Resources Research, 38, 1268-1290.

Y. Avramenko, A. Kraslawski, 2008, Case Based Design. Applications in Process Engineering, 51-108.

Y. Avramenko, A. Kraslawski, 2006, Similarity concept for case-based design in process engineering, Computers and Chemical Engineering, 30, 548–557.

G.J.A.A. Broer, 1991, Alarm system for accidental pollution on the river Rhine, Hydrology for the Water Management of Large River Basins, 201, 329 – 336.

C.C., Botar-Jid, 2007, Selective catalytic reduction of nitrogen oxides with ammonia in forced unsteady state reactors. Case based and mathematical model simulation reasoning. Acta Universitatis Lappeenrantaensis, 283, 1-184.

K.W. Chau, 2007, An ontology-based knowledge management system for flow and water quality modelling. Advances in Engineering Software, 38, 172–181.

F.M. Dunnivant, E. Anders, 2005, Fate and transport of pollutants in rivers and streams, A basic introduction to pollutant fate and transport. An integrated approach with chemistry, modeling, risk assessment, and environmental legislation, 187 – 209.

A. Eatherall, D.B. Boorman, R.J. Williams, R. Kowe, (1998). Modelling in-stream water quality in LOIS. The Science of the Total Environment, 210/211, 499-517.

H.B. Fischer, E. Liet, C. Koh, J. Imberger, N. Brooks, 1979, Chapters 3 and 5, Mixing in Inland and Coastal Waters.

W.A. House, M.S. Warwick, 1998, Intensive measurements of nutrient dynamics in the River Swale, The Science of the Total Environment, 210/211, 111-137.

F.A. Kilpatrick, R.E. Rathbun, N. Yotsukura, G.W. Parker, L.L. Delong ,1989, Determination of stream reaeration coefficients by use of tracers. In: U.S.G.S., Techniques of Water-Resources Investigations, Book 3, Chapter A18, 1-61.

F. Laszlo, 2006, Lessons learned from the cyanide and heavy metal accidental water pollution in the Tisa River basin in the year 2000. In: G. Dura, F. Simeonova (Eds), Management of international and accidental water pollution, 43-50.

M.E. Lee, I.W. Seo, 2007, Analysis of pollutant transport in the Han River with tidal current using a 2D finite element model, Journal of Hydro-Environment Research, 1, 30–42.

M. Mestres, J.P. Sierra, A. Sanchez-Arcilla, J. Gonzalez del Rio, T. Wolf, A. Rodriguez, S. Ouillon, 2003, Modelling of the Ebro River plume. Validation with field observations, Scientia Marina, 67, 4, 379 – 391.

Z. Moldovan, G. Schmutzer, F. Tusa, R. Calin, A.C. Alder, 2007, An overview of pharmaceuticals and personal care products contamination along the river Somes watershed, Romania, Journal of Environmental Monitoring, 9(9), 986-993.

Ponta, M., Frentiu, T., McCormick C., Bela, A., Roman, C., Cordos, E. , 2003, An investigation into the Cu, Pb and Zn pollution of the Somes and Lapus rivers of Baia-Mare area, Romania. Studia Chemia, 1, 16.

Ll. Pujol, J.A. Sanchez-Cabeza, 2000, Use of tritium to predict soluble pollutants transport in Ebro River waters (Spain), Environmental Pollution, 108, 257-269.

P.M. Rowinski, I. Guymer, A. Bielonko, J.J. Napiorkowski, J. Pearson, A. Piotrowski, 2007, Large scale tracer study of mixing in a natural lowland river, Proceedings of the 32nd IAHR Congress, paper 297.

J.C. Rutherford, B.L. Williams, 1992, Transverse mixing and surface heat exchange in the Waikato River: a comparison of two models, New Zealand Journal of Marine and Freshwater Research, 26, 435-452.

J.A. Sanchez-Cabeza, L. Pujol, 1999, Study on the hydrodynamics of the Ebro river lower course using tritium as a radiotracer, Water Research, 33, 10, 2345-2356.

P. Serban, I. Rosu, 2005, Management plan of the Somes-Tisa hydrographical area. 2004 Report.

S.A. Socolofsky, G.H. Jirka, 2005, Special Topics in Mixing and Transport Processes in the Environment. Engineering – Lectures. 5th Edition. Texas A&M University, 1-93, 125-193.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

MINLP synthesis of reactive distillation using a disjunctive, hybrid model

Marcel Ropotar[a], Zorka Novak Pintarič[a], Jean-Michel Reneaume[b], Zdravko Kravanja[a]

[a]*University of Maribor, Faculty of Chemistry and Chemical Engineering, Smetanova 17, 2000 Maribor, Slovenia*
[b]*Université de Pau et des Pays de l'Adour, ENSGTI, LaTEP Rue Jules Ferry, BP 7511, 64075 Pau Cedex, France*

Abstract

In order to develop optimization models for process industries, detailed hybrid modelling providing accurate descriptions of processes is required, where the simplest part of the model is represented explicitly in an equation-oriented environment whilst detailed calculations, e.g. enthalpies, kinetics, equilibrium and efficiencies, are programmed by external implicit functions. The two main objectives of our work have been: i) the development of such hybrid modelling environment, and ii) implementation of efficient solution strategies in an MINLP process synthesizer MIPSYN, the successor of PROSYN (Kravanja and Grossmann, 1994), which could make solving such complex industrial problems possible. The present study focuses on the catalytic distillation processes. Since the use of a multilevel MINLP approach reduces the effects of nonconvexities, better solutions were obtained by multilevel than by single level MINLP strategies.

Keywords: Reactive distillation, NLP, MINLP, MIPSYN.

1. Introduction

Catalytic Distillation (CD) is a multifunctional reactor concept which combines distillative separation with a chemical reaction catalyzed by a solid. The combination of distillation and chemical reaction may result in some advantages over the conventional process (reactor + distillation column), namely, investment cost reduction, higher conversion, improved selectivity, or heat integration.

The MINLP optimization of a CD column has already been studied by several authors (e.g. Ciric and Gu, 1994; Frey and Stichlmair, 2000; Cardoso et al., 2000; Poth et al., 2001; Jackson and Grossmann, 2001; Sand et al., 2004). Due to problem complexity, simplified models have mostly been used in the optimization so far, which might compromise the true optimality of the solutions. Only few attempts have been made in order to use more rigorous models. In 2003, Poth et al. proposed an extension of their previous works: the MINLP problem is solved using the general algebraic modelling systems (GAMS) in combination with external functions. Gomez et al. (2006) have proposed an optimization strategy to solve the Catalytic Distillation problem using a rate-based model and stochastic optimization algorithms.

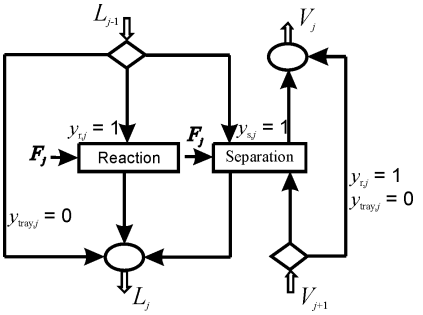
Although GAMS, as MIPSYN's interface, provides a powerful language for manipulating data and defining highly structured collections of variables and equations, there are situations when some parts of the model should be defined using FORTRAN,

C++ or some other language as user defined routines. This could be favourable because the size of the original optimization model could be significantly reduced if extensive calculations were accomplished outside the optimization model. Thus, the objective of the research has been to take advantage using such hybrid models through different advanced capabilities of MIPSYN. For example, MIPSYN allows the user to interfere and assist the optimization subiterations during the main iterations of the optimization algorithm, which significantly improves the robustness and the solution of the MINLP search. The hybrid distillation model consists of the equation-oriented part, defined by classical MESH equations, and the implicit part where rigorous sub-models are used to describe the thermodynamics, kinetics and hydrodynamics behaviour of the system, which are written in the external FORTRAN procedures, connected to MIPSYN.

2. Modelling of the tray superstructure of reactive distillation

The MINLP model of the reactive distillation column is based on the tray superstructure shown in Figure 1, where binary variables $y_{\text{tray},j}$, $y_{\text{r},j}$ and $y_{\text{s},j}$ are assigned to a tray j, a reaction stage j and a separation stage j, respectively. At every tray j, $i \in J$, three exclusive alternatives are available: a reaction stage ($y_{\text{r},j} = 1$), a separation stage ($y_{\text{s},j} = 1$), and a by-pass, ($y_{\text{tray},j} = 0$), which are combined with the following logical relation:

$$y_{\text{r},j} + y_{\text{s},j} = y_{\text{tray},j} \tag{1}$$



L ... liquid flowrate
V ... vapour flowrate
F ... feed flowrate

Figure 1: Superstructure for tray.

Many other logical constraints have to be applied as well in order to model such a tray. When separation is not selected, molar fractions for components i, $i \in C$, at a tray j in a vapour flow are the same as the ones below the tray:

$$x^{\text{v}}_{i,j} \le x^{\text{v}}_{i,j+1} + y_{\text{s},j} \quad \text{and} \quad x^{\text{v}}_{i,j} \ge x^{\text{v}}_{i,j+1} - y_{\text{s},j} \tag{2}$$

and similar constraints are given for tray temperatures, enthalpies and flow-rates. The same type of inequalities is applied also for a liquid flow when a by-pass is selected. In addition, the following constraints are applied when the reaction tray is selected:

$$r_{i.j} \le r^{\text{UP}}_{i,j}\, y_{\text{r},j} \quad \text{and} \quad r_{i.j} \ge r^{\text{LO}}_{i,j}\, y_{\text{r},j} \tag{3}$$

$$m_{\text{cat},j} \le m^{\text{UP}}_{\text{cat},j}\, y_{\text{r},j} \tag{4}$$

where r_{ij} and $m_{\text{cat},j}$ denote the reaction rate for a component i and the mass of the catalyst at tray j, respectively. There are also other logical equations, which prevent the structural multiplicity of the solutions:

$$y_{\text{tray},j} \ge y_{\text{tray},j+1} \tag{5}$$

and possibly the selection of two successive reaction trays:

$$y_{r,j} + y_{\mathrm{r},j+1} \le 1 . \tag{6}$$

The MESH equations for reaction trays are different from the ones for separation trays because there is no vapor-liquid interface, and the vapour characteristics of the tray are the same as below the tray. Therefore, some terms in the MESH equations are subtracted and, furthermore, the mass of catalyst has to be considered, e.g. as in the partial material balance:

$$L_{j-1}x_{i,j-1} - L_j x_{i,j} + m_{\mathrm{cat},j} r_{i,j} + F^{l}_{i,j} = 0 \tag{7}$$

An objective function, given below, is a cost function taken from Gomez et al. (2006) where C_i is the cost for raw materials, C_{cat} for the catalyst, C_{reb} and C_{con} for the utilities, and C_{T} and C_{SH} are the investment costs for the trays and shell, and C_{prd} for a product, respectively:

$$\min \left(\begin{array}{l} C_0 + \sum_i C_i \sum_{j=\mathrm{feed}} x_{i,j} L_j + A_{\mathrm{F}} \cdot C_{\mathrm{cat}} \cdot m_{\mathrm{cat}} + C_{\mathrm{reb}} Q_{\mathrm{reb}} + C_{\mathrm{con}} Q_{\mathrm{con}} + C_{\mathrm{T}} D^{1.55} \cdot \\ N\left(h_{\mathrm{tray}} + 1.27 \frac{W}{D^2} \right) + C_{\mathrm{SH}} D \left(h_0 + N\left(h_{\mathrm{tray}} + 1.27 \frac{W}{D^2} \right) \right)^{0.802} - C_{\mathrm{prd}} \cdot L_{n+1} \end{array} \right) \tag{8}$$

3. Connection between MIPSYN and FORTRAN

As rigorous sub-models are too complex to be represented explicitly, they are written in FORTRAN as external implicit functions. On the other hand, the conventional MESH equations are modelled in the GAMS language and represent the explicit part of the resulting hybrid model. In this way, constraints in the problem (MINLP) are decomposed into equation-oriented (EO) constraints used in GAMS, e.g. ($h_u(x_u, x_v, y) = 0$), and constraints represented in the external program, e.g. ($h_v(x_v, x_u, y) = 0$), where the suffix u is used for EO and v for the external part of the model. Similarly, the vector of continuous variables $\boldsymbol{x}$ is decomposed into $\boldsymbol{x}_u$ and $\boldsymbol{x}_v$, and the dimensionality n into n_u and n_v. When $\boldsymbol{x}_v$ is substituted by implicit functions $\Phi_v(x_u, y)$ through external procedures, the hybrid (MINLP-H) problem is defined.

$$\begin{array}{ll} \min Z = c^{\mathrm{T}} y + f(x_u, x_v) & \\ \text{s.t.} & \text{(MINLP)} \\ \quad h_u(x_u, x_v, y) = 0 & \\ \quad h_v(x_u, x_v, y) = 0 & \\ \quad g_u(x_u, x_v, y) \le 0 & \\ \quad g_v(x_u, x_v, y) \le 0 & \\ \quad x = [x_u, x_v] \in X \in R^n & \\ \quad X = X_v \cup X_u & \\ \quad x_u \in X_u \in R^{n_u}, x_v \in X_v \in R^{n_v} & \\ \quad y \in [0,1]^m & \end{array}$$

$$\begin{array}{ll} \min Z = c^{\mathrm{T}} y + f[x_u, \Phi_v(x_u, y)] & \\ \text{s.t.} & \text{(MINLP-H)} \\ \quad h_u[x_u, \Phi_v(x_u, y), y] = 0 & \\ \quad g_u[x_u, \Phi_v(x_u, y), y] \le 0 & \\ \quad x_u \in X_u \in R^{n-n_v} & \\ \quad y \in [0,1]^m & \end{array}$$

4. Strategies for MINLP synthesis of reactive distillation in MIPSYN

In our earlier attempts to solve the hybrid model for reactive distillation we were able to obtain feasible solutions only for NLPs at fixed structures with a moderate number of

trays (up to 20) applying the GAMS solver CONOPT3. Due to a variety of severe numeric problems, we could not solved the entire problem (MINLP-H), using the solver DICOPT.

In order to overcome the drawbacks, MIPSYN's outer-approximation (OA) algorithm was upgraded for the use of external routines in combination with disjunctive programming and convex hull formulation of the problem where subproblems $(NLP)^k$ are solved only for selected trays and not for the entire superstructure. This substantially reduces the effects of nonconvexities and singularities. In addition, since MIPSYN allows the user to interfere and assist the MINLP optimization, the robustness and hence chances of achieving the convergence of the MINLP search significantly increase. If the desired product purity increases, more trays have to be included in the model and the chances of obtaining the optimal solutions significantly decrease. Powerful strategies have to be implemented in MIPSYN in order to facilitate the solution procedure. They consist of different options that can easily be accommodated in the synthesizer MIPSYN:

- Interactive and automated mode of operation where the interactive one is perhaps the most important option since it enables the user to dynamically adjust the course of the optimization.
- A capability to perform the mentioned disjunctive programming with execution of only the "existing" part of the model.
- The use of different extensions of the OA algorithm with a possibility to use convexity test and the validation of the outer approximation, in order to reduce the size of the MILP master problem and reduce the undesirable effects of the nonconvexities.
- The possibility to solve feasibility problems with an augmented penalty function in the NLP and MILP master problems.
- Using different NLP and MILP solvers in sequence with different option files for the better tuning of solvers' parameters.
- Different multilevel MINLP can be implemented in the synthesizer, e.g. in our problem it consists of the execution of several sequences of MINLP iterations, each level restarting from an optimal solution obtained at the previous level.
- Additional constrained integer-cuts can also be imposed in order to limit the search to a narrower space of topological solutions.

5. Illustrative example

The illustrative example presented in this contribution is the cost minimization of ETBE production as main product. A minimum 95% molar purity of the ETBE is required. The significance of this reaction resides in the worldwide consumption of ETBE. We only consider the main reaction:

$$(CH_2)C{=}CH_2 + C_2H_5OH \leftrightarrow (CH_3)_3COC_2H_5 \tag{9}$$

The Vapour/Liquid equilibrium constant of component *i* at stage *j* $K_{i,j}$, present in the MESH equation is evaluated by an external procedure, as a function of the optimization variables (temperature, vapor and liquid composition, at a given total pressure, *P*).

$$K_{i,j} = \frac{\gamma_{i,j} P_i^{\text{sat}} \phi_i^{\text{sat}}}{\hat{\phi}_{i,j} P} \tag{10}$$

The vapour pressure (P_i^{sat}) is estimated using the Antoine equation. The activity coefficient ($\gamma_{i,j}$) is calculated using the modified UNIFAC (Dortmund) model. The

Soave, Redlich and Kwong equation of state is used in order to compute the fugacity coefficients of pure component *i* at vapor pressure (ϕ_i^{sat}) and the fugacity coefficients of the vapor mixture ($\hat{\phi}_{i,j}$). The effect of pressure on the liquid mixture is neglected.

The liquid and vapour molar enthalpies are also evaluated using external FORTRAN procedures. Specific heat capacities are given by the polynomial expression as a function of the temperature. The excess enthalpy is assumed to be zero. The reaction term arising in the MESH equations (mass balances) is a pseudo homogeneous kinetics based on the model proposed by Langmuir, Hinshelwood, Hougen and Watson (Jensen and Datta, 1995). The liquid phase is strongly nonideal and has to be taken into consideration by using activities instead of molar fractions.

The minimum column diameter recommended to avoid entrainment flooding is calculated using the correlations of Kister and Hass (Perry and Green, 1997; Gomez et al., 2006).

6. Results and discussion

The objective was to minimize the annual cost and identify an optimal column configuration: optimal location of two feeds, optimal number and location of reaction and separation trays, and optimal values of operating and design parameters. Three different strategies were applied: a 1-level MINLP with multiple restarts, a multiple level MINLP, and a multiple level MINLP with constrained integer cuts. It should be noted that in both multilevel MINLPs each level restarts from an optimal solution obtained at the previous level. The procedure continues to the next level until there is no better solution obtained at the current level. Table 1 shows the results. The optimal solution of the 1-level MINLP (8.970 k$/year) was taken as a starting point at the 2nd level MINLP. In both multilevel cases this gave rise to better solutions with 8.809 and 8.571 k$/year, respectively. Since the solutions at the 3rd level were not improved, the procedure was stopped. Thus, the best solution with 8.571 k$/year was obtained by the third strategy where constrained integer-cuts were applied in the multilevel MINLP. It should be noted that without the initialization of each NLP subproblem any feasible solution could hardly ever be obtained.

Table 1: Solution for three different strategies.

Process parameters	1-level MINLP with multiple restarts	Multiple level MINLP (2nd level)	Multiple level MINLP with constrained integer-cuts (2nd level)
Position of the feeds	8, 36	8, 37	10, 37
Position of reaction trays	3, 5, 7, 9, 11, 13, 15, 18, 36	2, 4, 6, 10, 14, 23, 25, 32, 38, 40	3, 5, 7, 10, 12, 14, 16, 21, 34, 37, 39, 41
Number of separation trays	37	36	35
Flow of distillate, mol/s	0.0648	0.0646	0.0642
Flow of product, mol/s	0.0281	0.0282	0.0284
Reboiler duty, W	4 024	3 687	3 377
Condenser duty, W	4 230	3 895	3 586
Isobutylene conversion, %	99.36	99.44	99.71
Annual cost, k$/year	8.926	8.809	8.571

In all three cases different trade-offs were obtained than the one by Gomez et al. (2006) using the same equilibrium model in stochastic optimization (annual cost of 8.970 k$/year, conversion of 99.77 %, reboiler duty of 6473 W, 44 separation and 11 reaction trays). It is interesting to note that although the conversions are now somewhat smaller, the annual costs are lower because of significant reduction in the reboiler duties and the number of separation trays (up to 48 % and 20 %). Note also that a hybrid model of distillation collumn with approximately 40 separation and 10 reactive trays typically comprises up to 3000 constraints and 1500 variables, one tenth of them being binary. About 500 equations and almost all variables were part of external procedures.

7. Conclusions

One of the main ideas of our research was to upgrade the process synthesis package MIPSYN for the use of implicit models using GAMS's external functions capabilities, since the convergence was not achieved when we tried to solve the complex hybrid reactive distillation model by DICOPT. To achieve sufficient purity of the final product in the MINLP synthesis of reactive distillation more trays had to be added to the superstructure. Consequently, the probability for obtaining an optimal solution significantly decreased. In order to solve problems with high product purities and hence a large number of trays, the main objective of our recent work has been to develop strategies by which it is now possible to solve such complex MINLP problems within the MINLP synthesizer MIPSYN. Since MIPSYN allows an interactive mode of optimization, different strategies have been implemented in the synthesizer. With the combination of multiple restarts, the convexity test at the first level of multilevel MINLP, and additional constrained integer-cuts without convexity test at the second level, the best solution was obtained.

References

M. F. Cardoso, R. L. Salcedo, S. Feyo de Azevedo, D. Barbosa, 2000, Optimization of Reactive Distillation Processes with Simulated Annealing, Chem. Eng. Sci., 55, 5059- 5078.

A. R. Ciric, D. Gu, 1994, Synthesis of Non Equilibrium Reactive Distillation Processes by MINLP Optimization, AIChE J., 40 (9), 1479-1487.

T. Frey, J. Stichlmair, 2000, MINLP Optimization of Reactive Distillation Columns, Eur. Symp. Comput. Aided Process Eng.-10, 115.

M. Gomez, J.M. Reneaume, M. Roques, M. Meyer, X. Meyer, 2006, A Mixed Integer Nonlinear Programming Formulation for Optimal Design of Catalytic Distillation Column Based on a Generic Non Equilibrium Model, Ind. Eng. Chem. Res., 45, 1373-1388

J. K. Jackson, I. E. Grossmann, 2001, A Disjunctive Programming Approach for the Optimal Design of Reactive Distillation Columns, Comput. Chem. Eng., 25, 1661-1673.

K. L. Jensen, R. Datta, 1995, Ethers from Ethanol. 1. Equilibrium. Thermodynamics Analysis of the Liquid-Phase Ethyl *tert*-Butyl Ether Reaction, Ind. Eng. Chem. Res., 34, 392-399.

Z. Kravanja and I.E. Grossmann, 1994, New Developments and Capabilities in PROSYN – an Automated Topology and Parameter Synthesizer, Comput. Chem. Eng., 18, 1097-1114.

H. Perry, D. Green, 1997, Perry's Chemical Engineers' Handbook; McGraw Hill: New York.

N. Poth, Th. Frey, J. Stichlmair, 2001, MINLP Optimization of Kinetically Controlled Reactive Distillation Processes, 34th Eur. Symp. Work. Party Comput. Aided Process Eng.-11, Supplementary Proceeding Volume, 79.

N. Poth, D. Brusis, J. Stichlmair, 2003, Rigorous Optimization of Reactive Distillation in GAMS with the Use of External Functions. Eur. Symp. Comput. Aided Process Eng.-13, 869.

G. Sand, S. Barkmann, S. Engell, G. Schembecker, 2004, Structuring of Reactive Distillation Columns for Non-Ideal Mixtures Using MINLP Techniques. Eur. Symp. Comput. Aided Process Eng.-14, 493.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Exergoeconomic Analysis Applied in the Optimization of Refrigeration System in Ethylene and Propylene Production Process

Francine M. Fábrega[a], Jakeline S. Rossi[a] and José V. H. d'Angelo[a]

[a]*DESQ - Department of Chemical Systems Engineering – University of Campinas, UNICAMP, Chemical Engineering School, Campinas-SP, CP 6066, 13083-970, Brazil*

Abstract

The expectation of economic growth for the next years concerns the industrial sectors with a possible shortage in the supply of electric energy. The inevitable exhaustion of fossil fuels has also been a constant source of concern. The use of fuels that emit large amounts of CO_2 and CO is becoming increasingly environmentally unacceptable. All these reasons led the most developed countries to optimize their production processes in order to reduce energy consumption and one way to do this is to perform an exergetic and exergoeconomic analyses of the processes. In this context, for ethylene and propylene production process, exergoeconomic costs were determined based on Exergetic Cost Theory from Lozano and Valero making use of software MatLab version 7 to solve the matrices of this method.

Keywords: exergoeconomic, refrigeration, ethylene, propylene

1. Introduction

The increasing consumption of energy to support the growth of chemical industries demands that processes must be operated in an optimized, making rational use of energy. Some actions in this direction are accessible and may be technical and economical feasible, leading to energy rationalization programs that with low costs when they are properly planned.

In Brazil, according to the National Energetic Balance of 2008 (MME, 2009), the chemical sector was responsible for 25.7% of the total industrial electrical energy consumed, which is equivalent to 4.8% of the total electrical energy consumed in the entire country.

In thermodynamics, the exergy of a system is defined as the maximum work possible during a process that brings the system into equilibrium with its surroundings. So, the exergy is the potential of a system to cause a change as it achieves equilibrium with its environment. Exergy is then the energy that is available to be used. After the system and surroundings reach equilibrium, its exergy is zero (Kotas, 1995).

The exergoeconomic analysis is defined as the union of exergetic analysis with economic concepts, which the main objective is to assign a cost to the exergetic content of a system or a stream of the system and when applied, this analysis provides a function of appropriate cost for that system (Valero *et al*., 1996).

The exergoeconomic analysis can be performed by applying the Theory of the Exergetic Cost (Lozano and Valero, 1993) which enables the determination of costs associated with the destruction (dissipation) of exergy, to evaluate the costs of production of each

product in order to: make possible studies of optimization and improvement of the process; help in decisions about actions for the operation, maintenance and resources allocation; compare alternatives techniques and exergoeconomic evaluation of each component.
In order to reduce the energy consumption of the process, exergoeconomic analysis has been used in many works (d'Accadia and Rossi, 1998a; Sahoo, 2008) and with this analysis it is possible to have a reference how the cost of exergy increases as the quality of energy is deteriorated during the process.
In this context, the exergoeconomic costs were calculated to the Refrigeration Systems in the Ethylene and Propylene Production Process.

2. Thermoeconomic analysis and optimization metodology

2.1. Refrigeration Systems in the Ethylene and Propylene Production Process

The simulation of the ethylene and propylene production process in this work was performed based in the production process of Quattor petrochemicals (former Petroquímica União S.A. – São Paulo). The simulation was done with the objective to obtain the thermal charges in the condensers of the distillation columns, which were validated comparing with industrial data and so they were used to simulate the refrigeration cycles required to provide cold utilities to the process. This simulation was performed using commercial simulator HYSYS© version 2.2 from Aspen Technology and the fluid package used in the determination of thermodynamic properties was Peng-Robinson equation of state, which is appropriated for systems consisting of hydrocarbons.
The simulated refrigeration cycles of each column were based in the Carnot cycle. Their configurations are detailed in Rossi *et al.* (2007) and are described below:

- in the demethanizer column three integrated cycles were used operating with the following refrigerants: methane, ethylene and propylene;
- in both deethanizer column and ethylene-ethane splitter, one cycle with propylene as refrigerant was used;
- in the propylene-propane splitter one cycle was used with ethane as refrigerant.

It is important to comment that in the simulation of these cycles were used the pressures given by Table 1.

Table 01 – Operational conditions of the refrigeration cycles in the Ethylene and Propylene Production Process.

Column		High Pressure (kPa)	Low Pressure (kPa)
Demethanizer	Cycle 1	2250	135
	Cycle 2	2000	50
	Cycle 3	1300	150
Deetanizer		1500	150
Ethylene-Ethane Splitter		1300	240
Propylene-Propane Splitter		4000	800

The values of the variables in Table 01 were obtained after adjusting the simulations of the refrigeration cycles and they were combined in order to perform exergetic analysis of all systems, after this the Theory of the Exergetic Cost proposed by Lozano and Valero (1993) was used to estimate exergoeconomic costs of the streams, using software MatLab version 7.1 to solve the matrix built from the equations of this theory. The reference state used in this work is T_o = 298.15 K e P_o = 101.325 kPa.

2.2. Calculation of exergoeconomic costs

In order to obtain the exergonomic costs of the process streams a model for the exergetic and exergoeconomic balance was developed, based on the Carnot cycle proposed by d'Accadia and Rossi (1998). This balance can be represented by Figure 1 and it is used to obtain the exergetic and exergoeconomic costs.

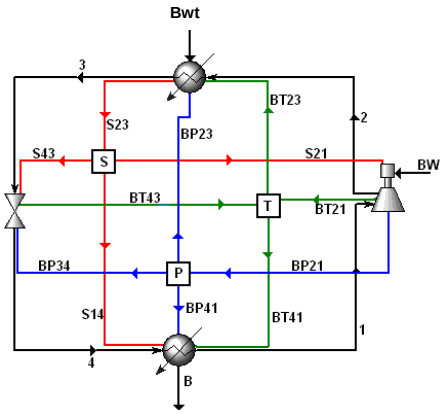


Figure 1 – Exergetic and exergoeconomic balance of a Carnot cycle.

As shown in Figure 1, there are three fictitious points that represent each of the thermodynamic variables in this study. There is an increase in the entropy of the refrigerant stream when it passes through the compressor, the expansion valve and evaporator. In the condenser it is necessary to reduce the entropy of the refrigerant. This is the reason for exergetic flow out of the condenser which is distributed to the others components of the cycle (red lines of Figure 1). Considering the exergetic flow referring to pressure (blue lines), it leaves the compressor and splits into three streams that represent the pressure drop in the heat exchangers and in the expansion valve. Finally, the exergetic flow related to temperature (green lines) leaves the compressor and the expansion valve and goes to heat exchangers.

In the balance showed in Figure 1 BT is the physical exergy of the thermal component which means that it is calculated using only the reference temperature (T_o) for each stream. BP is the physical exergy of the pressure component and it is calculated using only the reference pressure (P_o) for each stream. S is a component of the exergy related to the irreversibilities, and it is obtained by multiplying the entropy variation by the reference temperature. BW is the component realated to the work, Bwt is the component related to the exergy variation of the heat flow of condenser (cold water stream) and B is the exergy variation that the refrigeration cycle, or in order words, the final stream-product of the cycle.

Using this model and the balance equations proposed by Lozano and Valero (1993), d'Accadia and Rossi (1998) have developed a methodology to determine the exergonomic cost of each stream represented in Figure 1. The same procedure was used in this work, quantifying investment costs, operation and maintenance related to each one of the equipments present in the refrigeration cycles of the ethylene and propylene process, which are: expansion valves, compressors and heat exchangers, and also the cost of refrigeration water and electricity.

The values for the investment costs were defined according to the equations proposed by d'Accadia and Rossi (1998):

$$C_{comp} = 12000 \times \left(\frac{BP21 + BT21}{100} \right) \times \left(\frac{\eta}{0,9 - \eta} \right)^{\eta} \quad (\$) \quad (1)$$

$$C_{cond} = 450 \times \left(\frac{-\ln(1 - \varepsilon_{co}) \times S23}{U_{co} \times A_{co}} \right) \quad (\$) \quad (2)$$

$$C_{evap} = 1140 \times \left(\frac{-\ln(1 - \varepsilon_{ev}) \times B}{U_{ev} \times A_{ev}} \right) \quad (\$) \quad (3)$$

$$C_{valv} = 37 \times BT43 \quad (\$) \quad (4)$$

In Equation (1), η is the efficiency of the compressor chosen as 80 %. In Equation (2), ε_{co} is the condenser effectiveness used as 70 %, U_{co} is the unit thermal condutance, adopted as being 0.15 kW/(m^2·K) and A_{co} is the reference area, adopted as 100 m^2. In Equation (3), ε_{ev} is the evaporator effectiveness, adopted as 70 %, U_{ev} is the unit thermal condutance, adopted as 0.20 kW/m^2·K and A_{ev} is the reference area, adopted as 100 m^2. These values are the same ones used by d'Accadia and Rossi (1998), which also considered the cost of utility water as $0.25.10^{-3}$ \$/kg and the value of electricity as \$10/GJ. It was done an estimation of the operation and maintenance costs, amortizing the general cost of the equipment. In this estimation the following factors were considered: the maintenance cost of each equipment (as a percentage of its investment cost - %C.I.); the useful life for each equipment (V.U.) and the annual time of operation (T.O.). The equipment cost is defined as:

$$Z = \left(1 + \frac{\%C.I.}{100} \right) \times \frac{\text{initial cost}}{V.U.} \times \frac{\text{lano}}{T.O.} \quad (5)$$

where %C.I. = 6%, V.U. = 15 years and T.O. = 8000 hours (Kwak *et al*, 2003).

From the exergetic balance devoleped and from the derived equations proposed by Lozano and Valero (1993), it was determined the exergoeconomic costs matrix presented in Figure 2.

BW	Bag	BP21	BP23	BP34	BP41	BT21	BT23	BT43	BT41	S21	S14	S23	S43	B			
				1				-1							BW		-Zval
					1				1		1			-1	Bag		-Zevap
1		-1				-1				1					BP21		-Zcomp
	1		1				1					-1			BP23		-Zcond
										-1	-1	1	-1		BP34		0
						1	-1	1	-1						BP41		0
		1	-1	-1	-1										BT21		0
	1														BT23	=	Cag*Bag
1															BT43		CW*BW
							-1/BT23		1/BT41						BT41		0
			-1/BP23		1/BP41										S21		0
				-1/BP34	1/BP41										S14		0
										-1/S21			1/S43		S23		0
											-1/S14		1/S43		S43		0
		-1/BP21				1/BT21									B		0

Figura 2 – Matrix refering to the exergoeconomic cost of refrigeration cycle stream.

3. Results and Discussion

The calculations of exergoeconomic costs were performed for the refrigeration cycles in the ethylene an propylene production process and the results are presented in Tables 2 and 3.

Initially, it is important to do some observations about de matrix A (represented by Figure 2) and in the solution of the linear system **A**.**x = b**. In this work it was observed that matrix $\mathbf{A} \in \Re^{nxn}$ is not singular, therefore it admits inverse and it has complete rank (rank = 15). The augmented matrix $\mathbf{[A,b]} \in \Re^{nxn+1}$ has the same rank of matrix **A**, therefore the system **A**.**x = b** is consistent and it has an unique solution. This solution is obtained directly by solving the system, without requiring an approximation of results.

The lowest unit exergoeconomic cost for the stream-product B (Figure 1) was for cycle 01 of the demethanizer column and the one for the propylene-propene splitter was the greatest of all.

For the other streams of the cycle (*S, BP, BT* – Figure 1), the exergoeconomic costs were in the same range. It can also be observed, from Tables 2 and 3, that the integrated cycle of the demethanizer column is the most expensive, this happens because methane is the refrigerant, with the lowest temperatures and it is one that has the highest costs. The second more expensive is the cycle of ethylene-ethane splitter column, showing that both are the most indicated to attempt optimization procedures.

Table 2 – Exergoeconomic costs (C), unit (C_{un}) and costs rate (Z) for each equipment of the demethanizer colum refrigeration cycles.

	Cycle 01		**Cycle 02**		**Cycle 03**	
	C ($/s)	C_{un} ($/GJ)	C ($/s)	C_{un} ($/GJ)	C ($/s)	C_{un} ($/GJ)
BW	0.04	10.00	0.11	10.00	0.09	10.00
Bwt	0.61	83.49	0.47	106.16	0.36	524.60
BP21	0.15	26.82	0.19	8.79	0.13	17.98
BP23	0.09	26.82	-0.02	8.79	0.01	17.98
BP34	0.08	26.82	0.14	8.79	0.04	17.98
BP41	0.16	26.82	0.06	8.79	0.08	17.98
BT21	-0.07	26.82	-0.01	8.79	0.00	17.98
BT23	-0.19	41.92	-0.06	34.88	0.01	14.26
BT41	0.12	31.62	0.20	29.86	0.07	14.21
BT43	0.24	41.92	0.25	34.88	0.06	14.26
S21	0.03	27.07	0.04	18.69	0.03	15.69
S14	0.26	27.07	0.29	18.69	0.33	15.69
S23	0.33	27.07	0.39	18.69	0.38	15.69
S43	0.04	27.07	0.06	18.69	0.03	15.69
B	0.66	56.59	0.61	134.22	0.47	259.70
Total	3.07	-	2.90	-	2.09	-
	$Z.10^3$ ($/s)		**$Z.10^3$ ($/s)**		**$Z.10^3$ ($/s)**	
Valve	0.36	-	0.39	-	0.46	-
Evaporator	1.95	-	0.49	-	0.30	-
Compressor	4.34	-	9.44	-	11.31	-
Condenser	1.07	-	2.22	-	2.17	-
Cwt*Bwt	608.96	-	371.32	-	364.49	-
Cw*Bw	39.47	-	73.66	-	88.67	-

Cwt = cost of cold water used in the condensers.

Tabela 3 – Exergoeconomic costs (C), unit (C_{un}) and costs rate(Z) for each equipment of the deethanizer column, ethylene-ethane splitter and propylene-propene splitter.

	Deethanizer		**Ethylene-Ethane Splitter**		**Propylene-Propene Splitter**	
	C ($/s)	C_{un} ($/GJ)	C ($/s)	C_{un} ($/GJ)	C ($/s)	C_{un} ($/GJ)
BW	0.06	10.00	0.07	10.00	0.01	10.00
Bwt	0.23	524.60	0.37	524.60	0.05	524.60
BP21	0.09	17.34	0.11	17.95	0.02	18.71
BP23	0.01	17.34	0.01	17.95	0.00	18.71
BP34	0.03	17.34	0.03	17.95	0.01	18.71
BP41	0.05	17.34	0.07	17.95	0.01	18.71
BT21	0.01	17.34	0.00	17.95	-0.00	18.71
BT23	0.01	17.55	0.01	12.16	0.00	31.16
BT41	0.05	17.58	0.05	11.94	0.02	29.71
BT43	0.05	17.55	0.05	12.16	0.02	31.16
S21	0.02	16.77	0.03	15.49	0.01	15.45
S14	0.21	16.77	0.34	15.49	0.04	15.45
S23	0.26	16.77	0.39	15.49	0.05	15.45
S43	0.02	16.77	0.02	15.49	0.01	15.45
B	0.30	241.69	0.46	155.79	0.06	1407.88
Total	1.40	-	2.01	-	0.31	-
	$Z.10^3$ ($/s)		**$Z.10^3$ ($/s)**		**$Z.10^3$ ($/s)**	
Valve	0.28	-	6.17	-	0.06	-
Evaporator	0.21	-	0.76	-	0.01	-
Compressor	8.23	-	31.52	-	1.62	-
Condenser	1.35	-	1.83	-	0.275	-
Cwt*Bwt	231.27	-	467.34	-	45.74	-
Cw*Bw	61.32	-	106.84	-	13.32	-

4. Conclusions

In this present work the exergoeconomic costs of the refrigeration cycles in the ethylene and propylene production process were determined.

The computational tool (HYSYS©) and the methodology used in this work have shown to be adequate for estimating exergoeconomic costs of the cycles.

It was observed that the integrated cycle of the demethanizer column is the most expensive, followed by the cycle of the ethylene-ethane splitter column, showing that both are the most indicated cycles for optimization procedures.

5. References

D'Accadia, M. D., Rossi, F., 1998, Thermoeconomic Optimization of a Refrigeration Plant, International Journal of Refrigeration, vol. 21, nº 1, 42-54.

Lozano, M. A., Valero, A., 1993, Theory of the Exergetic Cost, Energy, vol. 18, nº 9, 939-960.

Kwak, H. Y., Kim, D. J., Jeon, J. S., 2003, Exergetic and Thermoeconomic Analyses of Power Plants, Energy, vol. 28, 343-360.

Kotas, T. J., 1995, The Exergy method of Thermal Plant Analysis, Butterworths.

Rossi, J. S., Fábrega, F.M., d'Angelo, J.V.H., 2007, Exergetic Analysis of Refrigeration Cycles of na Ethylene Production Process, 8º Congreso Interamericano de Computación Aplicada a la Industria de Processos – CAIP, 415-418.

Sahoo, P.K., 2008, Exergoeconomic analysis and optimization of a cogenaration system using evolutionary programming, Applied Thermal Enginnering, nº 28, 1580-1588.

Valero, A., Lozano, M. A., Muñoz, M, 1996, A General Theory of Exergy Saving: Part I. On the Exergetic Cost, Part II. On the Thermoeconomic Part II. Energy Saving and Thermoeconomics. In Computer-Aided Engineering Energy Systems, vol. 3 Second Law Analysis and Modelling, R.A. Gaggioli, ed. ASME, New York, 1996.

MME – Ministério de Minas e Energia: https://ben.epe.gov.br/ - site accessed in January, 2009.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)

Use of Excess Reactant in the Design of Reactive Distillation

Li-Chiang Lin, Jian-Kai Cheng, Cheng-Ching Yu

Dept. of Chemical Engineering, National Taiwan University, Taipei 106-17, TAIWAN

Paper withdrawn at the request of the author after final proofing stage

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Published by Elsevier B.V.

Determining Reaction Networks

Samantha C. Burnham[a] and Mark J. Willis[b]

[a]*Department of Chemical Engineering, Curtin University of Technology, GPO Box U1987, Perth, WA 6845, Australia*
[b]*School of CEAM, Newcastle University, Newcastle-upon-Tyne, NE1 7RU, UK*

Abstract

This work introduces a method to identify chemical reaction networks through the analysis of process data obtained in a laboratory environment (batch reactors operating away from steady state), assuming no *a priori* information about reaction stoichiometries or species structures. This is a three stage process achieved through: a) the determination of conservation relations, number of linearly independent reactions b) the partitioning of the matrices of conservation relations, thus reaction stoichiometry to reduce the parameter identification problem c) screening for a unique chemical reaction network. A case study is used to demonstrate the method.

Keywords: reaction networks, mathematical modelling, stoichiometry.

1. Introduction

Aris and Mah[1] were the first to demonstrate that by calculating the rank of a matrix of the process data it is possible to infer the actual number of reactions. More recently a number of contributions have attempted to infer stoichiometric information from process data. Hamer[6] and Bonvin and Rippin[2] demonstrated that the number of reactions and observed stoichiometry may be obtained from a singular value decomposition (SVD) analysis[7] of the matrix of molar change of chemical species. Bonvin and Rippin[2] also incorporated the chemometric method target factor analysis (TFA), to check if a proposed (target) stoichiometry was consistent with the measured data. Fotopoulos *et. al.*[4] and Georgakis and Lin[5] propose and use structured target factor analysis (STFA) to systematically check a pre-defined set of structured, sparse vectors against the observed stoichiometry. The procedure introduced in this paper compliments these earlier techniques and is consistent with the work completed in biochemical systems theory (e.g. see Palsson, 2006[10]).

2. Background

Assuming constant density, the material balance expression for a (homogeneous phase) well mixed, batch reactor that is operating isothermally, may be written as (1)

$$\frac{d[x_i]}{dt} = f_i \quad i = 1,...,n \tag{1}$$

where $[x_i]$ is the molar concentration of species i at time t. (1) is a set of coupled ordinary differential equations (ODEs) that describe the dynamic behaviour of the reactive species due to chemical reactions, as represented by the n reaction fluxes f_i.
The flux terms f_i are directly linked to the stoichiometries of the N_R reactions taking place and the kinetic rate terms of these reactions. The (n x N_R) stoichiometric matrix $\mathbf{V}$ describes the topology of a chemical reaction network. The stoichiometries of the

chemical species in the reactions, which make up the network, provide the entries of the stoichiometric matrix. For a network with n species and N_R chemical reactions, the stoichiometric relationships, which are assumed to be time invariant, are given by (2).
Where $v_{i,l}$, $i = 1, \ldots, n$, $l = 1, \ldots, N_R$, are the stoichiometric coefficients for the i^{th} species in the l^{th} reaction. By convention, $v_{i,l} < 0$ for a species that undergoes net consumption in a reaction, $v_{,il} > 0$ for a species that undergoes net production and $v_{i,l} = 0$ for a species that is either not involved in the l^{th} reaction or has no net change in it.
The component balances may now be expressed in terms of **V** and the N_R individual reaction rates; in matrix-vector form, (3).

$$\mathbf{V} = \begin{bmatrix} v_{1,1} & \cdots & v_{1,N_R} \\ \vdots & & \vdots \\ v_{n,1} & \cdots & v_{n,N_R} \end{bmatrix} \qquad (2) \qquad \frac{d\mathbf{x}}{dt} = \mathbf{V}\mathbf{r} \qquad (3)$$

Where **r** is the ($N_R \times 1$) dimensional vector of reaction rates and **x** is the ($n \times 1$) dimensional vector of species concentrations $[x_1],\ldots,[x_n]$ at time t. The N_R reaction rates constituting **r** are, in general, non linear functions of the concentrations **x**. If elementary reactions are assumed then the form of the N_R rate terms in **r** is determined uniquely by the reactants in each of the N_R elementary reactions.
The law of mass action kinetics states that the rate of an elementary reaction may be assumed to be directly proportional to the product of the reactant concentrations. Consider the chemical reaction (4): a molecules of species x_1 and b molecules of species x_2 form c molecules of species x_3 and d molecules of species x_4, with rate coefficient k.

$$ax_1 + bx_2 \xrightarrow{k} cx_3 + dx_4 \qquad (4)$$

The rate r of this reaction at any time, according to the law of mass action, is (5).

$$r = k[x_1]^a[x_2]^b \qquad (5)$$

Thus, given that the correct stoichiometric matrix **V** may be determined it is possible to generate the structure of the vector of reaction rates. Leaving the kinetic rate constants to be identified (possible by standard kinetic fitting).

3. Case Study

3.1. Data Generation
For example, consider the following chemical reaction network comprising five reactive species involved in three elementary reactions with rate coefficients k_1 and k_2 and k_3

$$(CH_3CO)_2O + H_2O \xrightarrow{k_1} 2CH_3COOH \qquad CH_3COOH + CH_3OH \underset{k_3}{\overset{k_2}{\leftrightarrow}} C_3H_6O_2 + H_2O \qquad (6)$$

In the first reaction acetic anhydride reacts with water to produce acetic acid. In the second (reversible) reaction acetic acid reacts with methanol to produce methyl acetate and water. The matrix **V** and vector **r** for this network are given by, (7).

$$\mathbf{V} = \begin{bmatrix} -1 & 0 & 0 \\ -1 & 1 & -1 \\ 2 & -1 & 1 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix} 1 \qquad \mathbf{r} = \begin{bmatrix} k_1[(CH_3CO)_2O][H_2O] \\ k_2[CH_3COOH][CH_3OH] \\ k_3[C_3H_6O_2][H_2O] \end{bmatrix} \qquad (7)$$

Hence the n ODEs, for known species initial conditions can be used to describe the temporal evolution of the species concentrations, (8).

$$\frac{d\mathbf{x}}{dt} = \mathbf{V}\mathbf{r} = \begin{bmatrix} -k_1[(CH_3CO)_2O][H_2O] \\ -k_1[(CH_3CO)_2O][H_2O] + k_2[CH_3COOH][CH_3OH] \\ \quad -k_3[C_3H_6O_2][H_2O] \\ 2k_1[(CH_3CO)_2O][H_2O] - k_2[CH_3COOH][CH_3OH] \\ \quad +k_3[C_3H_6O_2][H_2O] \\ -k_2[CH_3COOH][CH_3OH] + k_3[C_3H_6O_2][H_2O] \\ k_2[CH_3COOH][CH_3OH] - k_3[C_3H_6O_2][H_2O] \end{bmatrix} \quad (8)$$

For reaction scheme (6), it was assumed that only initial concentrations of acetic anhydride, methanol and water were present. A simulated batch experiment was performed, with each of the initial concentrations specified as 0.10moldm^{-3}. Note that the volume of the batch is specified as 1 litre and that the reaction takes place in an inert solvent. The kinetic rate constants were specified as k_1 = 0.10, k_2 = 0.15 and k_3 = 0.05 min^{-1}mol^{-1}dm^3. Eleven concentration measurements of each species were generated by numerical integration of the true system ODEs, (8), using the initial reactant concentrations specified above. The total simulation period was specified as 1200 mins and measurements of all species were recorded every 120 mins, therefore the number of time points, *N*, is eleven. In this case no simulated measurement error was added to the signals. The generated data is depicted in figure 1. Curves were fitted to the process data (also figure 1 – using MATLAB ® v.7.1 curve fitting toolbox) for the purpose of estimating the rate of change of species concentrations (figure 2 – calculated from derivatives of the fitted curves). The actual derivatives are also shown in figure 2, where the mean squared errors for the respective approximations are less than 3.840x10^{-6}.

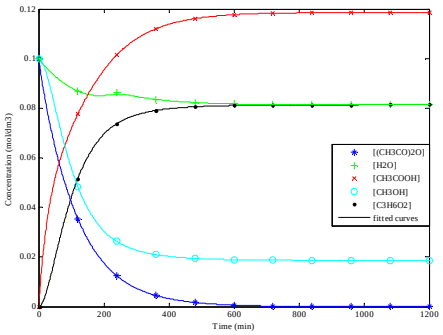


Figure 1: Measured concentration values and fitted curves

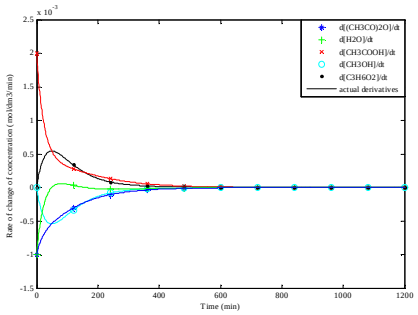


Figure 2: Approximate and accurate derivative values for the concentration data in figure 1.

3.2. Stage a) - Conservation Relations

3.2.1. Moiety Conservation

The term conserved moieties[11] defines molecular subgroups that are conserved during the evolution of a chemical reaction network. A conserved moiety is time invariant and is solely determined by the initial conditions imposed on the system, in a batch reactor by the initial concentrations of the species. A practical reason why conservation relations should be identified as a preliminary step in chemical reaction network analysis is that they allow the set of species, S_S, to be divided into independent and dependent species, S_I and S_D, respectively. The concentrations of the S_D can be calculated from the concentrations of the S_I. Thus the conserved moieties represent dependencies which can be removed to reduce a system's dimensionality. It should be noted that for each individual reaction one species involved in that reaction (not already in the set of independent species) must be included in the set of independent species.

3.2.2. Determining conservation relations from process data

The data in figure 2 was compiled to give the matrix $\dot{\mathbf{X}}$ (N x n) which represents rate of change of concentration of all species (each column) at all measured time points (each row). It is documented[10], that the conservation relation matrix, **C** (N_C x n), lies in the left null space of the stoichiometric matrix, **V**. Given that the transpose of the data matrix ($\dot{\mathbf{X}}'$) lies in the same vector space as **V**, it follows that **C** may be estimated from the (right) null space of $\dot{\mathbf{X}}$ and is estimated to be (9)[i].

Note that there are N_C conservation relations, equal the estimated number of reactions, $\hat{N}_R$, in this case two. The actual number of reactions is three (6), but only two are detected as the second and third reactions are fully correlated.

Note: given that **C** lies in the left null space of **V**; it is also true that **V** lies in the null space of **C**. Therefore by calculating the null space of **C** it is possible to gain an estimate of **V**; this will be referred to as an observed stoichiometry, $\mathbf{V_o}$, (10) – see footnote i.

$$\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ -1 & 1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 1 \end{bmatrix}^{\mathbf{T}} \qquad (9) \qquad \mathbf{V_o} = \begin{bmatrix} -0.5 & -0.5 & 1 & 0 & 0 \\ -0.5 & 0.5 & 0 & -1 & 1 \end{bmatrix}^{\mathbf{T}} \qquad (10)$$

3.3. Stage b) – Reduction of the parameter identification problem

3.3.1. MatrixPartitioning

Given that $\dot{\mathbf{X}}$ is ordered so that the first $\hat{N}_R$ = *rank*($\dot{\mathbf{X}}$) columns ($\dot{\mathbf{X}}_\mathbf{I}$) correspond to $\hat{N}_R$ independent species (S_I) and, the remaining $N_c = n - \hat{N}_R = n$ - *rank*($\dot{\mathbf{X}}$) columns ($\dot{\mathbf{X}}_\mathbf{D}$) correspond to the dependent species (S_D), $\dot{\mathbf{X}}$ is partitioned as $\dot{\mathbf{X}}$ = [$\dot{\mathbf{X}}_\mathbf{I}$ ¦ $\dot{\mathbf{X}}_\mathbf{D}$].

By partitioning $\dot{\mathbf{X}}$ it is true that **C** is also partitioned as $[-\mathbf{L_0} \,¦\, \mathbf{I}]^\mathbf{T}$, where $\mathbf{L_0}$ is a link-zero matrix ($\hat{N}_R$ x N_C), given that $N_C = n - N_R$, and **I** is an identity matrix (N_C x N_C). Further the stoichiometric matrix **V** may be partitioned to $[\mathbf{V_I} \,¦\, \mathbf{V_D}]^\mathbf{T}$ where $\mathbf{V_I}$ is a ($\hat{N}_R$ x $\hat{N}_R$) matrix representing the stoichiometries of the independent species and $\mathbf{V_D}$ is a (N_C x $\hat{N}_R$) matrix representing the stoichiometries of the dependent species. $\mathbf{V_D}$ can be calculated by (11)[9].

$$\mathbf{V_D} = \mathbf{L_0 V_I} \qquad (11)$$

Therefore the parameter estimation problem is reduced from determining a (n x N_R) stoichiometric matrix **V** to determining a ($\hat{N}_R$ x $\hat{N}_R$) matrix, $\mathbf{V_I}$, representing the stoichiometries of the independent species.

3.3.2. Defining a set of possible stoichiometric vectors

Assumption 1: reactants and products are restricted to be either unimolecular or bimolecular with a maximum of two reactants and products[ii].

Given assumption 1, the stoichiometric coefficients of the vectors within $\mathbf{V_I}$ belong to the set, S_C = {-2, -1, 0, 1, 2}. Given the choice of independent species, this integer representation is simply a description of the reactants (negative integers), products

[i] In this work the null spaces are calculated in MATLAB using the reduced row echelon form as a basis. The resulting matrices that this procedure produces are not unique. Applying a different basis will lead to different, but equally valid matrices.

[ii] This representation could be easily modified to include reaction types of greater or lesser complexity.

(positive integers) or species not involved (zeros) in each reaction. For $\hat{N}_R$ = 2, the set of all (uncorrelated) vectors, S_{VI}[iii], is (12).

$$\begin{matrix} 0 & 1 & 1 & 1 & 2 & 1 \\ 1 & -2 & -1 & 1 & -1 & 0 \end{matrix} \quad (12)$$

Given **C**, (9), $\mathbf{L_0}$ may be defined as [-1, 1, -1; -1 -1 1]. Therefore, given $\mathbf{L_0}$ and S_{VI}, the respective stoichiometries for the dependent species, S_{VD}, may be calculated from (11). Combining these with S_{VI} with S_{VD} a set of target stoichiometries is produced (13). Applying assumption 1, a number of these vectors can be discounted – highlighted by a shaded background. Leaving three feasible vectors – but as these vectors may be in the opposite direction they represent six possible reactions (14).

0	1	1	1	2	1
1	-2	-1	1	-1	0
-1	1	0	-2	-1	-1
-1	3	2	0	3	1
1	-3	-2	0	-3	-1

(13)

$$CH_3COOH + CH_3OH \leftrightarrow H_2O + C_3H_6O_2$$
$$2CH_3COOH \leftrightarrow (CH_3CO)_2O + H_2O \qquad CH_3COOH + C_3H_6O_2 \leftrightarrow (CH_3CO)_2O + CH_3OH \quad (14)$$

3.4. Stage c) – Screening for a Coupled System

3.4.1. Best Subsets Regression

From (14) there are six possible reactions, whose rate vectors may be identified, with unknown rate constants (15).

$$r_i = \{k_a[CH_3COOH][CH_3OH] \quad k_b[H_2O][C_3H_6O_2] \quad k_c[CH_3COOH]^2 \quad k_d[(CH_3CO)_2O][H_2O] \quad k_e[CH_3COOH][C_3H_6O_2] \quad k_f[(CH_3CO)_2O][CH_3OH]\} \quad (15)$$

Given that the data in figure 1 was compiled to give the matrix **X** (*N* x *n*), concentration of all species at all measured time points, then the values of r_i/k_i in (15), for *i* = a:f, at the measured time points may be calculated.

Given that the correct r_i/k_i terms will be proportional to the respective columns of $\dot{\mathbf{X}}$ (8), with proportional constants of the rate coefficients, k_i, it is possible to use regression analysis to determine the relevant r_i terms. Best subsets regression allows a number of models containing different predictors and different numbers of predictors to be compared. As $\hat{N}_R$ was estimated to be two, we know that N_R is a minimum of two (uncorrelated reactions) and a maximum of four (including the reverse reactions of the two uncorrelated ones). Hence, it is possible to compare the models comprised of all possible combinations of two, three or four terms from the set of rate terms in (15), when regressed against the rate of change of CH_3COOH (chosen as it is present in all six possible reactions)[iv].

It is usually not feasible to model each combination of predictors individually; alternatively, a statistical algorithm is used to produce summary statistics for the

[iii] The set of all possible (N_R x 1) vectors were generated using the set of stoichiometric coefficients, $\boldsymbol{S_C}$. The set were then reduced by removing any that implied reactions that were not unimolecular or bimolecular. All sets of correlated vectors were identified and the minimum norm solution was retained within the set, $\boldsymbol{S_V}$.

[iv] In other examples it may be necessary to perform regression against the rate of change of concentration of more than one species, given that one species may not appear in all possible reactions. Also it may be performed against the rate of change of all species to confirm any results.

possible combinations of predictors.
Performing best subsets regression the following rate terms were identified to produce the best model $k_a[CH_3COOH][CH_3OH]$, $k_b[H_2O][C_3H_6O_2]$ and $k_d[(CH_3CO)O][H_2O]$ with an adjusted R^2 value of 100% and Mallows' *C-p* statistic[8] of 2.5. Thus the reactions are correctly identified as (16) – *cf* (6). It would then be possible to use standard regression or standard kinetic fitting techniques to determine the values of kinetic rate coefficients[3].

$$CH_3COOH + CH_3OH \underset{k_b}{\overset{k_a}{\leftrightarrow}} C_3H_6O_2 + H_2O \qquad (CH_3CO)_2O + H_2O \overset{k_d}{\rightarrow} 2CH_3COOH \tag{16}$$

4. Discussions, Conclusions and Future Work

This paper proposes a method for the determination of chemical reaction networks. The procedure is demonstrated to be successful, through a worked example, consisting of five chemical species and three reaction steps. The presented analyses are relatively simple from a mathematical viewpoint and use concepts and tools that will be familiar to most researchers and engineers, for example curve fitting and regression analysis. The majority of methods from the literature are based on system perturbations around stationary states, thus do not lend themselves to the type of experimentation performed in most chemical laboratory contexts. In comparison the techniques proposed here, are based solely on the analysis of process data obtained from batch reactors. There are a number of potential limitations to the approach in its current form. For example, in accordance with other approaches in this area, and in the field of system identification in general, adequate experimental data (i.e. data from a system with sufficient excitation and appropriate data collection) is a pre-requisite to the success of the presented approach. Currently, however, the approach lacks diagnostic tools to systematically evaluate whether the acquired data is sufficient, in terms of information content, for successful network identification.

References

1. R. Aris and R.H.S. Mah (1963) Independence of chemical reactions, *Industrial and Engineering Chemistry Fundamentals*, 2, 90-94
2. D. Bonvin and D.W.T Rippin (1990) Target factor analysis for the identification of stoichiometric models, *Chemical Engineering Science*, 44, 3417-3426.
3. S.C. Burnham, D.P Searson, M.J. Willis and A.R. Wright (2007) Inference of chemical reaction networks, *Chemical Engineering Science*, 63, 4, 862-873
4. J. Fotopoulos, C. Georgakis and H.G. Stenger (1994) Structured Target Factor Analysis for the Stoichiometric Modeling of Batch Reactors, *Proceedings of the 1994 American Control Conference*, 495-499
5. C. Georgakis and R. Lin (2005) Stoichiometric Modeling of Complex Pharmaceutical Reactions, *Proceedings of the Annual AIChE meeting in Cincinnati, OH, November*
6. J.W. Hamer (1989) Stoichiometric interpretation of multireaction data: application to fed-batch fermentation, *Chemical Engineering Science*, 44, 2363-2374
7. C.L. Lawson and R.L Hanson (1974) *Solving Least Squares Problems*, Prentice-Hall, New Jersey
8. C.L. Mallows (1973) Some Comments on Cp, *Technometrics*, 15, 661-675
9. G. Maria (2005) Relations between apparent and intrinsic kinetics of "programmable" drug release in human plasma, *Chemical Engineering Science*, 60, 6, 1709-1723
10. B.O Palsson (2006) *Systems Biology: Properties of Reconstructed Networks*, Cambridge University Press, New York
11. J. Reich and E.E. Selkov (1981) *Energy Metabolism of the Cell: A Theoretical Treatise*, Academic Press, New York

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Comparison of Functional Modeling Techniques for Autonomous Systems Development

José Luis de la Mata[a] and Manuel Rodríguez[b]

[a,b]*Autonomou System Laboratory-Technical University of Madrid, José Gutiérrez Abascal, Madrid 28006, Spain*
[a]*joseluis.delamata@upm.es*
[b]*manuel.rodriguezh@upm.es*

Abstract

When developing autonomous systems one important step is to have integrated different views of the system modelled. One essential view of the system behavior is the functional view, so the first step is to choose a functional modeling technique. In order to be able to have an integrated model an open, neutral, domain and platform independent language is needed. The Systems Modeling Language (SysML) complies with these requirements. The SysML is a subset and an extension of the well-known Unified Modeling Language (UML).

Keywords: Functional modeling, autonomous systems.

1. Autonomous Systems Development

Etymologically autonomy means “giving itself its own laws” (Greek auto-self and nomos-law). That meaning, translated to technical language, can be read as the system decides about the control strategies to follow. Being autonomous seems to be at the very central objectives of most engineers. The model-driven approach, a generalization of formal development methods, is the current engineering trend trying to tackle the autonomy problem [1].
When developing autonomous systems, a very important aspect is that the system has “knowledge” about itself, i.e., its structure, behavior, etc. For this reason, a model-driven approach is used. A complete model will require integrated different views of the system. One essential view of the system behavior is the functional view, given by a functional model. Models are the core around which everything is developed.
This paper has been realized in the framework of the ASys project of the Autonomous Systems Laboratory research group. ASys is a long-term research project focused in the development of technology for the construction of autonomous systems.

2. Functional Modeling

A model is a simplified representation of the reality. It may have as many views as reality has. The question is to define the more interesting (useful) views to be included in a model. It seems that two philosophically fundamental views are mostly important and have to be considered: functional and structural. A functional model is a structured representation of the functions, activities or processes within the modeled system [2].

2.1. Multilevel Flow Modeling (MFM)

Multilevel Flow Models (MFM) are graphical models of technical processes [3]. The goals describe the purposes of a system and its subsystems, and the functions describe

the system's abilities in terms of flows of mass, energy and information. MFM also describes the relations between the goals and the functions that achieve those goals, and between functions and the subgoals that provide conditions for these functions. Mass and energy flow structures are used to model the functions of the plant and activity and information flow structures are used to model the functions of the operator and control systems. Using these concepts it is possible to represent knowledge of complex process plants. The basic MFM functions were presented by Lind [3], the use of information functions was introduced by Larsson [4], finally, Petersen [5] presented causality considerations between functions. Later, in [6] some extensions were proposed in order to perform control system reconfiguration.

2.2. Goal Tree – Success Tree (GTST)

The GTST is a functional decomposition framework for modeling complex physical systems. It is a functional hierarchy of a system organized in levels beginning with an objective at the top. The objective describes the main purpose of the system in an unambiguous term [7]. The relations in hierarchies represent a connection between different nodes of a hierarchy or between nodes of two correlative hierarchies. The type of relations in the GTST technique can be tagged as logical, physical or fuzzy [8].

The lowest level of the GTST models explains how system parts interact with each other to achieve higher level functions. In order to present the success logic of a very complex interacting system in a compact and transparent way, the Master Logic Diagram (MLD) is used. The MLD is such a model to represent the relationship between the main and support functions [8].

3. Functional Modeling Techniques Comparison

A. Jalashgar in [9] provides an overview of both methods but no comparison is made. According to the differences presented and their different performance, it is possible to compare both techniques.

Although both techniques are functional goal oriented techniques, there are several differences between them. While MFM provides a good general view (even of very complex systems), GTST doesn't, essentially because MFM models are not restricted to hierarchical structures. Relations between different hierarchical levels and different parts of the model are allowed and can be represented explicitly i.e., hierarchy does not constrain the model.

GTST represents all kinds of knowledge of the technical system and it has got textual appearance. This could be an advantage because it simplifies the implementation of the model, but it can also be a disadvantage because it makes complicated the identification of similarities between the parts of the model and between different models.

On the other hand, MFM represents in an explicit way the objectives and the system functions as well as their relations. A MFM model captures the functions and the goals of a system but there is a lack of homogeneous semantics because there are different ways of representing the same system and different levels of abstraction can be considered.

But the most important feature of MFM is that it is a *flow ontology*. Mass, energy and information flows are responsible for causal interactions in chemical plants. Safety, productivity and economy are managed by administrating these flows. The concepts of mass and energy flows can be applied on multiple levels of physical aggregation and to a wide variety of physical processes. According to the previous paragraphs, MFM seems to be the most suitable functional modeling technique for the implementation of the functional view of a chemical plant.

4. SysML

The Systems Modeling Language is a domain-specific modeling language for systems engineering that is defined as a dialect (Profile) of UML 2 [10]. The language provides graphical representations with a semantic foundation for modeling system requirements, behavior, structure, and parametrics, which is used to integrate with other engineering Modeling Language [11]. The block is the basic unit of structure in SysML and can be used to represent hardware, software, facilities, personnel, or any other system element [12]. The four pillars of SysML are:

1. *Structure*. Block definition diagrams and internal block diagrams represent the system structure.
2. *Behavior*. Use case diagram, activity diagram, sequence diagram, and state machine diagram.
3. *Requirements*. Provides a bridge between the typical requirements management tools and the system models.
4. *Parametrics*. Represents constraints on system property values as a mean to integrate the specification and design models with engineering analysis models.

As aforementioned, the block is the basic unit of structure in SysML and the basic units of MFM models are goals, structures and flow functions. A SysML block will model each flow concept of MFM. For example, a transport function will be a block with a flow input, a flow output and a condition port. Its behavior can be modeled with the SysML state diagram. In the same fashion, a structure can be modeled by a block with various achieve ports, an achieve-by-control port, a control port and various condition ports. The internal structure of a MFM structure is represented in the SysML internal block diagram.

5. Industrial process

5.1. MMAN production

MMAN is an explosive obtained from the reaction of nitric acid and monomethylamine (MMA) in a stirred tank reactor. The reaction is exothermic and it must be kept under 60ºC by the cooling jacket which is fed with water. MMA is stored as liquefied gas so it is vaporized before being fed to the reactor. To avoid crystallizations and product decomposition, MMAN has to be kept above 55ºC and below 70ºC. A P&I diagram of the process is shown in Fig. 1.

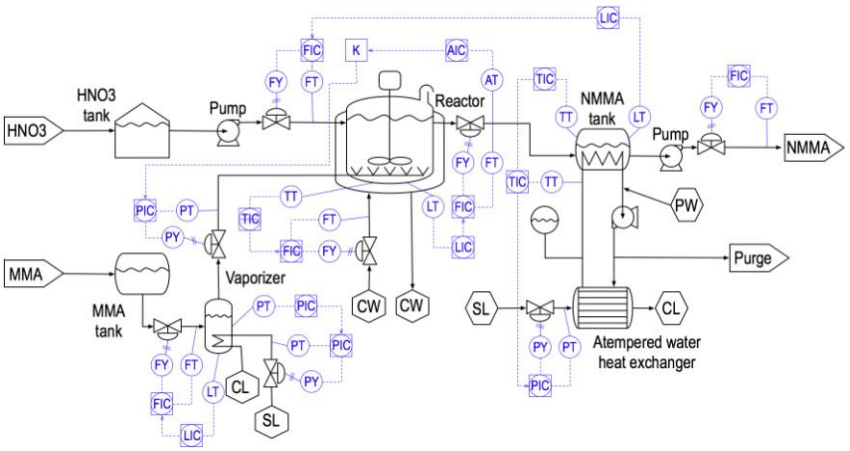


Fig. 1. MMAN process P&I diagram.

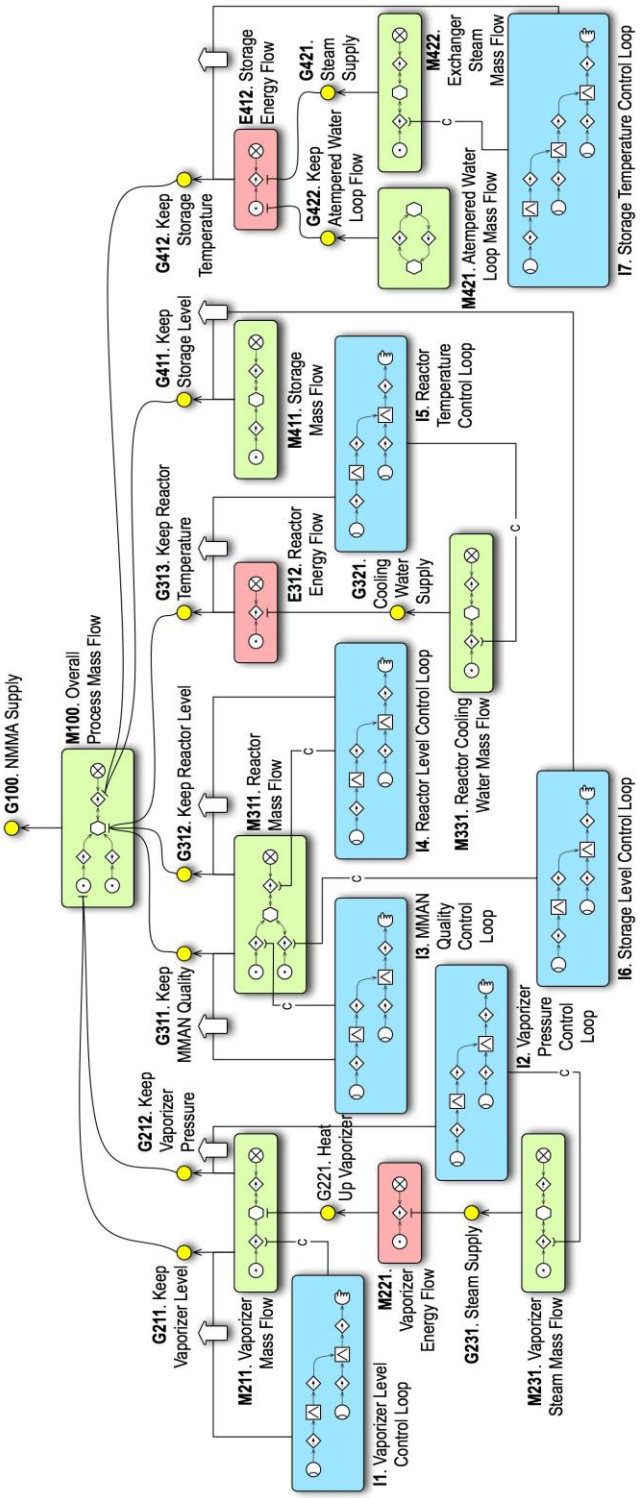


Fig 2. MMNA process MFM model.

A set of goals and subgoals has been identified in order to guarantee a good plant behavior. The main goal is to "supply MMAN" (G100), this goal is achieved in a mass structure (M100) which is conditioned by lower level goals. There are eight mass structures, three energy structures and seven information structures. Related to these structures there are thirteen goals and more than one hundred flow functions. All the relations between structures, goals and flow functions are collected in Fig. 2.

5.2. SysML Block Diagram

The MMAN MFM model is too big to represent it fully in a SysML block diagram, so only a part of it will be analyzed. The block diagram shown in Fig. 3 has been developed considering goal G313 (Keep Reactor Temperature) in Fig. 2 and its related structures and goals (E312, M331, I5 and G321). Keeping the reactor temperature is achieved in energy structure E312 which is conditioned by goal G321 (Cooling water supply). This goal is achieved in structure M331 which is controlled by the information structure I5. The goal G312 is achieved by control in structure I5. These concepts are represented in the block diagram shown in Fig. 3.

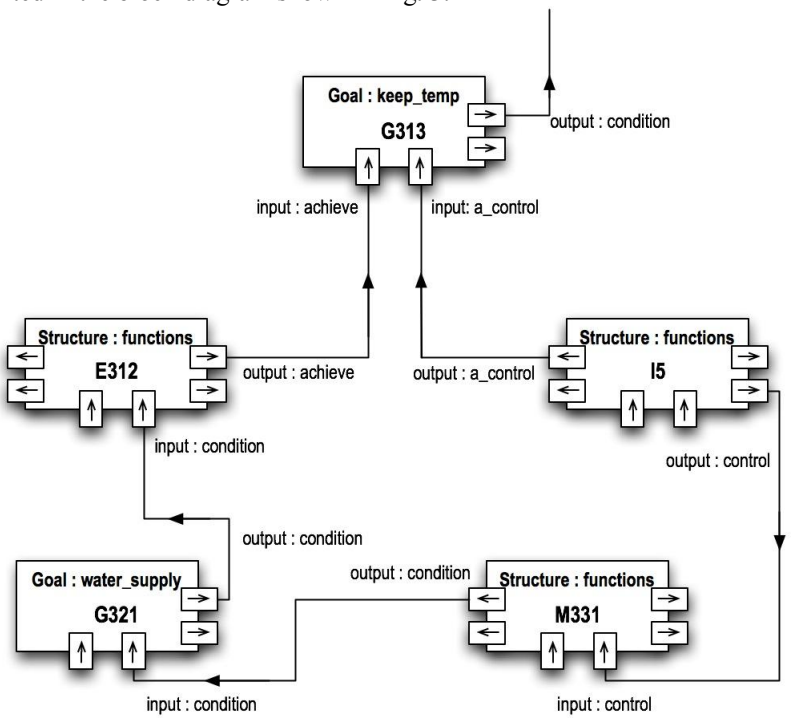


Fig. 3. SysML Block Diagram

In Fig. 4 the Internal Block Diagram of the energy structure is depicted. The structure consists of a source function, a transport function and a sink function, which are represented by the related blocks. There are different flows in both diagrams (figures 3 and 4): mass, energy, information, condition, control, achieve and achieve-by-control.
The MFM vertical relations are modeled as SysML connections, in fact, flow connections in the block diagram. The block diagram captures the main structure and relations of MFM models.

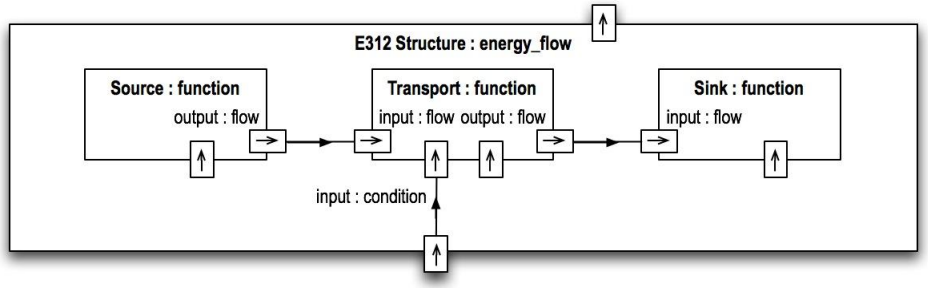


Fig. 4. SysML Internal Block Diagram of structure E312.

6. Conclusions and Further Work

In this paper a comparison between functional modeling techniques (MFM and GTST) has been made. Both techniques provide modeling features that are useful for developing the functional view of process plants but the flow ontology feature makes MFM more suitable for process systems.

An MFM model of the MMAN production process has been implemented. This model also has been developed in SysML using blocks to model MFM concepts and connections to model MFM relations.

MFM models implement the functional view of a system. These models can be used for multiple purposes such as risk analysis, fault diagnosis and control system reconfiguration. Further work includes the development of these features in SysML along with quantitative information integration.

References

Ricardo Sanz and Manuel Rodríguez. The ASys Vision, Engineering Any-X autonomous system. ASLab Technical Report. 2008.

Birgitte Rasmussen and Kurt E. Petersen. Plant functional modeling as a basis for assessing the impact of management on plant safety. Reliability Engineering and System Safety. 1999.

Morten Lind. Modeling goals and functions of complex industrial plant. Applied Artificial Intelligence, Vol. 8 No. 2, April-June 1994.

Jan Eric Larsson. Diagnosis based on explicit means-ends models. Artificial Intelligence, vol. 80. 1996.

Johannes Petersen. Knowledge based support for situation assessment in human supervisory control. Ph. D. Thesis. Technical University of Denmark. 2000.

José Luis de la Mata and Manuel Rodríguez. Accident prevention by control system reconfiguration. Submitted to ESCAPE 19. 2009.

Mohammad Modarres and Se Woo Cheon. Function-centered modeling of engineering systems using the goal tree – success tree technique and functional primitives. Reliability Engineering and System Safety. 1999.

Mohammad Modarres and Yu Shu Hu. Evaluating system behavior through dynamic master logic diagram (DMLD) modeling. Reliability Engineering and System Safety. 1999.

Atoosa Jashagar. Goal-oriented systems modelling: justification of the approach and overview of the methods. Reliability Engineering and System Safety, vol 64. 1999.

R.S. Peak, R.M. Burkhart, S.A. Friedenthal, M.W. Wilson, M. Bajaj, I. Kim. Simulation-based design using SysML. INCOSE Intl. Symposium, San Diego. 2007.

T.A. Johnson. Integrating models and simulations of continous dynamic system behavior into SysML. Master Thesis. Georgia Institute of Technology. 2008.

The Official OMG SysML site. http://www.omgsysml.org

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Development of integrated functional-structural models

Manuel Rodriguez[a], Ricardo Sanz[a]

[a] *Autonomous System Laboratory- Technical University of Madrid, Jose Gutierrez Abascal, Madrid 28006, Spain*

Abstract

Chemical processes are usually large complex systems. Models concerning the topology and structure have existed for decades and more recently functional models have appeared as a complementary way to describe these systems. In this paper a new modeling formalism to capture structure as well as functionality of a system is presented. Based on the visual formalism higraphs, it has been adapted and extended to be applicable to process systems description. Although it can be used for several applications, this paper focuses on its use in risk analysis.

Keywords: Higraphs, functional modeling, complex systems.

1. Introduction

General systems theory (as introduced in [1], [2]) can be seen as a universal framework for abstract modeling, the philosophy of this theory can be summarized as pointed out by Le Moigne:

"*an object with a set of goals, in a well defined environment, exerts an activity (function) and at the same time experiences how its internal structure evolves through time keeping its identity*".[3]

This idea will be kept in the remaining of the paper along with the idea of model based systems engineering methodology. An extension of the higraph formalism is presented and it will be applied to the development of models that capture the functional (activity) as well as the structural (ontologic) aspects of a process. The inclusion and need of evolutionary (genetic) aspects in a model is not treated in this paper although it could, a priori, be represented using this same formalism.

The methodology described complete some of the existing techniques used for functional analysis (as Multilevel Flow Modeling [4] or Goal Tree Success Tree [5]), and although it is not conceived to be an alternative to the classical modeling environments, a simple translation could be made from one to the other.

The paper is organized as follows: section two explains the higraphs formalism and its extension for process systems, section three applies the proposed methodology to a benchmark process, section four deals with implementation issues and finally section five draws conclusions on the presented work.

2. Integration of functional-structural models

2.1. Higraphs

Higraphs were first presented by Harel in [6]. They are an extension and combination of conventional graphs and Venn diagrams. They constitute a visual formalism of topological nature that can represent set enclosure, exclusion and intersection and the Cartesian product. Higraphs consist of two elements: blobs and edges. Blobs are represented as rounded-corner rectangular shapes. They represent mutual exclusive sets, they may intersect and be arranged in an inclusion hierarchy. Blobs of different levels in the hierarchy may be connected. Edges, represented as arrowed lines, are used to connect blobs. Higraphs can be defined as the result of:

graphs + depth (hierarchy) + orthogonality (Cartesian product). [7]

This visual formalism is well suited for the behavioral specification and design of complex concurrent systems. Another formalism, statecharts, derived from higraphs has been widely used on this purpose for discrete-event systems. In statecharts, blobs represent system states and edges represent the transitions between the states.

Figure 1. illustrates higraphs components and features. It represents the `stopwatch` state of a watch [8].

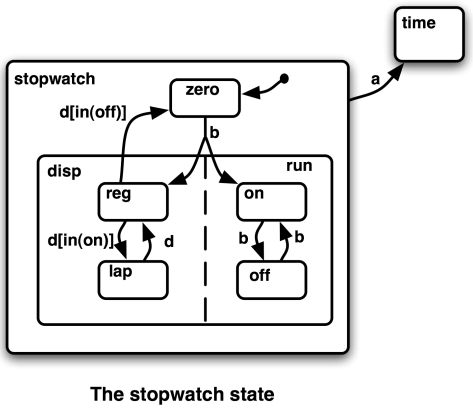


Fig 1. Statechart diagram representing the stopwatch state of a watch.

The `stopwatch` blob (state) has two states, the `zero` state and the `disp/run` state. These states are OR components. The `disp/run` state has a `disp` state and a run state, these are AND states, so both happen at the same time. The `run` state has the `on` and `off` states (which, again, are OR states). Transitions between states are labeled and they can contain a condition. The transition from `reg` to `lap` happens in the event of d, and when the condition (in `on` state) is true.

2.2. D-Higraphs: An extension for process systems

Classic higraphs or statecharts are not useful to specify functional models of process systems because they provide depth in states. When describing the functionality, depth is required for functions (in order to have functional dependency) and system states enable these functions. So, in order to use higraphs for process systems they will have to

be used in a dual way as they are used in statecharts. Blobs will represent now transitions and edges will represent states.

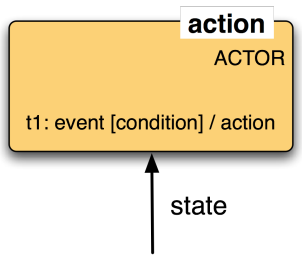


Fig. 2. Basic blob (action), performed by ACTOR when state (event) is enabled and condition is true.

The name of the function will appear in the border of the blob, and the actor (usually equipment) that performs or allows that function will be indicated inside the blob. Transitions are enabled when the state (incoming arrow) is achieved and, as in statecharts, transitions can contain conditions. Firing the transition causes new states, represented by the edges coming out from the blob. Edges represent flow of material (solid line), or energy (dashed line) or information (dotted line). Figure 2 shows the basic components of the D-higraph. Being dual to statecharts, blobs included in a blob are AND blobs, so all the functions have to happen in order for the more global function to happen. Usually OR blobs are not necessary but if needed they would be represented separated by a dashed line. For the edges, an additional representation is added. When the occurrence of a state is necessary for the function, the edge arrow must be solid black, when the event to fire the transition (function) is one of several states (OR condition) the edge arrow is empty. For more complex combinations a black solid bar is added before connecting to the blob. Figure 3 exemplifies some cases.

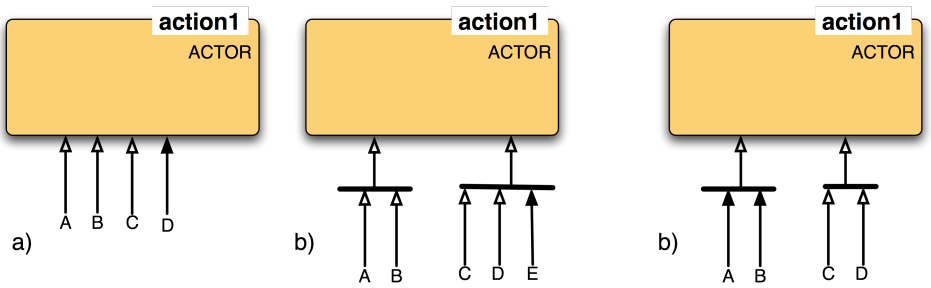


Fig 3. Case a) A or B or C and D must be true to fire action1. Case b) A OR B OR (C OR D AND E) must be true to fire action1. Case c) A AND B or C OR D must be true to fire action1.

A simple example (fig. 4) illustrates what has been explained so far. In this case a Jacketed Continuous Stirred Tank Reactor with two feeds is represented. The overall goal is to produce in operating conditions, this goal is decomposed in `produce` and `remove_heat`. `Produce` (in this kind of reactor) has another function in it, `stir`. As all the arrows are solid, it means that all the states have to be true before the transition is fired. In the case of the CSTR, the `produce` transition will happen if reactants A and B are flowing into the reactor.

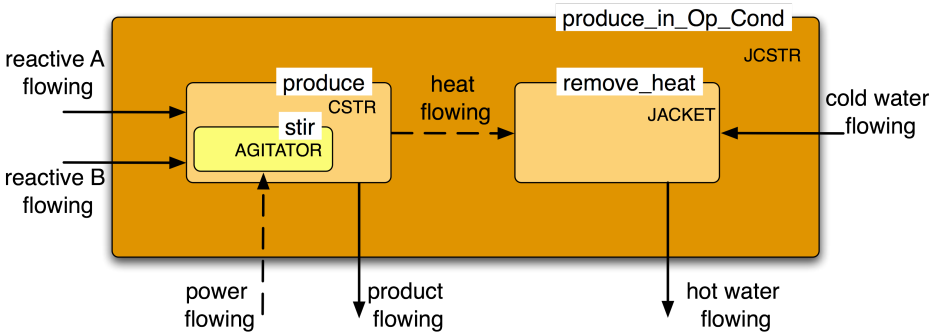


Fig 4. D-Higraph simple example. A Jacketed continuous stirred tank reactor.

3. Application

The Tennessee Eastman Process consists of a reactor, condenser, separator, compressor and stripper with a gas recycle stream, a more detailed description can be found in [9]. This process was initially developed as an industrial plantwide control test problem. In the control structure considered, (on-demand product) we assume that the flowrate of the product stream leaving the base of the stripper is set by a downstream customer as in [10]. Control loops selected are production rate, product quality, reactor temperature, pressure, liquid inventories and two compositions. A ratio is used to feed G and H reactants accordingly to the desired production.

Considering the process description a set of goals and subgoals have been identified in order to guarantee a good plant behavior. All these goals are related with the mentioned control objectives and with guaranteeing plant safety, being the overall goal to produce in safety conditions and with the desired quality. Figure 5 represents the Eastman process using the D-higraph formalism.

It can be observed that control loops are represented in a straightforward way including advanced configurations as cascade control or ratio control. In this case, both configurations are present in the feed to the reactor. The process can be followed without difficulty even without having the classic flowsheet representation.

Risk analysis can be easily performed. For example, if `heat transmission` state from the reactor to the jacket fails, then it causes the `remove heat` function to fail and, consequently, the `react in op. conditions` does fail which implies that, finally, the overall goal `produce safely and with desired quality` is not accomplished. Notice that the failure of the `heat transmission` does not affect the `produce G,H` function of the reactor as this state is not an input to that blob and therefore it is not a necessary condition to fire it. In the other hand, in the cases of the heat exchanger and the vaporizer `heat transmission` affects both functions (`vaporize` and `heat` and `cold` and `heat`) as showed with the double arrowed line.

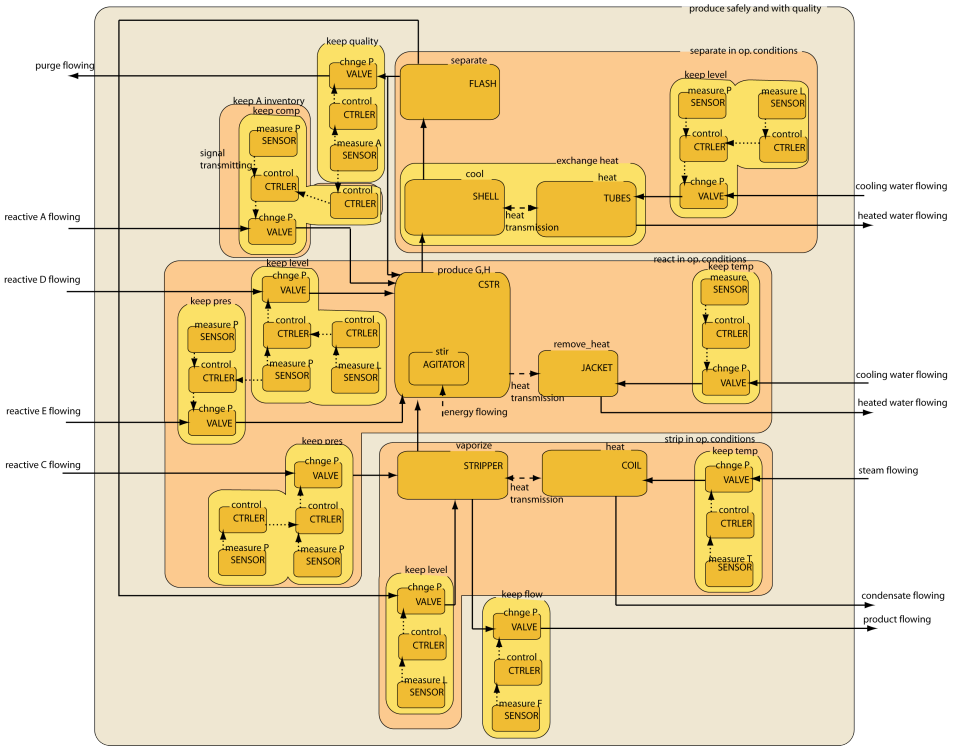


Fig 5. D-Higraph diagram of the Tennesse Eastman process.

4. Implementation

The implementation will be based on a model driven approach for systems engineering. of the visual formalism D_Higraphs. To support this model based design and implementation, sysML (Systems Modeling Language) will be used as it is suited for modeling different aspects (functional/behavioral, structural, operational) of large scale engineering projects. SysML is an extension of the Unified Modeling Language (UML) and it is specified in [11]. SysML models can be transformed into models used by simulation languages.

Figure 6 shows the implementation diagram. Model development will be based on existing ontologies, namely a domain -structural- ontology (as can be ONTOCAPE [12] for the process systems domain) and a functional ontology. The built (D-higraph) model can be mapped to a sysML model (with different views) that can be itself transformed to a model for many (simulation) languages. For example an interface definition language model (.idl extension) can be generated from the sysML description and this can be embedded in a simulator as a Cape Open component. A direct mapping from the D-higraph model to a commercial modeling environment is as well possible.

Following the systems engineering approach a product will be generated from the requirements specification using the commented components.

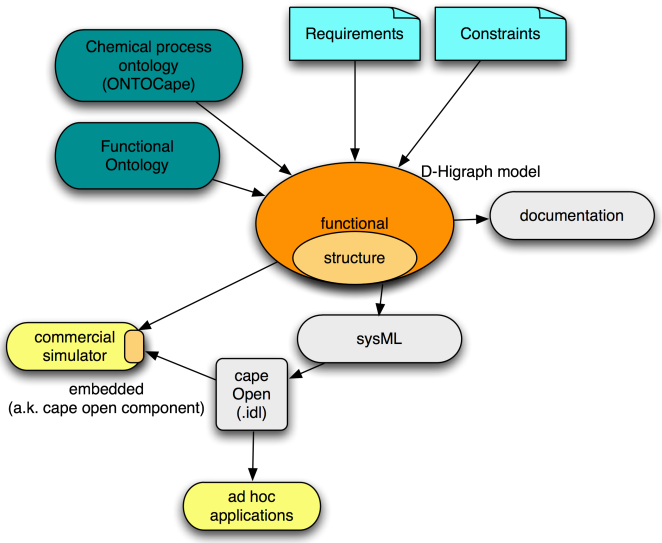


Fig 6. Implementation diagram. Model based engineering.

5. Conclusions

In this paper a new formalism has been described, based on higraphs it extends and adapts them to be used for process systems description. Structure and functionality are captured by the resulting diagram, which can be understood without many additional explanations. Not only process information but control information as well is displayed in the diagram. The diagram can be used for multiple purposes from fault diagnosis or risk analysis, to process documentation although only risk analysis have been outlined in this paper. Being a layered diagram, different views can be displayed, allowing for more abstract or specific views, or for just a functional or structural view. Further work includes representing (in a similar fashion as in existing commercial simulators) quantitative models, mapping ontologies to models and integrating the components in a single architecture with additional features, mainly cognitive.

References

1. Von Bertalanffy, 1968, L. General Systems Theory. New York, George Braziller.
2. G. Klir, 1969, An Approach to General Systems Theory, Van Nostrand Reinhold, NY.
3. J.L. Le Moigne, 1990, La théorie du système général, 3 ed, Paris, PUF.
4. M. Lind, 1994, Modeling goals and functions of complex industrial plant, AAI , 8,
5. M. Modarres, 1999, Function centered modeling of engineerig systems using the goal tree-success tree technique and functional primitives. Reliab. Eng. and Sys. Safety, 64
6. D. Harel, 1988, On visual formalisms, Comm. Assoc. Comput. Mach., vol 3, 514-530,
7. D. Harel et al., 2002, An algorithm for blob hierarchy layout, Visual computer.
8. D. Harel, 1987, Statecharts: A visual formalism for complex systems. Sci. Comput. Program., 8(3):231–274.
9. J. J. Downs and E. F. Vogel, 1993, A plant-wide industrial process control problem. Computers and Chemical Engineering 17 , pp. 245–255.
10. W. L. Luyben et al. , 1998, Plantwide Process Control. McGraw-Hill.
11. OMG, 2006, OMG Systems Modeling Language Specification.
12. J. Morbach, et al, 2007, OntoCAPE – a large-scale ontology for chemical process engineering, Engineering Applications of Artificial Intelligence 20:2, pp 147–161

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Fast and Systematic Procedure to Develop Dynamic Models of Bioprocesses – Application to Microalgae Cultures

Johan Mailier and Alain Vande Wouwer

Control Department, Polytechnic Faculty - University of Mons
31 Boulevard Dolez, 7000 Mons, Belgium.

Abstract

The purpose of this paper is to report on the development of a procedure for inferring black-box, yet biologically interpretable, dynamic models of bioprocesses based on sets of measurements of a few external components (biomass, substrates, products of interest). The procedure has three main steps: (a) the determination of the number of macroscopic biological reactions linking the measured components, (b) the estimation of a first reaction scheme, which has interesting mathematical properties, but might lack a biological interpretation, and (c) the "projection" (or transformation) of this reaction scheme onto a biologically-consistent scheme. The advantage of the method is that it allows the fast prototyping of models for the culture of micro-organisms that are not well documented. The good performance of the third step of the method is demonstrated on an application example of vegetal cell culture.

Keywords: reaction networks, mathematical modeling, parameter estimation, bioprocesses

1. Introduction

The increasing demand in process monitoring, control, and optimization encourages the development of various techniques for modeling bioprocesses. Among these techniques, the one exposed by Bastin and Dochain [1990] relies on the concept of macroscopic reaction schemes. These schemes consist of a set of reactions that represent simple relations between key-components like substrates, biomass, and products of the process. The general form of a macroscopic reaction scheme is

$$\sum_{i\in\mathcal{R}_k}(-\nu_{ik})\xi_i \xrightarrow{\varphi_k} \sum_{i\in\mathcal{P}_k}\nu_{jk}\xi_j, \quad k\in[1,M], \tag{1.1}$$

where $\mathcal{R}_k$ and $\mathcal{P}_k$ are respectively the sets of reactants and products of reaction k, φ_k is the k^{th} reaction rate expressed as a function of the reacting species concentrations, and ν_{*k} are the yield (or pseudo-stoichiometric) coefficients of reaction k. This simple formulation provides a dynamic model of the process that requires no knowledge on its metabolic pathways, yet preserving some biological meaning. In the absence of gaseous outflow, a mass balance applied to (1.1) gives

$$d\xi(t)/dt = K\varphi(\xi) - D(t)\xi(t) + Fe(t), \tag{1.2}$$

where $\xi(t) \in \mathbb{R}^N$ is the vector of species concentrations, $K \in \mathbb{R}^{M \times N}$ the pseudo-stoichiometric matrix, $\varphi(t) \in \mathbb{R}^M$ the reaction rate vector, $D(t) \in \mathbb{R}$ the input dilution rate, and $Fe(t) \in \mathbb{R}^N$ the external feed rates.
The considered problem is to use the measurements of $\xi(t)$ in order to identify the elements ϑ_K of the pseudo-stoichiometric matrix K, and the parameters ϑ_φ of the kinetic function $\varphi(\xi)$, the structure of which is assumed to be known as soon as the number of reactions M is known.

2. Robust Methodology of identification

In this study, the assumption is made that the whole vector $\xi(t)$ is measured at discrete times. A continuous version of the measurements can be obtained through simple filtering and interpolation.

2.1. Number of Reactions

As described by Bernard and Bastin [2005a, 2005b], the number of reactions M needed by the model to reproduce the dynamic observations is successfully estimated by principal component analysis (PCA) applied to the measurements, possibly filtered and interpolated. In the sequel, this first stage is assumed to be completed, so that the number of reactions M is known.

2.2. Identification of C-Identifiable Schemes

A systematic data-driven method has been suggested by Hulhoven et al. [2005] to identify those reaction schemes that satisfy the mathematical property known as C-identifiability. This procedure relies on a state transformation, operated via the partitioning of the pseudo-stoichiometric matrix K into matrices $K_a \in \mathbb{R}^{M \times M}$ and $K_b \in \mathbb{R}^{(N-M) \times M}$. The dynamics of the transformed state vector $z = CK_a + K_b$ (with matrix $C \in \mathbb{R}^{(N-M) \times M}$ chosen so that $CK_a + K_b = 0$) is independent of the reaction rates $\varphi(\xi)$. Moreover, the dynamics of z involves C, which can then be estimated from the measurements, independently of the kinetics. Those reaction schemes whose matrix K uniquely derives from the knowledge of C, are called C-identifiable (Chen and Bastin [1996]).

The procedure of Hulhoven et al. [2005] forces the C-identifiability of the identified model, by testing specific values of K_a, and selecting the one that gives the lowest cost function after identification of the matrix C. Matrix K is then simply derived from K_a and C.

However, the majority of the real bioprocesses are not C-identifiable, so that there is little interest in determining an "optimal" K_a, which anyway would lead to a reaction scheme lacking biological interpretability. Therefore the present approach simply sets K_a equal to the identity matrix I_M for identifying the matrix C independently of the kinetics (Bogaerts et al., 2008). In the sequel, it is assumed that C has been identified, so that the model now writes

$$d\xi(t)/dt = \begin{pmatrix} I_M & -C^T \end{pmatrix}^T \varphi_C - D(t)\xi(t) + Fe(t), \tag{2.1}$$

where φ_C is a reaction rate vector which may not be biologically consistent (for instance it could be negative). However, it can be estimated from measurements. The introduction of biological consistence into the model is left to the next stage of the

procedure, which does so by identifying the matrix K_a that projects the model (2.1) onto the following biologically consistent model, i.e. this model satisfies biological constraints on its pseudo-stoichiometric and kinetic parameters:

$$d\xi(t)/dt = \begin{pmatrix} K_a^T & -K_a^T C^T \end{pmatrix}^T (K_a^{-1}\varphi_C) - D(t)\xi(t) + Fe(t). \tag{2.2}$$

Note that following Delcoux et al. [2001], (2.1) and (2.2) are equivalent models because they describe the same time-evolution of the state vector $\xi(t)$.

2.3. Identification of Biologically Relevant Schemes

This last stage, to which the main contribution of the present study is devoted, can itself be decomposed into two intermediate steps: a quick estimation step, followed by a refining step that consists of a classical maximum-likelihood identification of all the parameters, starting from the values obtained by the quick estimation. Only the first step is described in this paper, since it is precisely the contribution of this work.

The purpose is to find K_a together with the kinetic parameters ϑ_φ, so that the projected reaction rates have a relevant biological structure: $K_a^{-1}\varphi_C = \varphi(\xi, \vartheta_\varphi)$. To do so quickly, (2.2) is rearranged by grouping the measured signals (i.e. $d\xi/dt + D\xi - Fe$, possibly filtered and interpolated) column-wise in the left-hand side. Doing so for the other signals in the right-hand side gives the following matrix equation:

$$U = \begin{pmatrix} I_M & -C^T \end{pmatrix}^T K_a(\vartheta_K)\Phi(\vartheta_\varphi), \tag{2.3}$$

from which a least-square estimator of K_a is built as a function of ϑ_φ (via Φ):

$$\widehat{K}_a = (I_M + C^T C)^{-1}(\mathrm{I}_M \quad -C^T)\Phi^T(\Phi\Phi^T)^{-1}. \tag{2.4}$$

Biological constraints on ϑ_K can also be added, using a classical constrained linear optimization for the estimation of $K_{\hat{a}}$

The quick estimation step then consists of a (constrained) nonlinear optimization of ϑ_φ that aims at minimizing the following cost function built from (2.3) and (2.4):

$$\mathcal{J} = \sum_i \sum_j \left(U - (I_M \quad -C)\widehat{K}_a\Phi \right)_{i,j}^2. \tag{2.5}$$

3. Simulation Example

The continuous cultivation of microalgae in a photobioreactor is considered, where growth is limited by a single nutrient. Attention is focused on the third stage of the identification procedure (consisting of a quick parameter estimation without integration of the model equations, and of a final global nonlinear identification) and its comparison with a standard identification of all the parameters in a single step.

3.1. Macroscopic Model

Plant cells generally behave differently from animal cells. Indeed, rather than directly depending on the external substrate concentrations, their growth rate is bound to the quantity of stored nutrients. The concept of nutrient quota, i.e. the quantity of nutrient per biomass unit, has therefore been widely used in algal growth models, like the well-

known Droop model. However, these quotas do not fit well in the formalism of the general dynamic model, which rather uses absolute quantities. Therefore, algal growth should be modeled in terms of total amount of stored nutrients, i.e. multiplying quotas by the biomass.

The considered model has been developed by Lemesle and Mailleret [2008], as a mechanistic reformulation of the Droop model. The first reaction is the storage of extracellular substrate *S* into the cell, and the second is the metabolism of stored nutrients *R* to produce biomass *X* and metabolized nutrients *B*:

$$\begin{aligned} &1S \xrightarrow{\varphi_1} 1R && \text{(substrate storage)},\\ &0.5R \xrightarrow{\varphi_2} 0.5B + 1X && \text{(biomass growth)}. \end{aligned} \tag{3.1}$$

with the reaction rates $\varphi_1 = S/(S+15)\,X$ and $\varphi_2 = (R/X)/(1+R/X)X$. A simple mass balance on (3.1) gives the following matrix system of differential equations:

$$\begin{pmatrix} dS/dt \\ dR/dt \\ dB/dt \\ dX/dt \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} \vartheta^*_{K_{1,1}} & \vartheta^*_{K_{2,1}} \\ \vartheta^*_{K_{1,2}} & \vartheta^*_{K_{2,2}} \end{pmatrix} \begin{pmatrix} \dfrac{S}{S+\vartheta^*_{\varphi_1}} X \\ \dfrac{R/X}{\vartheta^*_{\varphi_2}+R/X} \end{pmatrix} - D \begin{pmatrix} S \\ R \\ B \\ X \end{pmatrix} + D \begin{pmatrix} S_{in} \\ 0 \\ 0 \\ 0 \end{pmatrix}, \tag{3.2}$$

with $\begin{pmatrix} \vartheta^*_{\varphi_1} & \vartheta^*_{\varphi_2} & \vartheta^*_{K_{1,1}} & \vartheta^*_{K_{2,1}} & \vartheta^*_{K_{1,2}} & \vartheta^*_{K_{2,2}} \end{pmatrix}^T = (15 \;\; 1 \;\; -1 \;\; 1 \;\; 0 \;\; -0.5)^T$.

3.2. Simulated Experimental Data

Four experiments are made, with a constant dilution rate $D = 0.1$, starting from the following measured initial conditions: $\xi_{0,1} = (2 \;\; 0.75 \;\; 5.9 \;\; 3)^T$, $\xi_{0,2} = (30 \;\; 0.25 \;\; 5.3 \;\; 2.6)^T$, $\xi_{0,3} = (10 \;\; 0.5 \;\; 2.6 \;\; 1.3)^T$, and $\xi_{0,4} = (4 \;\; 1 \;\; 2.2 \;\; 1)^T$.

Sixty equally-spaced measurements of $\xi(t)$ are made for each experiment, from $t = 0$ to $t = 60$. Gaussian noise is then added, with a constant relative standard deviation $\sigma = 0.05^{1/2}$. Interpolation is performed by the Matlab function *csaps*, with a smoothing parameter $p = 0.9$, while filtering is achieved by a moving average filter with a five-measurement moving window.

3.3. Identification Results

Fifty initial guesses (ten for each value of $\Delta\vartheta$) are generated using the formula $\vartheta_0 = \vartheta^*(1+\Delta\vartheta.\alpha)$, where α is a standard normally distributed random vector. The first step of the optimization is carried out by a Fletcher version of the Levenberg-Maquardt algorithm (Fletcher [1971]), while direct one-step optimizations use the Nelder-Mead simplex method (Matlab function *fminsearch*).

Figure 1 illustrates the repartition of the relative distances between the estimates and their corresponding optimum values, for ϑ_{φ_1} (left column) and ϑ_{φ_2} (right column). Compared to direct identification, excellent performances are obtained with the new method, since the estimates are always close to their optimum value. In this specific example, the second step of the new algorithm does not improve the estimation provided by the first step, due to the limited information content of the measurements.

Because there is no numerical solution of the model equations, the quick estimation step of the new method takes only a few seconds, which is much faster than the few minutes of the direct identification. Figure 2 illustrates these observations.

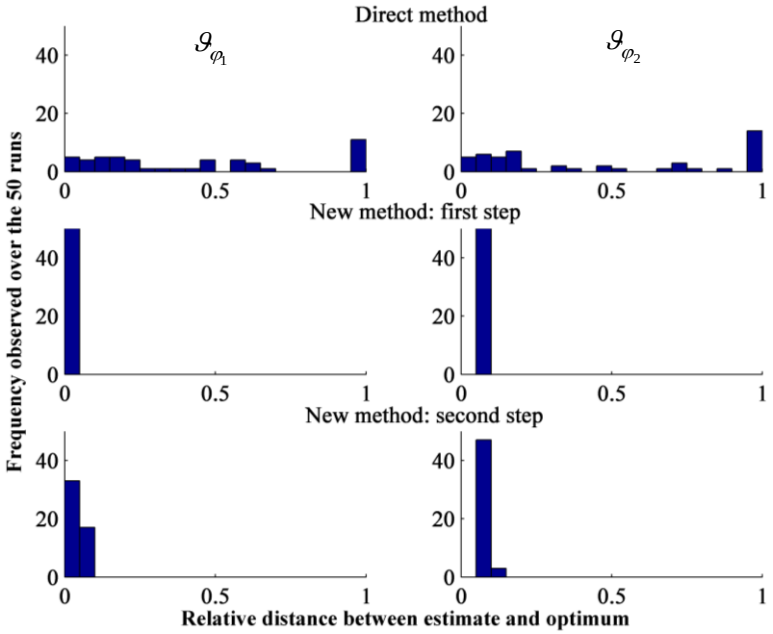


Figure 1. Histogram of the relative distance between estimates and their corresponding optimal values, after direct identification (top), quick estimation (middle), and final identification (bottom)

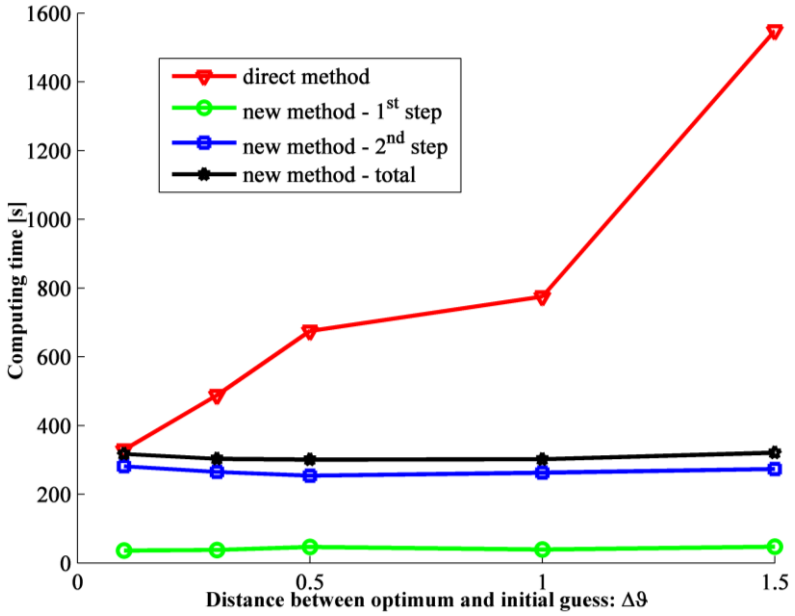


Figure 2. Mean computing time as a function of the distance $\Delta\vartheta$ between the initial guesses and the optimum parameters.

4. Conclusions

The systematic identification procedure developed in this paper uses and extends previous research results to the projection of C-identifiable schemes onto biologically consistent schemes. The strong point of the new approach mainly lies in the quick estimation step of the projection stage. Indeed, it has a low computation time because it integrates no differential equation, and also because it operates on a limited number of parameters, i.e. the set of kinetic parameters. Moreover, a simulation example of continuous microalgae culture has shown that the new method is more robust than direct global identification with regard to uncertainties on initial guesses.

Acknowledgements

Johan Mailier is a Research Fellow by the Fonds de la Recherche Scientifique (FNRS – Belgium). This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.

References

G. Bastin and D. Dochain, 1990, On-line Estimation and Adaptive Control of Bioreactors, Elsevier, New York.

O. Bernard and G. Bastin, 2005a, On the estimation of the pseudo-stoichiometric matrix for macroscopic mass balance modelling of biotechnological processes, Mathematical Biosciences, 193:51-77.

O. Bernard and G. Bastin, 2005b, Identification of reaction networks for bioprocesses: determination of an incompletely known pseudo-stoichiometric matrix, Bioprocess and Biosystems Engineering, 27(5):293-301.

Ph. Bogaerts, M. Rooman, V. Vastemans, and A. Vande Wouwer, 2008, Determination of macroscopic reaction schemes: towards a unifying view, Proceedings of the 17th World Congress of IFAC, Seoul, Korea.

L. Chen and G. Bastin, 1996, Structural identifiability of the yield coefficients in bioprocess models when the reaction rates are unknown, Mathematical Biosciences, 132:35-67.

J.-L. Delcoux, R. Hanus, and Ph. Bogaerts, 2001, Equivalence of reaction schemes in bioprocesses modelling, Proceedings of the 4th International Symposium on Mathematical Modelling and Simulation in Agricultural and Bio-Industries, Haifa, Israel.

R. Fletcher, 1971, A modified Marquardt subroutine for nonlinear least squares, Technical Report R-6799, Theoretical Physics Division, AERE Harwell.

X. Hulhoven, A. Vande Wouwer, and Ph. Bogaerts, 2005, On a systematic procedure for the predetermination of macroscopic reaction schemes, Bioprocess and Biosystems Engineering, 27(5):283-291.

V. Lemesle and L. Mailleret, 2008, A mechanistic investigation of the algae growth "Droop" model, Acta Biotheoretica, 56(1-2):57-102.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Reliability vs. efficiency when solving multiphase equilibrium problems with hybrid optimization codes

Paula B. Staudt and Rafael de P. Soares

Department of Chemical Engineering, Federal University of Rio Grande do Sul – UFRGS, Rua Engenheiro Luis Englert, s/n, CEP 90040-040, Porto Alegre - Brasil

Abstract

Computation of phase equilibrium is a very important and frequently encountered problem in process systems engineering. Accurate and robust flash routines are at the core of the major part of chemical engineering design applications, ranging from pipelines to distillation columns, chemical reactors, and oil and gas production. The multiphase equilibrium calculation by the direct minimization of the Gibbs free energy has some potential advantages over classical methods, mainly because the equality of chemical potentials is a necessary but not a sufficient condition for equilibrium. Recent developments in global optimization and the availability of very fast computers have stimulated the research on the direct minimization approach. In this work, a hybrid method (genetic algorithm for the global search and interior point for refinement) was implemented and the efficiency/reliability relation was studied. For some systems the global minimum can be found in a NFE compatible with the methods currently available in process simulators. But, it was also found that, for some systems, there is an asymptotic behavior. Apparently, a reliability of 100 %, when using global search algorithms with some random nature, would be found only for an infinite number of function evaluations.

Keywords: Global optimization, Gibbs minimization, multiphase equilibrium.

1. Introduction

Phase equilibrium calculations and phase stability problems have to be solved a very large number of times in the design and analysis of chemical processes. For a system with specified components, composition, temperature and pressure, a phase equilibrium problem involves the calculation of number of moles of each phase and its composition at equilibrium whereas phase stability analysis determines the stability of the system (Srinivas and Rangaiah, 2007).

Today, the availability of fast computers has stimulated the research of reliable but efficient methods for resolving phase equilibrium calculations. Methods for solving phase equilibrium problems can be classified into two main categories: equation-solving and direct minimization of the Gibbs free energy. A minimum in the Gibbs free energy implies that chemical potential, activity, and fugacity of each component must be equal in all phases of the system. In the equation-solving approach this isoactivity (or fugacity matching) conditions coupled with mass balances and mole fraction summations for an assumed number of phases are solved numerically (Teh and Rangaiah, 2002). Even though, equation solving approach seems to be faster and simple, the solution obtained by this method may not correspond to the true minimum of Gibbs free energy function.

A major drawback is that the equality of chemical potentials represents only the stationary point conditions for the minimum. Any method based on these conditions may fail to find the global minimum, being attracted by local minima. Also, it needs *a priori* knowledge of phases coexisting at equilibrium (Srinivas and Rangaiah, 2007). Then it is usual to combine equation-solving and phase stability algorithms to determine the number of existing phases and their composition.

A second approach, which has gained much attention recently, is the treatment of the phase equilibrium problem by the direct minimization of Gibbs free energy (Nichita et al., 2002). This optimization problem consists in the minimization of the overall system Gibbs free energy:

$$\frac{G(n)}{RT} = \sum_{k=1}^{N_p} \sum_{i=1}^{N_c} n_i^k \, ln f_i^k \tag{1}$$

satisfying the appropriate material balance constraints:

$$\sum_{k=1}^{np} n_i^k = n_i \qquad i = 1,2,3, \dots, N_c \tag{2}$$

and bound constraints:

$$0 \leq n_i^k \leq n_i \quad i = 1,2, \dots, N_c \quad k = 1,2, \dots, N_p \tag{3}$$

where n_i is the overall (or feed) number of moles of the component i, n_i^k and f_i^k are the number of moles and the fugacity of the component i in the phase k, respectively. The decision variables for minimizing the free energy are n_i^k. The maximum number of phases at equilibrium N_p is considered to be known. If N_p is greater than the actual number of phases at equilibrium, the solution will contain empty phases (all n_i^k zero for a phase k) or identical phases (phases with nonzero n_i^k values but identical composition). Identical phases need to be identified and agglutinated into a unique phase in a post processing step.

The constrained minimization problem (Equations 1-3) can be simplified to a bound-constrained problem replacing the decision variables n_i^k by β_i^k (for $i=1,2,\dots,N_c$; $k=1,2,\dots,N_p-1$). The new variables, β_i^k are bounded between 0 and 1 and are related to n_i^k by:

$$n_i^1 = \beta_i^1 n_i \tag{4}$$

$$n_i^k = \beta_i^k \left(n_i - \sum_{j=1}^{k-1} n_i^j \right) \quad i = 1,2, \dots, N_c \quad k = 2, \dots, N_p - 1 \tag{5}$$

The equality constraints are eliminated and the number of decision variables is reduced from N_c by N_p to N_c by N_p-1. The optimization problem (Equation 1) with the decision variables β_i^k is a bound-constrained problem. This problem is much more suited for the solution using stochastic methods than the constrained optimization problem (Equations 1-3). Regardless the formulation (constrained or bound-constrained), the difficulty of phase equilibrium calculations using the direct optimization approach comes from the highly nonlinear and non-convex form of the objective function (the fugacity coefficients in Equation 1) that gives no guarantee that the global minimum will be found.

In this work, the solution of multi-phase equilibrium problem by the minimization of the bound-constrained optimization problem by a hybrid algorithm (global + local) is studied. Our objective is to obtain an optimum balance between global and local search reducing the number of function evaluations in the global search step as much as possible without sacrificing the reliability.

2. Robustness and Reliability

In order to solve a multi-phase equilibrium calculation by direct minimization of the Gibbs free energy, an optimization problem needs to be solved. The objective function of this optimization problem is highly nonlinear and non-convex. The complexity of the problem increases near critical points and phase boundaries. In order to reach the global minimum a global optimization strategy is necessary, but it is common sense that global optimizers are very computationally expensive.

In many practical applications, heuristic stochastic methods have been successfully used. These algorithms do not use any gradient information and are usually initialized with random values. Examples are: genetic algorithms, tabu-search, simulated annealing, tunneling and differential evolution. This class of methods has the property to descend the objective function value very fast at the beginning, but as they lose diversity information of the space, they tend to be extremely slow to actually get the global solution (Nichita et al., 2002). As a consequence, the global minimization tends to be very computationally intensive. Because of this, the use of hybrid procedures is becoming more frequent in literature (Teh and Rangaiah, 2003; Balogh et al., 2003; Srinivas and Rangaiah, 2007).

In this work, we solve the optimization problem in two steps. A genetic algorithm (GA) *evolves* a *population* of fixed size (randomly initialized) for a fixed number of evolutions. The *individual* with the best objective function is used as the starting point for an interior point algorithm which does the local optimization. The GA implementation used was the open-source package JGAP (http://jgap.sourceforge.net). For the local search the interior point open-source optimization package IPOPT from Wächter and Biegler (2006) was used. The first-order derivatives of the objective function required by the local minimizer were obtained analytically.

GA implementations usually do not have a convergence criteria, the user just select a *population* size and a number of *evolutions* to be executed. In our approach, the GA *fittest individual* (the best point found) is the starting point to the local optimization method (that uses gradient information). The objective is to refine the solution of the global optimization in few extra function evaluations.

The most important aspect of this kind of approach is how it balances global *vs*. local searches. Emphasis on global search increases the number of function evaluations significantly. While emphasis on local search can decrease the total number of function evaluations, it usually decreases the reliability in finding the global solution (Saber and Shaw, 2008). When using GA the emphasis on the global search is mainly controlled by the population size and the number of evolutions.

In order to discover the reliability/efficiency relation for our hybrid implementation several experiments varying the population size and number of evolutions were executed for the Problem 1 (Table 1). The reliability (RE) is the rate of success in finding the true global minimum of the problem and it is obtained by running the same problem a certain number of times. The term efficiency is measured as the inverse of the number of function evaluations (NFE) required for the solution (including both steps global and local). The Problem 1 is known to have a global minimum very difficult to

find. To compute the reliability, 2500 independent optimizations were executed for each pair of population and evolutions. The results can be seen in Figure 1.

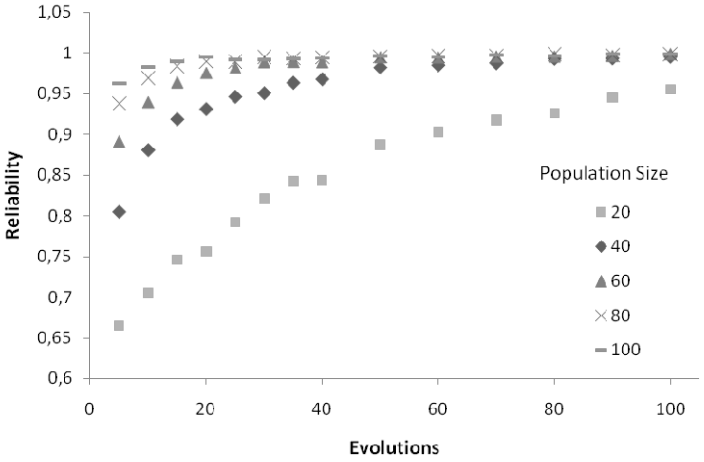


Figure 1: Variation of reliability with number of evolutions and population for Problem 1

The different curves illustrate the variation of the number of evolutions for a fixed population. As expected, large population sizes and large number of evolutions produces larger reliabilities but an asymptotic behavior is noticed (no experiment in Figure 1 actually reached reliability equals 1). This indicates that no matter how big is the population size or the number of evolutions, there will always be a small probability of missing the true global minimum.

In Figure 2 we can observe the relation between NFE and reliability. As can be seen, there is room for a great improvement in reliability when the NFE is small. But in the vicinity of 100% reliable, better results can be obtained only by a great increase in the NFE.

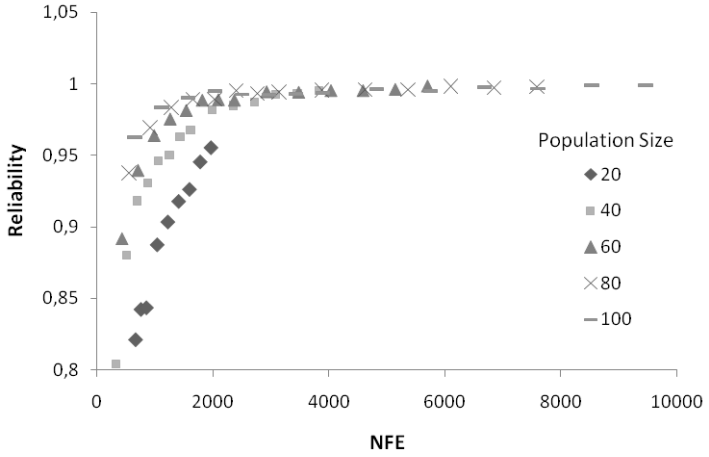


Figure 2: Reliability vs. NFE for Problem 1, including all tests of Figure 1.

This means a great additional effort for a very small gain, because no matter how large is the NFE, there will always be a small probability of failure in finding the true global minimum. Large population sizes and number of generations are very often in the literature, for instance in Teh and Rangaiah (2003), they are set to 200 and 100,

respectively. From the results in Figure 2, we can conclude that this is a conservative set of parameters. Still, according to the Figure 2, we can see that in the low reliability region, a bigger population size offers a better reliability in less NFE. This relation is not clearly observed when reliability around 100 %. Thus, we decided to set the population size as 50 and the number of evolutions as 30 for solving the benchmark problems in Section 3.

3. Results with Benchmark Problems

Using the optimum set for the population size and number of evaluations obtained in the previous section, the performance of our hybrid implementation was tested using benchmark problems drawn from the literature.

Table 1: Benchmark problems used to test the implementation.

	Mixture	**Model**	**Eq.**	**Reference**
1	Toluene, water, aniline	NRTL	LLE	McDonald and Floudas (1995)
2	Hydrogen sulfide, methane	SRK	VLE	Hua et al. (1998)
3	N-butyl acetate, water	NRTL	LLE	McDonald and Floudas (1996)
4	Nine hydrocarbons	SRK	VLE	Rangaiah (2001)
5	Nitrogen, methane, ethane	PR	LLE	Hua et al. (1998)
6	Benzene, acetonitrile, water	NRTL	VLLE	McDonald and Floudas (1995)

The efficiency of our implementation was evaluated through the NFE necessary to achieve the solution of the minimization problem. The NFE of each example above were compared to those found by Teh and Rangaiah (2003) with GA and Tabu search for global search and a quasi-Newton method for local refinement, referred as GAQN and TSQN, respectively. Also, a comparison was made with the differential evolution (DE) approach in Srinivas and Rangaiah (2007) and simulated annealing (SA) results from Rangaiah (2001). In Teh and Rangaiah (2003), the number of generations used was 100 and the population size equal to 200. The reliability number was found running only 25 times each example. In this work the reliability was computed running 100 times each problem.

Table 2: Efficiency comparison with other algorithms.

	This work		***GAQN***		***TSQN***		***DE***		***SA***	
	NFE	**RE**	**NFE**	**RE**	**NFE**	**RE**	**NFE**	**RE**	**NFE**	**RE**
1	**1493**	**99%**	20069	96%[a]	1719	96%[a]	-	-	-	-
2	**1320**	**99%[b]**	20024	80%	1187	96%[a]	-	-	-	-
3	1336	73%	20018	96%[a]	**1425**	**96%[a]**	7600	99%[b]	34563	99%[b]
4	**2855**	**82%**	20515	96%[a]	10040	96%[a]	-	-	134767	99%[b]
5	1449	87%	20238	96%[a]	**1777**	**96%[a]**	-	-	-	-
6	**1991**	**99%[b]**	20212	96%[a]	5615	96%[a]	-	-	-	-

[a] 100% in 25 runs [b] 100% in 100 runs

The NFE shown is the total number of function evaluations to solve the problem, including both global and local search steps. As can be seen in Table 2, our implementation presents a very good reliability in much less function evaluations for most problems. Even for a bigger problem like the example 4, with nine components, the NFE was much smaller than the other implementations. Reliability larger than 82 % can be achieved for this problem if the population size or number of evolutions is

increased, this can be done without much performance penalty. For Problems 3 and 5 Tabu search shown a better reliability with NFE only slightly greater.

4. Conclusions

The correct prediction of multi-phase equilibrium problems is a very important task. The minimization of Gibbs free energy is a promising approach to solve this class of problems. There are many methods to solve the problem and hybrid algorithms have a great advantage because they merge global search and local search becoming more efficient and robust. These two aspects, efficiency and robustness, are closely connected and have to be analyzed together. The reliability of a method is inversely proportional to the efficiency, thus a balance is necessary to get the better performance.

When a stochastic method is used, like genetic algorithm, the reliability analysis requires some attention. Because of the random character of the method, two independent optimization runs can lead to different solutions. Even if a first execution succeeded to find the global minimum, there is no guarantee that the next independent run will. This effect is more evident if a complicated problem is solved. Analyzing the results of a phase equilibrium problem with a very difficult to find global minimum, it was found that a reliability of 100 % would require an infinite number of function evaluations. Besides, after a certain number of function evaluations (which is problem dependent) an increase in the size of the optimization problem (populations and evolutions) does not guarantee a better reliability; just make the approach less efficient.

References

Balogh, J., Csendes, T., Stateva, R. P., 2003, Application of a stochastic method to the solution of the phase stability problem: cubic equations of state, Fluid Phase Equilibria, 212 (1-2), 257-267.

Hua, J. Z., Brennecke, J. F., Stadtherr, M. A., 1998, Enhanced interval analysis for phase stability: Cubic equation of state models, Industrial & Engineering Chemistry Research 37 (4), 1519-1527.

McDonald, C. M., Floudas, C. A., 1995, Global optimization for the phase and chemical equilibrium problem: application to the NRTL equation, Computers & Chemical Engineering 19 (11), 1111-1139.

McDonald, C. M., Floudas, C. A., 1996, Glopeq: A new computational tool for the phase and chemical equilibrium problem, Computers & Chemical Engineering 21 (1), 1-23.

Nichita, D. V., Gomez, S., Luna, E., 2002, Multiphase equilibria calculation by direct minimization of gibbs free energy with a global optimization method, Computers & Chemical Engineering 26 (12), 1703-1724.

Rangaiah, G. P., 2001, Evaluation of genetic algorithms and simulated annealing for phase equilibrium and stability problems, Fluid Phase Equilibria 187-188, 83-109.

Saber, N., Shaw, J. M., 2008, Rapid and robust phase behaviour stability analysis using global optimization, Fluid Phase Equilibria 264 (1-2), 137-146.

Srinivas, M., Rangaiah, G. P., 2007, A study of differential evolution and tabu search for benchmark, phase equilibrium and phase stability problems. Computers & Chemical Engineering 31 (7), 760-772.

Teh, Y. S., Rangaiah, G. P., 2002, A study of equation-solving and gibbs free energy minimization methods for phase equilibrium calculations. Chemical Engineering Research and Design 80 (7), 745-759.

Teh, Y. S., Rangaiah, G. P., 2003, Tabu search for global optimization of continuous functions with application to phase equilibrium calculations. Computers & Chemical Engineering 27 (11), 1665-1679.

Wächter, A., Biegler, L. T., 2006. On the implementation of an interior-point filter linesearch algorithm for large-scale nonlinear programming. Mathematical Programming 106 (1), 25-57.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Thermohydraulic Simulation of Heat Exchanger Networks Subjected to Fouling

Viviane B. G. Tavares[a], André L. H. Costa[a], Eduardo M. Queiroz[b], Fernando L. P. Pessoa[b], Fábios S. Liporace[c], Sérgio G. Oliveira[c]

[a] *Rio de Janeiro State University – UERJ, Institute of Chemistry - Department of Industrial Operations and Designs, Rua São Francisco Xavier, 524, Maracanã, Rio de Janeiro CEP 20550-900, Brazil*
[b] *Federal University of Rio de Janeiro – UFRJ, School of Chemistry - Department of Chemical Engineering, Av. Horácio Macedo, 2030, CT, Bl. E, Cidade Universitária, Rio de Janeiro CEP 21941-909, Brazil*
[c] *PETROBRAS/CENPES/PDEDS/Gás Natural/Célula de Otimização, Av. Horácio Macedo, 950 - Cidade Universitária, Rio de Janeiro, 21949-915, Brazil*

Abstract

Fouling is a complex problem which affects thermal equipment. The deposits over the heat exchanger surface diminish the thermal effectiveness and increase flow resistances. The consequences of fouling are related to important economic penalties, such as, more expensive equipment with larger thermal surfaces, higher energy consumption costs (heat/power), shutdown costs for exchanger cleaning, fluid treatment costs, etc. Despite the considerable research efforts about this subject, the engineering practice is still based on traditional approaches, e.g., fouling factors. Several recent papers have been focused on the schedule optimization of heat exchanger cleaning during a certain time horizon. In this case, optimization algorithms demand the availability of a heat exchanger network model in order to predict the network behavior. Aiming to this important problem, this paper presents the modeling and simulation of heat exchanger networks subjected to fouling. A fundamental aspect of the proposed simulation scheme involves the capacity to represent the network behavior considering heat transfer and fluid flow aspects simultaneously. For a given set of fouling models, it is possible to predict the temperature and flow rate along a heat exchanger network during a certain time span in the future. The model is represented by a set of matrix equations, where the network structure is parameterized using graph theory concepts. The potentiality of the proposed simulation scheme is illustrated through a typical example of an industrial heat exchanger network.

Keywords: fouling, heat exchangers, simulation.

1. Introduction

The modern concern of the society with the rational use of fossil fuel energy is not only with respect to environmental laws, but mainly because the costs are increasing every year. In process plants, an important issue about energy efficiency involves fouling management. Fouling is the accumulation of deposits on the surface of heat exchange equipment; as a consequence, this phenomenon imposes higher energy consumption to reach the required process temperatures. The traditional design of heat

exchangers take into account a factor for fouling, but due to the complexity of predicting its behaviour, often occurs unnecessary oversizing of equipment. Other important economic impacts due to fouling involve heat exchanger cleaning costs, shutdown costs, necessity of treatment of certain streams, etc.

Besides the fouling aspects related to heat transfer, fouling also brings problems to fluid flow. In this context, this paper discuss a heat exchanger network algorithm involving heat transfer equations coupled to fluid flow equations in order to predict the network behavior. The model is represented by a set of matrix equations, where the network structure is parameterized using graph theory concepts, allowing an expansion of the purely thermal approach of Oliveira Filho et al., 2007.

2. Model of network elements

The main network elements considered in this analysis are heat exchanger, pipe sections and flow machines (e.g. centrifugal pumps). The following subsections discuss the model employed for the description of the behaviour of each element. Since, the network model will be employed for optimization and time series analysis, unnecessary complex models were discarded for the representation of each element.

2.1. Heat Exchangers

The heat transfer in heat exchangers can be modeled by linear equations in relation to inlet and outlet temperatures. This linear structure can be built employing the ε-NTU method.

In this context, the equations can assume the following form (Oliveira Filho et al., 2007):

$$\begin{bmatrix} \varepsilon + (y-1) & -(y-1) & y-\varepsilon & -y \\ C_R(y-1)+y & -C_R(y-1)-y & C_R y+(y-1) & -C_R y-(y-1) \end{bmatrix} \underline{T} = \underline{0} \qquad (1)$$

where the parameter y is equal to 1, if the minimum fluid is the hot stream (i.e. the heat capacity flow rate of the cold stream is larger than the hot stream one), otherwise it is equal to 0, C_R is the ratio between maximum and minimum heat capacity flow rates and $\underline{T}$ is a vector of heat exchanger temperatures, $\underline{T}^T = [\ T_{c,i}\ \ T_{c,o}\ \ T_{h,i}\ \ T_{h,o}\ ]$ (c and h represent cold and hot streams, i and o represent equipment inlet and outlet.

The film coefficients of each stream in the heat exchanger is calculated using the equation below, with a base case and a correction factor to account for the variation of the stream flow rates:

$$h = h_{base}(m \,/\, m_{base})^n \qquad (2)$$

where h is the film coefficient, m is the mass flow rate and the exponent n is 0.8 for tube side flow and 0.6 for shell side flow, considering turbulent regime.

The hydraulic equation for heat exchangers is also built employing a base case correction:

$$P_i - P_o = \Delta P_{base}(m \,/\, m_{base})^n \qquad (3)$$

where P is the stream pressure and n is equal to 1.8 for tube side and shell side flow, in turbulent regime.

2.2 Flow machines

Flow machine model corresponds to the representation of a characteristic curve, e.g., centrifugal pumps/compressors:

$$P_i - P_o = -\rho g \sum a_i (m/\rho)^i \tag{4}$$

where a_i are coefficients of the polynomial equation of the characteristic curve. Eventual thermal effects during fluid flow along the machines are not considered

2.3. Pipe sections

The model equation for a pipe section corresponds to the Bernoulli equation, with the head loss described by the Darcy equation:

$$P_i - P_o - [(8L)/(\rho \pi^2 D^5)]m^2 + \rho g(z_i - z_o) = 0 \tag{5}$$

where L is the pipe length, D is the inner pipe diameter and z is the elevation. Head losses due to fittings can be included using the concept of equivalent length.

3. Network representation

The network flowsheet is described through a digraph, where the network connectivity is parameterized by an incidence matrix. In the flowsheet digraph, pipe sections and flow machines correspond to the edges and heat exchangers correspond to the vertices. The flowsheet digraph is employed to model energy balances along the network together with heat exchanger equations (ε-NTU). Figure 1 illustrates the process digraph related to the example discussed later in the Results section.

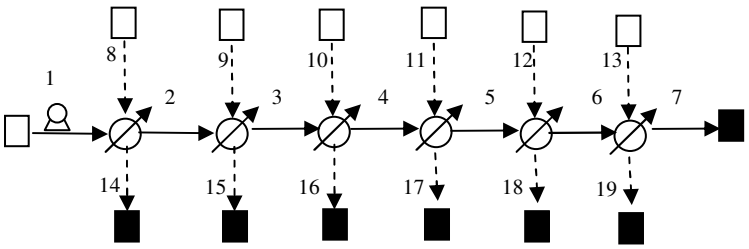


Figure 1. Process flowsheet digraph – Cold streams: continuous lines – Hot streams: dashed lines – Supply units: white squares – Demand units: black squares

However, the hydraulic model of the network, demands a different representation, where the heat exchangers become edges instead of vertices. Using an adequate reorganization algorithm of the incidence matrix, the hydraulic digraphs can be generated automatically. In this process each heat exchanger vertex is transformed in two new edges connected to four new terminal vertices. Figure 2 presents the corresponding digraphs originated from the flowsheet digraph of Figure 1.

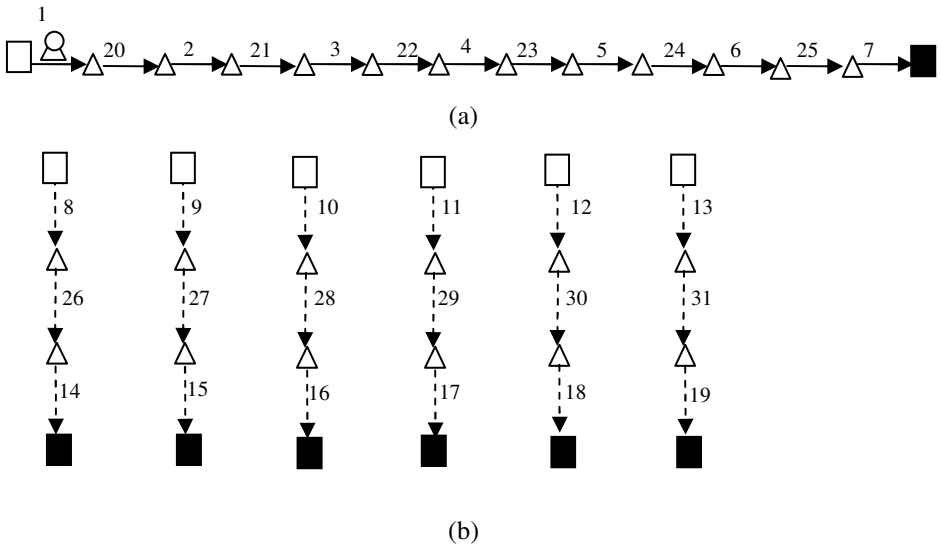


Figure 2. Hydraulic digraphs – Cold streams (a) and Hot streams (b) – Supply units: white squares – Demand units: black squares – Triangles: modified vertices

4. Network modeling and simulation

The models of the network elements are interconnected using a matrix structure based on the corresponding digraphs. The set of equations are organized in two systems: a hydraulic model, composed by a system of nonlinear algebraic equations and a thermal model, composed by a linear system.

The network simulation is conducted in two steps, the flow rates and pressures along the network are determined by the resolution of the hydraulic model through a Newton-Raphson algorithm and, then, the network temperatures are evaluated by the solution of the linear system of the thermal model.

4.1. Fouling

Since fouling is much slower than the temporal behavior of the process variables, the dynamic simulation of a fouled network can be executed using a pseudo-stationary approach. Thus, a time series of the network behavior due to fouling can be generated through a sequence of steady-state simulations, where fouling factors are updated after each individual run. The equation below presents a typical linear model for fouling growth, after a temporal discretization:

$$R_f \Big|_{t+1} = R_f \Big|_t + c\Delta t \tag{6}$$

For a given thermal conductivity of the deposits (k), the fouling thickness (δ) can be given by:

$$\delta = 0{,}5[D_{t,i} - D_{t,i}\exp(-2R_f k / D_{t,i}) \tag{7}$$

where $D_{t,i}$ is the tube inner diameter. Since, the fouling thickness is known; a correction of the tube diameter of the heat exchanger allows the simulation of the respective increase of the pressure drop (Ishiyama et al., 2008).

5. Results

The network example investigated is illustrated by the flowsheet in Figure 1. This structure is typical from a preheat train in petroleum refineries.

The thermal surface of the heat exchangers are presented in Table 1, organized in a countercurrent configuration. The inlet temperature of the cold stream is 30 ºC and the hot stream ones are 90 ºC, 140 ºC, 175 ºC, 220 ºC, 255 ºC and 276 ºC. The cold streams flow inside the tubes with a diameter of 3/4 in. The base conditions of pressure loss and flow rates in the heat exchangers are 0.5 bar and 80 kg/s for hot streams and 0.5 bar and 210 kg/s for cold streams. All heat exchangers are interconnected by pipes with a diameter of 6 in and an equivalent length of 50 m.

The supply and demand points of the cold streams are associated to a pressure of 2 bar and 3 bar, respectively. The specifications of the hydraulic circuit of the hot streams correspond to a pressure of 6 bar at the supply node and the following set of flow rates at the demand point: 40 kg/s, 30 kg/s, 30 kg/s, 80 kg/s, 120 kg/s, 100 kg/s, according to the exchanger ordering. The coefficients of the characteristic curve of the pump are a_0 = 429 m, a_1 = -331 s/m^2 and a_2 = -257 s^2/m^5 (Eq. 6). Fouling growth involves a linear model with $c = 1\cdot10^{-10}$ m^2/K·J (Eq. 8) and deposit thermal conductivity of 0.5 W/m^{2}K.

Table 1. Heat exchanger data

Heat Exchanger	1	2	3	4	5	6
Surface Area (m^2)	681	683	229	281	406	407
Overall coefficient (W/m^{2}K)	1000	3000	2500	3200	2600	4100

For the described problem, two simulations were conducted for a period of 36 months: a traditional thermal simulation and the proposed thermohydraulic simulation.

The cold stream flow rate calculated by the thermal simulation does not change, maintaining a constant value of 479 kg/s during the whole simulation period. Differently, the thermohydraulic model is able to predict the impact of fouling on the head loss, implying the reduction of the flow rate at the end of the simulation from 479 kg/s to 465 kg/s. Additionally, Figure 3 compares the results of both approaches in relation to the final cold stream temperature. As can be seen, there is a considerable difference between the profiles (a smaller cold stream flow rate reaches a higher temperature).

It is important to stress that the traditional thermal simulation does not predict any of these mentioned effects.

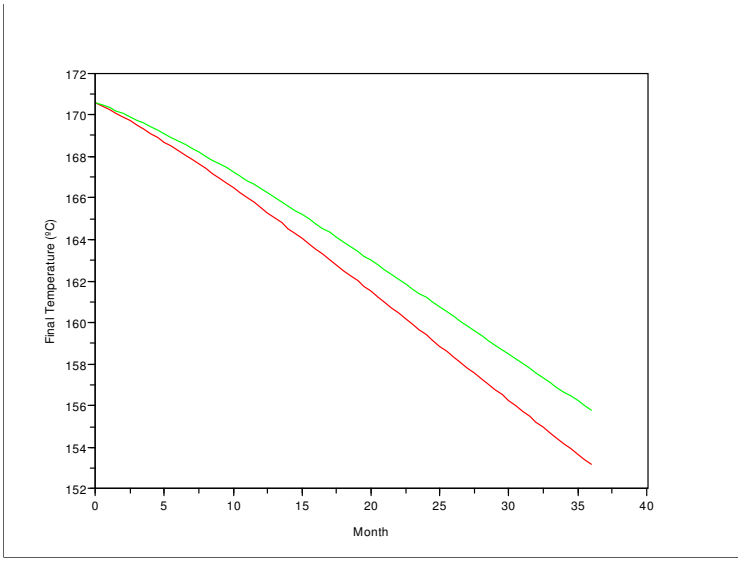


Figure 3. Final temperature of the cold stream
(green – thermohydraulic simulation, red – thermal simulation)

6. Conclusions

This paper presents a thermohydraulic model for heat exchanger network simulation. The proposed algorithm is explored seeking to predict the behavior of fouled networks. A numerical example shows a significant difference of simulation results between a traditional thermal model and the thermohydraulic proposal. This gap illustrates the importance of this approach for an accurate investigation of fouling in networks, especially when there is pump power limitations.

Acknowledgments

All authors thank Petróleo Brasileiro SA – PETROBRAS for the financial support. E. M. Queiroz and F. L. P. Pessoa are grateful to CNPq for the research scholarship and A. L. H. Costa thanks Prociência/UERJ.

References

Ishiyama, E. M., Paterson, W. R., Wilson, D. I.; 2008, Thermo-hydraulic chanelling in parallel heat exchangers subject to fouling, Chemical Engineering Science, 63, 3400-3410.

Oliveira Filho, L. O., Queiroz, E. M., Costa, A. L. H.; 2007, A matrix approach for steady-state simulation of heat exchanger networks, Applied Thermal Engineering, 27, 2385-2393.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal Planning of the Sustainable Supply Chain for Sugar and Bioethanol Production

Fernando D. Mele,[a] Gonzalo Guillén-Gosálbez,[b] Laureano Jiménez,[b] Alberto Bandoni[c]

[a]*Universidad Nacional de Tucumán, Av. Independencia 1800, S. M. de Tucumán T4002BLR, Argentina, fmele@herrera.unt.edu.ar*
[b]*Universitat Rovira i Virgili, Av. Països Catalans 26, Tarragona 43007, Spain*
[c]*PLAPIQUI, CONICET, Camino La Carrindanga Km.7, Bahía Blanca 8000, Argentina*

Abstract

The interest in renewable fuels such as bioethanol has significantly increased in the last years all over the world. In this work, a decision-support tool based on a mixed integer linear model for the design of sugarcane supply chains (SC) is developed. The approach considers the minimization of the total SC costs and environmental impact, which is determined according to the Life-Cycle Assessment principles. The solution is a set of Pareto optimal alternatives, i.e. SC configurations with associated strategic planning decisions. The capabilities of the proposed framework are illustrated through a case-study designed on the basis of the current situation in Argentina.

Keywords: bioethanol supply chain, Life-Cycle Assessment.

1. Introduction

Following the global trend, Argentina approved the National Act 26 093 to foment the production of bio-ethanol for fuel blending in 2010. However, given the fiscal-technical uncertainty generated by this measure, it is crucial that the industry turns flexible to satisfy the market, using proper energetic and environmental policies (Borrero et al., 2003; von Blottnitz and Curran, 2007).

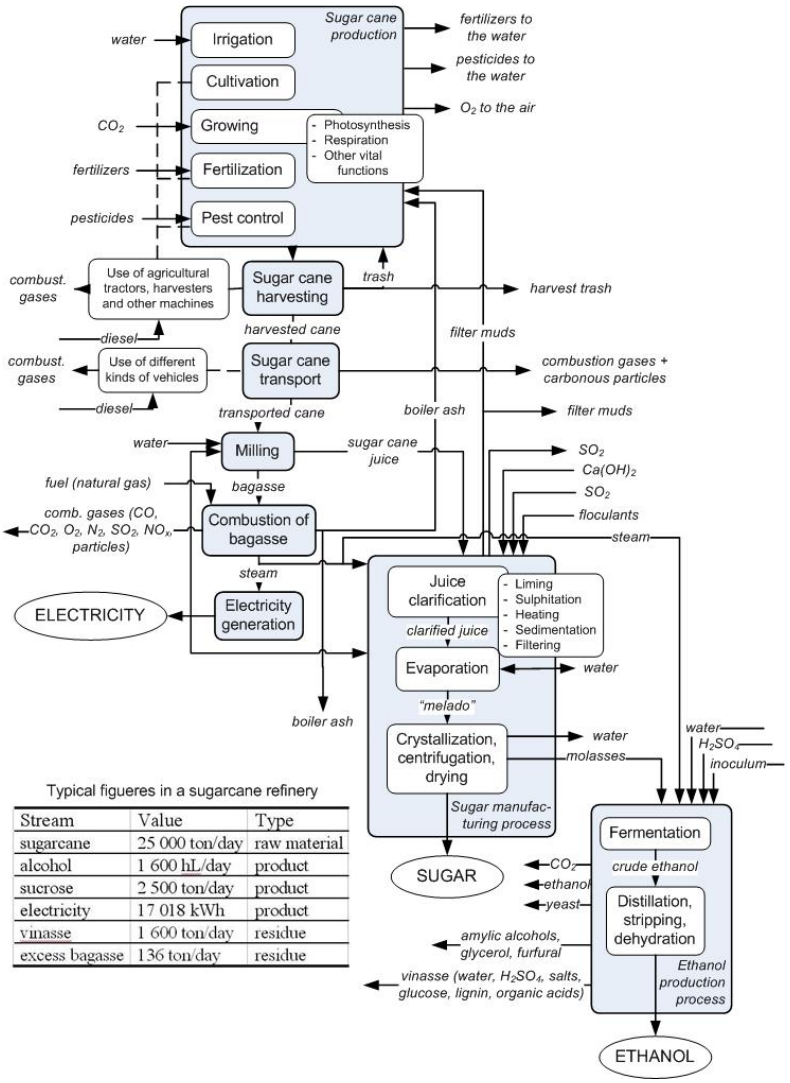
Ethanol production from sugar cane has a clearly positive energy balance which can mitigate global warming, but the fast expansion of the ethanol production/consumption has affected the global market of sugar, whose production is closely linked.

A key issue that still remains open in this area is how to address the design of efficient supply chains (SC) capable of fulfilling the sugar and bioethanol demand in the growing markets. This is not a trivial task, since it requires the understanding of the temporal and capacity interdependencies among the different entities of the network. The problem complexity is further increased by the fluctuating nature of the sugar/alcohol demand in the markets, and the need to consider different conflicting criteria at the design stage (Buddadee et al., 2008).

The aim of this work is to develop a quantitative decision-support tool based on mathematical programming for tackling this last issue, the design of SCs for sugar and bioethanol production. The solution obtained for such a problem will be a set of Pareto optimal alternatives, each of which entailing a particular SC configuration and associated strategic planning decisions (i.e., number, location, type and capacity of the production and storage facilities; sugar and alcohol production at each location; average inventory levels; and type and capacity of the transportation units).

2. Motivating Case Study

The optimization model has been inspired in the sugarcane industry of the NW of Argentina. The area occupied by the sugarcane SC has been divided into sectors, which must be provided as input data to the model. This SC includes a number of plants producing or storing sugar and bio-ethanol; and transportation links between sectors. It is assumed that sugar and alcohol are produced from sugarcane cultivated in the vicinity of the factories by using the current regional technologies. Sugar is elaborated in two different types: white and raw sugar. Ethanol is also made in two qualities: anhydrous (for fuel blending) and azeotropic.



Typical figueres in a sugarcane refinery

Stream	Value	Type
sugarcane	25 000 ton/day	raw material
alcohol	1 600 hL/day	product
sucrose	2 500 ton/day	product
electricity	17 018 kWh	product
vinasse	1 600 ton/day	residue
excess bagasse	136 ton/day	residue

Figure 1. System boundaries for the LCA-based environmental impact assessment

The SC of interest includes (I) a number of plants that produce sugar and bio-ethanol, (II) a set of warehouses where these products are stored before being delivered to the

markets and (III) a set of transportation links that connect the SC entities. The model is used to find different configurations that satisfy the customers' requirements along a time horizon divided into periods. Decisions to be made include the number, type and capacity of the storage and production facilities, the sugar and ethanol production rates and inventory levels at each facility, and the transportation flows. Since this industry combines continuous and batch processes, the periods have been taken large enough so that the operations can be considered to be continuous. Most of the parameters used have been taken from industry, whereas missing data has been supplied from the literature (Hugot, 1982).

The standard production scheme of sugar and ethanol considered is showed in Figure 1. For sugar production, it includes milling, clarification, evaporation (vacuumed quadruple effect tanks), crystallization (three cooked masses), centrifugation and drying. For ethanol production, the process includes: substratum preparation, fermentation, centrifugation, distillation, rectification and dehydration. To obtain ethanol, different combinations (blending) of the sugar production process streams are allowed (sugarcane syrup, molasses A, B or C).

3. Mathematical model

The approach relies on assuming a superstructure that embeds all possible logistic alternatives to the SC design problem. This mathematical representation is then converted into a bi-objective MILP problem that is solved via standard branch and bound techniques. To construct such a superstructure, the overall region of interest (i.e., NW of Argentina) is first divided into a set of sectors, each featured by a specific demand of sucrose and ethanol. The production plants and storage facilities can be established in any of these sectors, and may be expanded in capacity over time in order to follow a specific demand pattern. The model includes three main sets of equations: mass balances, capacity constraints and objective function calculations. A brief outline of each of these blocks of equations is next given.

3.1. Mass Balance Constraints

A total mass balance must be written to relate inputs and outputs of a particular region, in a given period *t*. The current inventory level in a region *f* (S^T_{isft}) is equal to the sum of the total flow rate of each product entering region ($Q_{ilf'ft}$), plus the amount of raw materials purchased PU_{ift}, plus the total production rate of the same region (P^T_{ift}), and the inventory level in the previous time interval ($S^T_{isf,t-1}$), minus the total flow rate leaving this region ($Q_{ilff't}$), the flow of wastes produced W_{irt}, and the total demand required by region r itself (D^T_{irt}). The following equation expresses this balance:

$$\sum_s S^T_{ift} = \sum_s S^T_{if,t-1} + P^T_{ift} + RM_{ift} + \sum_{f' \neq f} (Q_{if'ft} - Q_{iff't}) + W_{ift} - D^T_{ift}, \forall i, f, t \quad (1)$$

3.2. Capacity Constraints

The actual capacities and capacity expansions of the production facilities are represented by a set of continuous variables (C^{PL}_{fit} and CE^{PL}_{fit}), whereas the number of capacity expansions executed in a sector are given by integer variables (N^{PL}_{fit}). Similar expressions are used for storage facilities.

$$C^{PL}_{fit} = C^{PL}_{fit-1} + CE^{PL}_{fit}, \quad \forall f, i, t \quad (2)$$

$$N_{fit}^{PL}\underline{C_{fi}^{PL}} \le CE_{fit}^{PL} \le N_{fit}^{PL}\overline{C_{fi}^{PL}}, \quad \forall f,i,t \tag{3}$$

Regarding the transportation constraints, there must be a product flow between sectors so that demands are fulfilled. The product flow between sectors *f* and *f'* only will exist if the transportation channel is set.

$$Q_i^{\min} X_{iff't} \le Q_{iff't} \le Q_i^{\max} X_{iff't} \qquad \forall i, f, f'(f \ne f'), t \tag{4}$$

being $X_{iff't}$ a binary variable for the existence of a channel.
The storage facilities absorb demand and supply fluctuations. The storage period *b* is introduced to define the average capacity S'_{ift}:

$$S'_{ift} = bD_{ift}^{T}, \ \forall i, f, t \tag{5}$$

3.3. Objective Functions

One of the objectives of the model is to minimize the total costs (*TC*) of the sugarcane SC. The capital costs are computed when manufacturing plants, storage facilities (FCC_t) and transportation channels (TCC_t) are established. The operational costs are computed daily and they correspond to the production and storage costs (FOC_t) and the transportation costs (TOC_t) throughout the SC:

$$TC = \sum_t \frac{FCC_t + TCC_t + FOC_t + TOC_t}{(1+IR)^{t-1}} \tag{6}$$

being *IR* the interest rate.
In order to define an environmental indicator, only two steps of a LCA study have been used (Consoli et al., 1993). In the first step, the process boundaries have been expanded to include the entire SC (see Figure 1), and the functional unit has been chosen as the tons of sugarcane consumed along the time horizon. The second step is an inventory of the emissions released and resources consumed as a function of the tasks carried out in the network. Among these, the biochemical oxygen demand (BOD_{20}) results to be the most relevant because of its importance as an indicator of the pollution of the watercourses. BOD_{20} is a variable depending on the continuous and discrete decision variables of the model (SC design and operation).

3.4. Multi-objective Optimization

Thus, two objective functions, f_1 and f_2, are included in the bi-objective formulation:

$$\min_{x,y} U \begin{cases} f_1(x,y) = TC \\ f_2(x,y) = BOD_{20} \end{cases} \tag{7}$$

U is the global objective function and *x* and *y* are vectors of the continuous and discrete model variables, respectively, defined into the feasible region of the constraints, such as material flows, facility capacities and binary variables related to the existence of a plant or distribution channel. The problem solution is a set of efficient or Pareto optimal points obtained by using the ε-constraint method, a well known method, whose solutions have to be checked for feasibility and efficiency.

4. Results

On solving the bi-objective MILP, the Pareto points showed in Figure 2 have been obtained, confirming the existent trade-off between the economic and environmental issues. The points with lower costs are those which have the highest environmental impact. It is hence responsibility of the decision-maker to choose which of the configurations (represented by a particular Pareto point) is what better fits her/his preferences. As can be observed, the slope of the Pareto curve is rather smooth on its left side becoming sharper as one moves to the right. These results indicate that it is possible to achieve significant environmental improvements at a marginal increase in cost in the Pareto points that are close to the minimum cost solution. The opposite situation occurs in the points on the other side. Figure 3 shows the configurations corresponding to the two extreme situations. As can be seen, the minimum cost solution entails a more decentralized network configuration. This is because the reduction in the total SC capital cost that is attained by opening a lower number of nodes compensates the increase in the transportation cost. On the other hand, in the minimum environmental impact solution, the transportation tasks, the most polluting ones, are minimized.

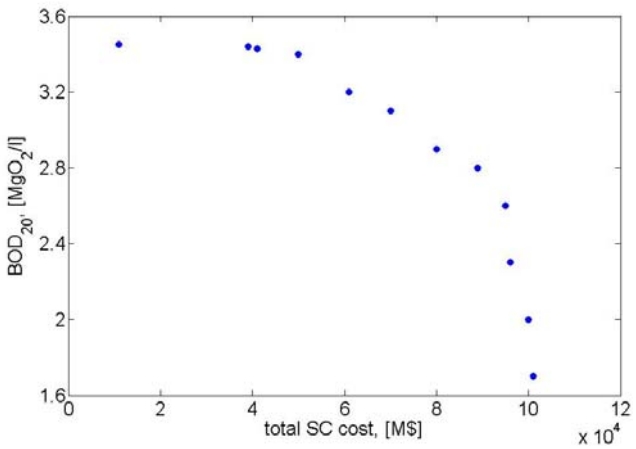


Figure 2. Pareto points between two objectives

The MILP models used have been solved on a Pentium M 730 machine, 1.6 GHz using CPLEX 9.0 through GAMS (Brooke et al., 1998). The optimization of one SC configuration involves about $3 \cdot 10^5$ equations, $5 \cdot 10^4$ continuous variables and $1.2 \cdot 10^3$ discrete variables, and it consumes 200 seconds on average.

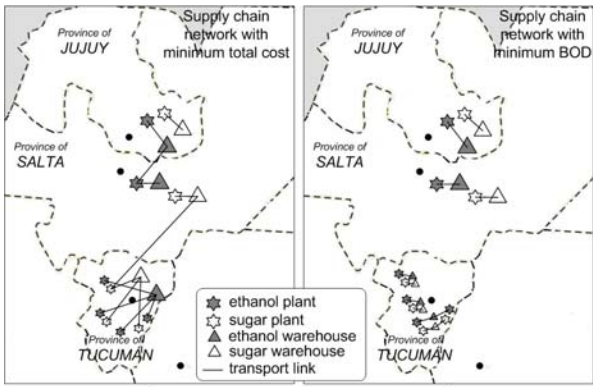


Figure 3. Two extreme cases

5. Conclusions

This work has addressed the design of SCs for sugar and ethanol production with economic and environmental concerns. The bi-objective MILP model of the sugarcane SC enables comparing the traditional economic indicators with the emerging environmental impact indicators. It has been shown how the solutions set can become valuable information from the environmental point of view for making SC strategic decisions.

Furthermore, the mathematical model has provided valuable insight into the design problem, suggesting the convenience of adopting more decentralized SC configurations in which the transportation tasks are reduced.

It is our current work to incorporate more details, decomposition strategies if necessary, as well as the treatment of the uncertain parameters and variables.

6. Acknowledgements

The authors wish to acknowledge support from CONICET (Argentina), the Spanish Ministry of Education and Science (project DPI2008-04099), and the Spanish Ministry of External Affairs (projects A/8502/07- HS2007-0006).

References

M. Borrero, J. T. V. Pereira, E. E. Miranda, 2003, An environmental management method for sugar cane alcohol production in Brazil, Biomass & Bioenergy, 25, 287-299.

A. Brooke, D. Kendrik, A. Meeraus, R. Raman, R. E. Rosenthal, 1998, GAMS-A User's Guide; GAMS Development Corporation, Washington DC.

B. Buddadee, W. Wirojanagud, D. J. Watts, R. Pitakaso, 2008, The development of multi-objective optimization model for excess bagasse utilization: A case study for Thailand. Environmental Impact Assessment Review, 28, 380-391.

F. Consoli, D. Allen, I. Boustead, N. de Oude, J. Fava, et al, 1993, Guidelines for Life-Cycle Assessment: A 'Code of Practice', SETAC.

E. Hugot, 1982, Manual para ingenieros azucareros. C. E. Continental: Mexico.

H. von Blottnitz, M. A. Curran, 2007, A review of assessments conducted on bio-ethanol as a transportation fuel from a net energy, greenhouse gas, and environmental life cycle perspective, J. Cleaner Prod., 15, 607-619.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Dynamic Simulation of a Compressor Located in a Natural Gas Processing Unit Using EMSO Simulator

Felipe C. Cunha[a], Maurício Bezerra de Souza Jr.[a], Amaro G. Barreto Jr.[a] and Domingos F. de S. Souza[b]

[a] *Department of Chemical Engineering, Federal University of Rio de Janeiro, Av. Athos da Silveira Ramos, 149, CT, Sala E-207, 21941-909, Rio de Janeiro – RJ, Brazil.*
[b] *Department of Chemical Engineering, Federal University of Rio Grande do Norte, Av. Sen. Salgado Filho, 3000, Lagoa Nova, 59072-970, Natal – RN, Brazil.*

Abstract

Nowadays, a profound knowledge of all the stages involved in large scale processing of oil, natural gas or other fuels is necessary, due to the increased preoccupation with how to use the natural resources in a responsible way. The facilities for processing natural gas can be divided in two main parts, the first one responsible for the condensation of larger hydrocarbons (ethane, propane, butane, and the like) and the second one for the distillation of these larger hydrocarbons obtained. The first part of these facilities achieves a minimum temperature of -90°C due to a propane refrigeration system. This work studies a centrifugal compressor situated in the propane refrigeration system of a Natural Gas Liquid Recovery Unit (LRU) located in Macaé (Rio de Janeiro, Brazil) – Petrobras Transporte S.A. (Transpetro). This compressor is one of the most complex and important equipments of the LRU. The main contribution of this paper is to propose a dynamic model for this centrifugal compressor based on mass and energy balance coupled with thermodynamics equations for the nonidealities of the gas. This dynamic model was used to simulate real conditions in some situations observed in the LRU, analyzing its open loop behavior and investigating its automatic control (controller tuning and closed loop behavior). The situations studied consisted in applying step disturbances to the centrifugal compressor and then interpreting the results in the light of the developed model. The modeling, simulation and control investigations were achieved using a Brazilian simulator called EMSO (Environment for Modeling, Simulation and Optimization), which is a graphical environment that can be freely modified by all users. In this environment, the language is fully object-oriented, allowing the user to develop complex dynamic or steady models by composing them with existent small models or build specific models by deriving standard ones. EMSO was developed by the ALSOC Project (Free Environment for Simulation, Optimization and Control of Processes) which supplies this software without costs for partners and universities, aiming therefore to standardize and spread out its use. The present work provides information on the operation and control of the compressor allowing the further development of abnormal situation management tools that will support the control system and provide an easier and safer operation of the LRU.

Keywords: centrifugal compressor, dynamic simulation, capacity control, natural gas

1. Introduction

This work focuses on two Natural Gas Liquid Recovery Units (LRU), which are located in Macaé (Rio de Janeiro, Brazil) – Petrobras Transporte S.A. (Transpetro). On these LRU's, the treated natural gas passes trough heat exchangers (with propane as the refrigerating fluid), where it is cooled and, then, expanded in a turboexpander. Because of this expansion, the larger natural gas components are liquefied. This is the way that ethane is obtained for the petrochemical industry (Mendonça, 2007). Mendonça (2007) concluded that, if the main LRU's goal is to recover a maximum ethane quantity, the best strategy is to use the centrifugal compressors as much as possible, trying to reach the lowest temperature in the vessel which separate gases for the turboexpander. As the centrifugal compressor is a key equipment of these LRU's, its dynamic modeling and its control investigation are extremely useful for the proper operation of these facilities. Therefore, this work has as its main goal the proposal of a dynamic model for the propane compressors situated in the natural gas refrigeration system and, consequently, the analysis of its behavior. All the real conditions used in this work were provided by Petrobras Transporte S.A. (Transpetro).

The simulator used in this paper was the EMSO - *Environment for Modeling, Simulation and Optimization* (SOARES & SECCHI, 2003). This simulator has many advantages, like the availability of a range of models, that can be freely modified by the user, and the simplicity of the language which permits the user to write almost as can be written in a paper, besides being fully object-oriented.

1.1. Natural Gas Processing

In sum, the facilities for processing natural gas can be divided in two main parts, the first one responsible for the condensation of larger hydrocarbons (ethane, propane, butane, and the like) and the second one for the distillation of these larger hydrocarbons obtained. The first part of the facility employs an auxiliary system, where the so called propane refrigeration system is found, as can be seen in Figure 1.

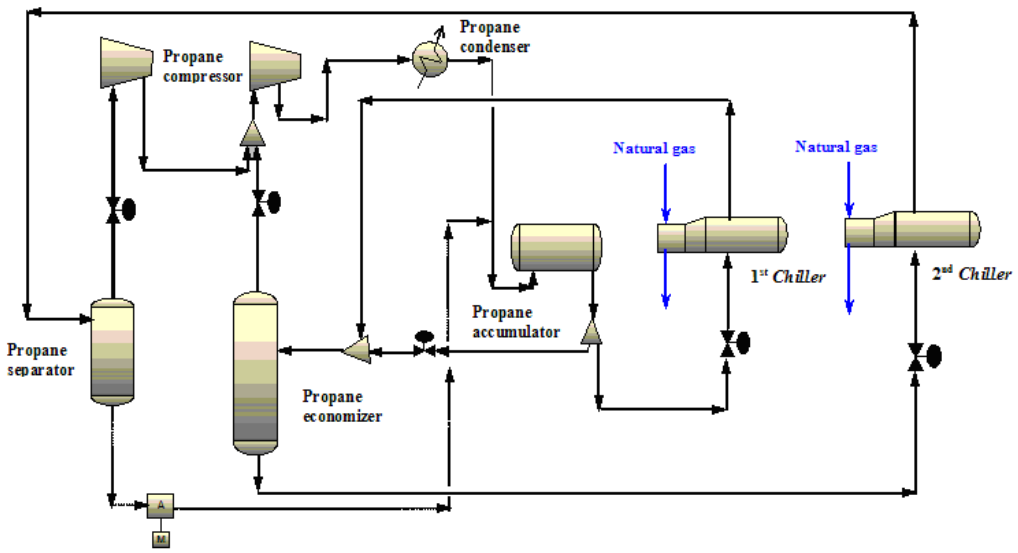


Figure 1. Simplified flowsheet of the propane refrigeration system (Source: Mendonça, 2007).

1.2. Thermodynamic Theory of Centrifugal Compressors

While the centrifugal compression takes place, the gas passes through the impeller and, hence, its energy increases (this energy is composed by enthalpy and, mainly, kinetics energy). After that, the gas enters in the diffuser, where the obtained kinetic energy is converted in enthalpy. This energy transfer that happens in the impeller can be represented by Equation 1 (Rodrigues, 1991)

$$dW = \frac{du^2}{2} + dh. \qquad (1)$$

where $u^2/2$ is the kinetic energy per mass unit, W is the work given to the gas per mass unit and h is the gas enthalpy per mass unit.

The work given to the mass unit of gas can be defined as (when compressors are the subject) head (H), therefore the enthalpy energy is also called static head (H_s) and the kinetic energy is also called dynamic head (H_d). When the gas exits the impeller, the dynamic head is converted into static head and, consequently, the pressure is increased; from this moment on the static head is called effective Head (which is nothing more than the total variation of enthalpy, after this gas passes through the impeller-diffuser). The effective head (or just enthalpy) definition allows the use of Equation 2:

$$H_{ef} = \int_{p_S}^{p_D} v.dp\,, \qquad (2)$$

where the interval between p_S e p_D are the pressure in the suction and discharge area, respectively and v represents the volume per mass unit. By the Equation 2, it is noticeable that a satisfactory relation between the pressure and the volume is necessary. Once defined this relation, Equation 2 can be integrated. The effective Head now can be called also as polytropic Head, since this one was computed in the polytropic base (the polytropic model $p.v^n=constant$, where n is the polytropic coefficient). Now the problem is to determine this polytropic coefficient. It is possible to demonstrate, that the polytropic and adiabatic coefficient have the following relation (Smith, Van Ness & Abbott, 2000):

$$\frac{n-1}{n} = \frac{k-1}{k.\eta_p} \qquad (3)$$

where η_p represents the polytropic efficiency ($\eta_p = H_p / H$). If α is defined as the followed ratio $\frac{n-1}{n}$ and R_c is defined as $\frac{P_D}{P_S}$, the Equation below is achieved:

$$H_p = z_m.\frac{R.T_S}{\alpha}.\left(R_c^{\alpha} - 1\right), \qquad (4)$$

where z_m is the compressibility factor, which intends to express the nonideality for the gas while it is compressed. This one can be computed as an arithmetic average between the calculated compressibility factor in the suction area and in the discharge area. The Virial Equations were used to calculate all the compressibility factors, because these are well fundamented in theories (Smith, Van Ness & Abbott, 2000). And the coefficient k

is obtained by its definition that means the ratio between the calorific capacities at constant pressure and constant volume.

1.3. Vessel Dynamics

Hypothetical accumulation vessels (a suction vessel and a discharge vessel) are used here to model the dynamics of the compressors. The dynamics of the vessels is described by:

$$\frac{d(P/z)}{dt} = \frac{R.T}{V.PM_{mist}}.(Q_{in} - Q_{out}), \qquad (5)$$

where PM_{mist} is the molecular weight of the mixture (kg/kmol) and Q_{in} and Q_{out} are the inlet and outlet mass flows to the vessel.

2. Results and Discussion

2.1. Analysis of the centrifugal compressor open loop

2.1.1. Step change in the inlet flow of the Suction Vessel in the first section

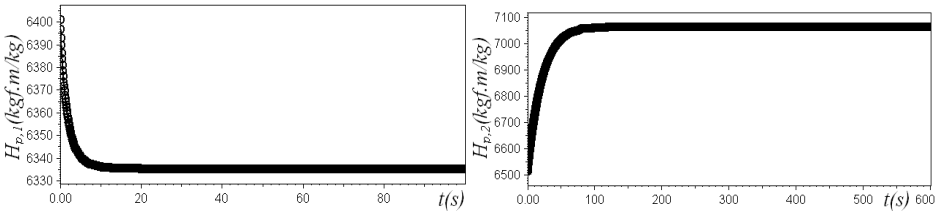
The centrifugal compressor has two sections. In order to analyze its behavior, a 1% positive step change was applied in the volumetric inlet flow (F_s) of the suction vessel in the first section, using the EMSO simulator. This step change affected the polytropic head, which was given for the gas in the first section, in a negative way as can be observed in the Figure 2(a). At the same time that the volumetric flow was increased, the ideal procedure would be to increase also the rotation of the impeller, trying to adjust the new flow as the operational flow of the compressor. Unfortunately, this procedure could not be realized, once the rotation of the impeller can not be changed. So the gas is going to enter in the impeller far from its ideal condition, which would be tangent to the blade of the impeller, therefore whirls will exist, dissipating, in this way, the energy and, consequently, decreasing the polytropic head that would be transferred from the blade of the impeller to the gas. The pressure in the discharge vessel located in the first stage also presented a satisfactory behavior and similar to the behavior of the pressure in the suction vessel.

The analysis of the polytropic head in the second section also showed that the step change did not approximate its flow to the operational flow and the reasons are the same explained above.

2.1.2. Step change in the outlet flow of the Discharge Vessel in the second section

This time a 10% negative step change was given in the outlet volumetric flow of the discharge vessel in the second section, using the EMSO simulator. In the first section practically no change was observed (Cunha, 2008), while in the second section the discharge pressure increased. However, the most interesting result was the polytropic head variation, as can be seen in the Figure 2(b).

Considering that the compressor was in steady-state before any disturbance, the compressor will not be anymore in steady-state after the negative step change, but now in a dynamic state, where the inlet flow will be greater than the outlet flow. Trying to achieve another steady-state condition, the compressor needs to decrease its inlet flow, what allows the gas to win more polytropic head. Therefore, it is concluded that the new flow adjusted is near to the operational flow.



(a) (b)

Figure 2. Dynamic behavior of the polytropic head. (a) Head in the first section after a positive step change in the inlet flow of the suction vessel; (b) Head in the second section after a negative step change in the outlet flow of the discharge vessel.

2.2. Analysis of the centrifugal compressor closed loop

A capacity control was implemented, where the controlled variable was the discharge pressure in the second section and the manipulated variable was the volumetric inlet flow of the suction vessel in the first section. The tuning procedure was as follows: the initial controller tunings were made using the Ziegler-Nichols method and, if this method was not efficient enough, the Tyreus-Luyben method (Edgard, Mellichamp and Seborg, 2004); if both methods still performed inefficiently, simple manual adjustments were used until the oscillatory behavior decreased significantly. The results for set-point (+ 1 %) change are discussed bellow.

The proportional control presented an expected strong oscillatory behavior behavior when the Ziegler-Nichols method was used, because this method is based in a one quarter decay ratio response. This oscillatory behavior shows that the gain controller is high, forcing the controller to act intensively in the system. Therefore, the gain was decreased, trying to obtain a better control over the discharge pressure. So, instead of multiplying the critic gain by 0.5, smaller values were used until the oscillation decrease significantly, as can be seen in Figure 3. It was also observed that the proportional control had no offset, due to the capacitive nature of the process.

The proportional-integral control was also tuned by Ziegler-Nichols, but instability was obtained with this set of parameters. Hence, the Tyreus-Luyben method was used. This one presented a better control, however oscillations were present. This problem was solved just with a manual adjustment, which made the behavior of the controller satisfactory. The results can be observed in Figure 3. Comparing the set of parameters obtained using the Ziegler-Nichols and the Tyreus-Luyben methods, it was noticed that the gain was decreased and the integral time was increased for the latter approach. Based on this rationale, the manual adjustment consisted of additional decreases in the gain and increases in the integral time.

A stable closed loop with the proportional-integral-derivative control tuned by the Ziegler-Nichols method was not obtained; therefore the Tyreus-Luyben method was used. This second method was sufficient and no manual adjustment was required. The result of this tuning is presented in the Figure 3. Analyzing again the set of parameters (Cunha, 2008), it was observed that the controller based on the Ziegler-Nichols method was acting strongly, what caused too much oscillation in the system. In the other hand, the Tyreus-Luyben method caused less oscillation, because the integral action was decreased significantly and the derivative action was increased, and this one has a stabilizing effect.

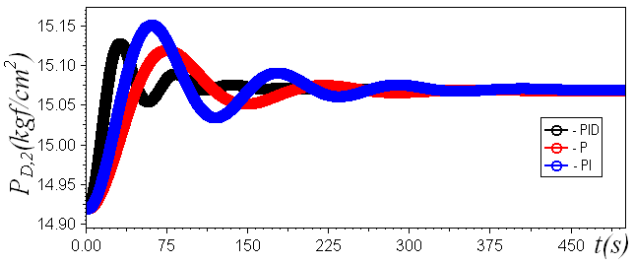


Figure 3. Results obtained for the three controllers.

It was noticed that the three controllers worked well, achieving the required set-point. Based on the set of parameters, the stabilizing effect of the derivative term allowed the increase in the proportional term. However it is important to keep in mind that the derivative action worked well, because the studied system is not subject to high frequency disturbances.

3. Conclusion

A dynamic mathematical model was developed for the propane compressor of the propane refrigeration system of the LRU. The EMSO simulations provided results that could be well analyzed, understood and explained in the light of the developed dynamic model. The feedback controllers employed for the capacity control had good servo performances, specially the PID control, which put the discharge pressure on its set-point in the shortest time. Regulatory tests were also performed and the controllers exhibited a similar performance (Cunha, 2008). More complete conclusions require the use of an anti-surge control coupled with the studied capacity control.

4. References

E. S. MENDONÇA. Monitoramento, Diagnósticos e Otimização Operacional de uma Unidade de Processamento de Gás Natural. Rio de Janeiro: Escola de Química – UFRJ, 2007. Dissertação de Mestrado em Engenharia Química.

P. S. B. RODRIGUES. Compressores Industriais. Rio de Janeiro: EDC – Editora Didática e Científica Ltda, 1991.

W. G. dos SANTOS. Processamento de Gás Natural. Vitória: Instituto Brasileiro de Petróleo e Gás – IBP, 2006.

D. E. SEBORG; T. F. EDGAR; D. A. MELLICHAMP. Process Dynamics and Control. 2ª Edição. Editora John Wiley & Sons, 2004.

J. M. SMITH; H. C. VAN NESS; M. M. ABBOTT. Introdução à Termodinâmica da Engenharia Química. Tradução: Eduardo Mach Queiroz e Fernando Luiz Pellegrini. Pessoa. 5ª Edição. Rio de Janeiro: LTC – Livros Técnicos e Científicos Editora S.A., 2000.

SOARES, R.P. and A.R. SECCHI. EMSO: A New Environment for Modeling, Simulation and Optimization. ESCAPE 13, Lappeenranta, Finlândia, 947–952 (2003).

F. C. CUNHA. Simulação Dinâmica de um Compressor de uma Unidade de Processamento de Gás Natural usando o Ambiente Gráfico EMSO. Rio de Janeiro: Escola de Química – UFRJ, 2008. Dissertação de Projeto Final em Engenharia Química.

Acknowledgments

This work was sponsored by 'Brazilian Research and Projects Financing Agency' (FINEP) and PETROBRAS (Grant 01.04.0902.00).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Control Oriented Dynamic Rigorous Model of a Fuel Processor System and Fuel Cell Stack

L. Nieto Degliuomini[a], S. Biset[b], J. M. Domínguez[b], M. S. Basualdo[a,b]

[a]*GIAIP-CIFASIS (CONICET-UNR-UPCAM)-27 de Febrero 210 bis S2000EZP – Rosario (Argentina).nieto, basualdo@cifasis-conicet.gov.ar*
[b]*UTN-FRRo Zeballos 1341 S2000EZP – Rosario (Argentina)*

Abstract

Dynamic rigorous model of a complete plant of a Fuel Processor System (FPS) that reforms bio-ethanol to hydrogen-rich mixture to feed the anode of a fuel cell stack, along with a Proton Exchange Membrane Fuel Cell (PEM-FC) is presented. This is the first work that deals with this kind of integrated plant and is used to analyze the dynamic behavior and then propose a suitable control structure. The model can be implemented thanks to the use of a specific communication protocol between HYSYS® and MATLAB® to coordinate the calculations. The heat integration is achieved with the 'LNG' tool, available at HYSYS®, which applies the pinch analysis and provides an approximation of the optimal heat exchanges that can minimize the auxiliary services in steady state mode. It allows to assume that the stream temperatures are always at their optimum values, so their complex dynamic effects are neglected in this first problem approximation. On the other hand, the dynamic behavior of the reactors are modeled in MATLAB® including the PEM. The main objectives of the FPS control are to maintain H_2 levels on the anode of the FC, retain low CO levels at the inlet stream of the anode, and keep the temperatures of the reactors set and FC under rigorous specifications.

Keywords: Fuel processor, PEM, Rigorous dynamic model, Hydrogen generation.

1. Introduction

Ethanol is a very good fuel cell energy source because of its high efficiency and concentration of H_2 that can be obtained, also because of its environmental friendliness. Usual methods of converting bioethanol are catalytic partial oxidation and steam reforming, the latter is chosen here because of the high concentrations of hydrogen that can be obtained to feed the fuel cell stack.

Dynamic rigorous models of fuel cell power plants are extremely important in understanding the variable interactions, implications on system performance and transient behaviors, hardly analyzed nowadays. Designs for the Catalytic Partial Oxidation (CPO) and Water Gas Shift reactors in fuel cell power systems are presented in Görgün et al. (2004) and Görgün (2005) using natural gas (methane) as fuel. However, these designs are based on the reactors invariants and do not take into account reaction rate expressions. A control-oriented model of a CPO of natural gas processor is developed in Pukrushpan et al. (2004). A mathematical model of the reformer is developed by Chuang et al. (2008) and selected parameters of the model are fit to match experimental data from the dynamic response of the process.

2. Fuel Processor System

The FPS (illustrated in Fig. 1) consists of an Ethanol Steam Reforming (ESR) plug flow reactor, where most of the conversion of ethanol to H_2 is made. Carbon monoxide (CO) which poisons the fuel cell catalyst is produced in the ESR, so additional processing is needed to remove this substance. There are three reactors that configure the cleaning system; these are two Water Gas Shift, one of high temperature (fast), and the other of low temperature, that favors the equilibrium of the reaction to higher conversion rates of CO. The third is a Preferential Oxidation of Carbon monoxide (CO-PrOx) reactor, where oxidation of CO into CO_2 is made; also, the undesired oxidation of H_2 occurs, so the catalyst is selected to improve the conversion of CO.

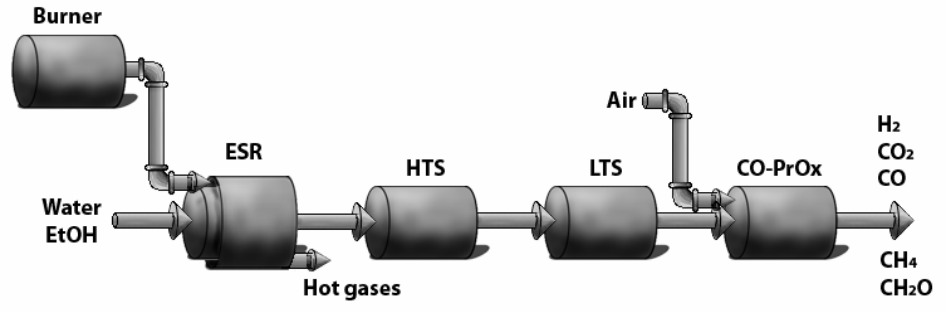


Figure 1. Fuel Processor System

The overall reaction is endothermic, and the heat requirement is supplied by a burner, which is fed with ethanol and compressed air, the transference of heat is achieved passing the hot gases through the jacket of the reformer.
There are two WGS reactors operating at different temperatures, and WGS reaction is produced inside them:

$$CO + H_20 \leftrightarrow H_2 + CO_2 \qquad \text{(Eqn. 1)}$$

This reaction produces heat and H_2. Levels of CO are still high even after the two WGS reactors, so the final elimination is made in the CO-PrOx reactor, which produces the oxidation of CO into CO_2 and the undesired combustion of H_2. The WGS reaction takes place in this reactor too. Oxygen is injected into the CO-PrOx, the amount needed is about twice the stoichiometric relationship to have a good selectivity and satisfy the requirements of the FC. The plug flow reactors are modeled with twenty lined-up Continuous Stirred Tank Reactors (CSTR). The molar flow between two volumes, and the exit molar flow of each reactor, is given by the orifice flow equation as a function of upstream pressure and downstream pressure. It is assumed that the ideal gas equation of state is valid, and for gas flow rate, a material balance on each species gives:

$$\frac{dy_{j,i}}{dt} = \frac{F_{i-1}y_{j,i-1} - F_i y_{j,i} + (1-\varepsilon)\sum_j \nu_{j,k} r_k \rho_{cat}}{vol_{slice}} \qquad \text{(Eqn. 2)}$$

Where y, F, ε, ν, r, ρ_{cat} and vol_{slice} are mole fraction, flow rate, void fraction, stoichiometric coefficient, reaction rate, catalyst density and volume of slice, respectively. Subscripts j, i and k refers to species, slice, and reaction.
The heat generated by the different reactions is supposed to vary with temperature. A thermal balance for each slice gives the equation for the temperature of the gas phase:

$$\frac{dT_i}{dt} = \frac{F_{i-1}\,\overline{cp}_{i-1}T_{i-1} - F_i\,\overline{cp}_i T_i + (1-\varepsilon)\sum_j (-\Delta H_k) r_k \rho_{cat} + Q}{vol_{slice}\,\overline{cp}_i} \quad \text{(Eqn. 3)}$$

Where T, $\overline{cp}$ and ΔH are temperature, specific heat, and heat of reaction, respectively. All units except the ESR are considered adiabatic; Q is a term that accounts for the exchanged heat so is the only calculated for the reformer.

2.1. Ethanol Steam Reforming

Despite the apparent simplicity of the stoichiometric reaction for the production of hydrogen:

$$CH_3CH_2OH + 3\ H_2O \leftrightarrow 6\ H_2 + 2\ CO_2 \quad \text{(Eqn. 4),}$$

actually, this reaction involves a complex system of reactions, and the selectivity towards H_2 is affected by a number of undesired secondary reactions.

The analysis made by Francesconi (2008) adjusting experimental data of a Cobalt catalyst supported in Zircon (CO/ZrO2), leads to the reaction scheme of acetaldehyde decomposition, according to the mechanism proposed by Benito et al. (2005).

To fulfill the heat requirements of the reactions, the ESR is provided with a hot gases chamber, it collects the products of combustion from the ethanol burner, and exchanges its heat with the main reactor through a jacket.

2.2. Water Gas Shift

In these reactors, the only reaction that takes place is the WGS. The main objectives of this reaction are to remove the CO from streams and adjust the H_2/CO molar rate. This reaction is moderately exothermic and its equilibrium constant decreases as temperature rises, so lower temperature favors high conversions. Reactors of the cleaning system are considered adiabatic, and this condition limits their performance, which is the reason that justifies the need of more than one WGS reactor with intermediate cooling.

2.2.1. High Temperature WGS

The High Temperature water gas Shift (HTS) reactor uses an iron based catalyst (Fe/Cr), and makes most of the conversion of CO.

2.2.2. Low Temperature WGS

The Low Temperature water gas Shift (LTS) reactor operates at lower temperatures and uses a $CuO/ZnO/Al_2O_3$ catalyst. It is thermodynamically favored, so higher conversions rates are obtained because the equilibrium constant is low.

2.3. Preferential oxidation of Carbon Monoxide

The main (desired) reaction given in this stage is the oxidation of CO into CO_2; the oxidation of H_2 is produced too, it is undesirable because reduces the efficiency of the FPS and increases the temperature of the stream. Although many publications present kinetic expressions for CO oxidation, only few of them consider the simultaneous H_2 oxidation. Choi and Stenger (2004) reported that WGS reaction should be considered at this stage and give equations for the three reactions, with a $Pt\text{-}Fe/Al_2O_3$ catalyst.

2.4. Burner

A stream of ethanol mixed with air is fed into this reactor to burn the ethanol and obtain hot gases that will go to the jacket of the ESR to provide the necessary heat to reach good yields. In addition, the exit gases from the PEM-FC are fed into this reactor to burn completely the remaining H_2 and methane.

3. The Mathematical Model Implementation

The pressure requirements are satisfied with compressors and turbines modeled in HYSYS®. In addition, it supports the important data bank information for the different components, and the LNG operation, useful for minimizing auxiliary services. On the other hand, the dynamic behaviors of the reactors are modeled in MATLAB®, which, supports the differential equations integration solving with the ode15s solver (variable order solver based on the Numerical Differentiation Formulas), together with the data obtained from the steady-state case in HYSYS®. The communication interface is performed by the use of the spreadsheets in HYSYS® and a specific library for doing the corresponding data transference and updating at scheduled sampling time between both programs.

3.1. Heat Integration

The heat integration needed for the model is achieved by the LNG tool in HYSYS®, it is an operation that solves material and heat balances for multi-stream exchangers and heat exchangers networks by means of the Pinch technology, as described in Linnhoff et al. (1994). With this tool the simulation of the heat exchangers network is defined without actually knowing its configuration. The minimum heat requirement of the system and the minimum heat to be evacuated can be computed for each operating point or with the system under different disturbances.

3.2. Proton Exchange Membrane Fuel Cell

A fuel cell stack is constituted by an anode, where the fuel is injected, and a cathode, where the oxidant, normally oxygen or air, is injected. The electrodes are separated by a semi-permeable membrane that allows the proton exchange and makes the oxidation reaction to produce electrical power. The cell generates an open-circuit voltage which is affected by a number of losses (activation, concentration and ohmic) that leads to a useful actual voltage. Pukrushpan et al. (2004) presents a rigorous dynamic model of a PEM-FC which is used in this work, adapted to produce a maximum power of 10 KW. Transient behavior of manifold filling, membrane hydration, air compressor and heat management are included in the model. Interactions between processes are also accounted. In that work, the necessary hydrogen production is calculated as a function of the stack current:

$$W_{H_2} = \frac{nI_{st}}{2F} \qquad \text{(Eqn. 5)}$$

Where W_{H_2} refers to H_2 production (taken as setpoint), n is the number of cells in the stack and F the Faraday constant.

4. Control Structure

The development of the control oriented dynamic model presented here allows a more efficient testing of the adopted control structure based on a previous analysis, completely performed in steady state model. It was described in detail in Biset et al. (2009) and is shown in Fig. 2. The most critical controlled variables were selected taking into account the objectives mentioned above, directly related to the optimal operating conditions of the FPS and the less sensitivity structure in terms of the regulatory problem. For analyzing the dynamic behavior, step changes of -5% at 0.25 seconds were made in the manipulated variables, that is, ethanol feed flow rate to the burner, heat exchanged in the ESR, and ethanol feed flow rate to ESR. The effects on the variables to be controlled were recorded, as can be seen in Fig. 3. Around this

operating point, the variables were linearized in order to obtain an acceptable nominal model for implementing internal model control (IMC) method. The tuning parameters were chosen following the recommendations given in Rivera (2007).

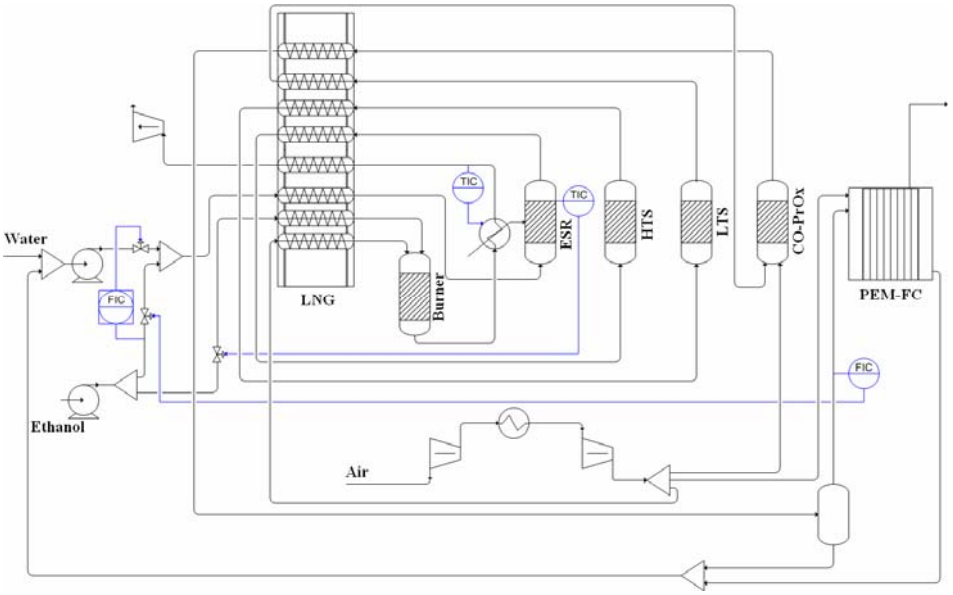


Figure 2. Implemented control structure.

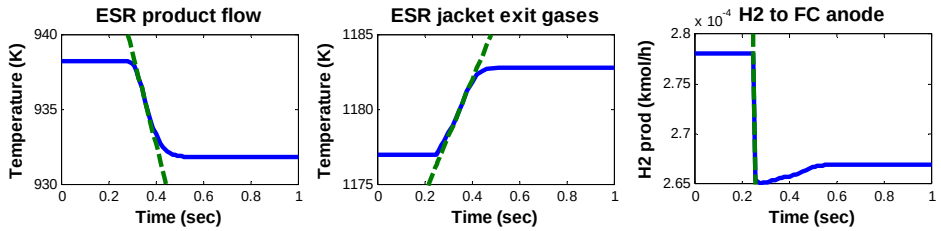


Figure 3. Linearization of controlled variables for controllers tuning.

5. Results

Since the authors of this work didn't find any reference of a dynamic model similar to the one presented here, the resulting responses were compared to the stationary states presented by Francesconi (2008) for the 10 KW FC, and the step changes considered at Biset et al. (2009). The comparison to those values is shown in Fig. 4. The applied steps were -1%, -5% and -10% at 20, 80 and 140 seconds, respectively. The first group of figures, on the left, corresponds to a step change disturbance in the current I_{st}, and the second group, on the right, in fresh ethanol molar fraction.

6. Conclusions

In this work, the main characteristics of the rigorous dynamic model of a Fuel Processor System along with a Fuel Cell Stack are presented. Both systems have been integrated so as to obtain good efficiencies and maximum heat recovery. Keeping the plant in desired conditions requires a well designed plant-wide control structure. In this context, the control oriented model makes it possible to obtain a preliminary control policy accounting with a good understanding of the dynamic interactions. The model is

flexible enough allowing to analyze other possible kinetics, sizing, etc., accounting different scenarios. As future work it is evaluated the use of this dynamic model for testing new methodologies of plant-wide control structure and optimal sensor location, including experimental data.

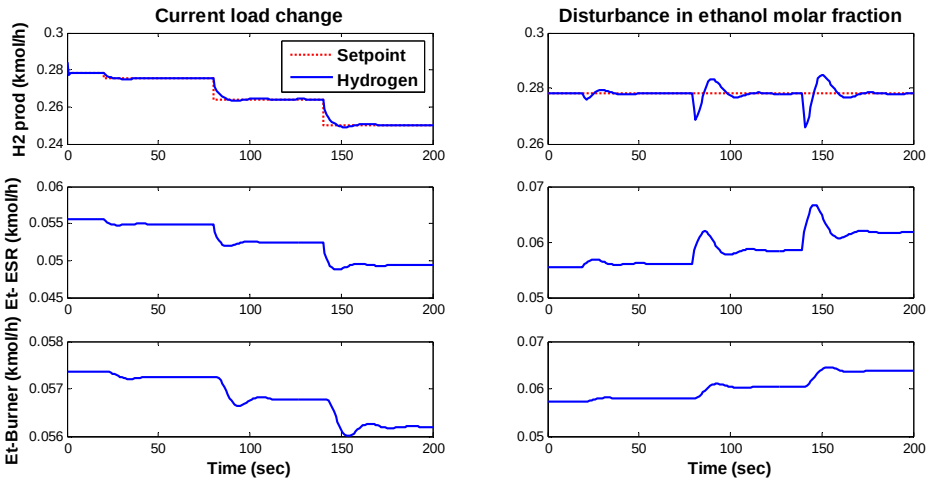


Figure 4. Dynamic simulations of H_2 production, controlled variable, and manipulated variables, feed flow rate ethanol to ESR and burner reactors under perturbation.

References

M. Benito, J.L. Sanz, R. Isabel, R. Padilla, R. Arjona, and L. Daza, 2005, Bio-ethanol steam reforming: Insights on the mechanism for hydrogen production, J. Power Sources, 151:11–17.

S. Biset, L. Nieto Degliuomini, M. S. Basualdo, V.M. García, and M. Serra, 2009, Analysis of the control structures for an integrated ethanol processor for proton exchange membrane fuel cell, J. Power Sources, DOI:10.1016/j.powsour.2008.12.099.

Y. Choi and H. G. Stenger, 2004, Kinetics, simulation and insights for CO selective oxidation in fuel cell applications, J. Power Sources, 129(2):246–254.

C. C. Chuang, Y. H. Cheng, J. D. Ward, C. C. Yu, Y. C. Liu, and C. H. Lee, 2008, Optimal design of an experimental methanol fuel reformer, International Journal of Hydrogen Energy, 33:7062–7073.

J. A. Francesconi, 2008, Modelado, Síntesis y Optimización del Proceso de Reformado de Bioetanol para la Producción de Hidrógeno grado PEM, PhD thesis, FIQ-Universidad Nacional del Litoral.

H. Görgün, 2005, Control-oriented modeling of fuel processing reactors in fuel cell power systems, Journal of Electrical and Electronics Engineering, 5(1):1279–1286.

H. Görgün, M. Arcak, S. Varigonda, and S. A. Bortoff, 2004, Nonlinear observer design for fuel processing reactors in fuel cell power systems, Proceedings of the American Control Conference, 845–850.

B. Linnhoff, P. Townsend, P. Boland, G. F. Hewitt, B. E. A. Thomas, A. R. Guy, and R. H. Marsland, 1994, A User Guide on Process Integration for the Efficient Use of Energy, Institute of Chemical Engineers.

J.T. Pukrushpan, A.G. Stefanopoulou, and H. Peng, 2004, Control of Fuel Cell Power Systems: Principles, Modeling, Analysis and Feedback Design, Springer.

D. E. Rivera, 2007, Una metodología para la identificación integrada con el diseño de controladores IMC-PID, Revista Iberoamericana de Automática e Informática Industrial, 4(4):129–134.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Multiperiod Production Planning and Design of Batch Plants under Uncertainty

Marta S. Moreno[a,b] and Jorge M. Montagna[a]

[a]*INGAR - CONICET, Avellaneda 3657, Santa Fe (S3002GJC), Argentina*
[b]*FCEFyN-UNC, Velez Sarsfield 1611, Córdoba (X5016GCA), Argentina*

Abstract

A general multiperiod LGDP (Linear General Disjunctive Programming) model in a multiproduct batch plant has been developed. Both overall production planning and design decisions of the plant are encompassed considering uncertainty in demands represented by a set of scenarios. Specifically, the proposed model considers simultaneously the plant structure (duplication of units in parallel and the new option "in series"), the decision variables of the batch plant (unit sizes, number of batches, total time for producing each product), and the production planning decisions (e.g. plan of production, policy of inventory, purchases of raw materials and sales of products) in each time period within each scenario. Also, this model allows the incorporation of new equipment items in different periods. The overall objective is to maximize the expected net present value of the benefit. The benefits of the simultaneous optimization of production planning and design decisions under uncertain product demands in a multiscenario approach are demonstrated through a particular extraction process that produces oleoresins.

Keywords: Multiproduct batch plants, Demand uncertainty, Units in series, Design and planning.

1. Introduction

An important area of concern in the design of multiproduct batch plants is their ability to meet production requirements and maximize profits given uncertainties in the market demand for the products. The formulations presented in previous works in the multiproduct batch design area (Moreno et al., 2007; Moreno and Montagna, 2007) have been limited to deterministic approaches, wherein product demands are assumed to be known with certainty. However, when long time periods are assumed a mathematical formulation including uncertainty is more appropriate. Uncertainty is usual when long term variations are considered, for example with product demands forecasting. The scenario analysis approach has been considerably used in the literature and has proven to provide reliable and practical results for optimization under uncertainty (Subrahmanyam et al., 1994; Ahmed and Sahinidis, 1998; Gupta and Maranas, 2003). Hence, in this work, this approach is adopted for describing uncertainty in the random variables of product demands. Despite that many contributions dealing with the design of batch plants under uncertainty have been published, the simultaneous optimization of the design and planning conditions with capacity expansion of the plant has not been sufficiently studied including all the elements considered in this work.

The goal of this work is to propose a scenario-based approach for the simultaneous design and production planning of multiproduct batch plants under uncertain demands, considering different duplication options of units (in series and in parallel). Also, this

model allows a capacity expansion adding a set of new units in parallel working out of phase in different periods. A linear general disjunctive programming model is formulated for the optimization problem. The proposed model is finally tested by solving an example considering the production of oleoresins.

2. Problem Definition

Consider a multiproduct batch plant with $p = 1,\ldots, P$ operations that processes $i = 1,\ldots, I$ products over a time horizon H. These products are produced during a finite number of time periods $t = 1,\ldots, T$. Let H_t be the length of each time period, which may be different. Forecasts for market demands for each product come in a finite number of possible scenarios $s = 1,\ldots, S$, each of which has an associated probability ψ_s, where $\sum_s \psi_s = 1$. Thus, the different possible scenarios describe lower and upper bounds on product demand levels in each time period d_{its}^L and d_{its}^U. In this first work, it is assumed that there is a procedure for generating a set of scenarios representative of both optimistic and pessimistic situations within a risk analysis strategy (Vanston et al., 1977; Mulvey, 1996).

For each operation p there may be different configurations of units in series $h = 1,\ldots, H_p$ to performed it. Also, the selected configuration of units in series can be duplicated in parallel working out of phase. The duplication in parallel can be different in each time period allowing the capacity expansion of the plant. Let M_p^U be the maximum number of units that can be allocated in parallel in operation p. In each time period, $m = 1,\ldots, M_p^U$ sets of units of the same size operate out of phase. This amount can vary taking into account that each operation can be expanded with new sets of units. N_{pt} represents the number of set of units in parallel in operation p at each period t. This value is modified taking into account that $g = 0, 1,\ldots, M_p^U$ set of units in parallel can be added in each period t. Furthermore, the problem involves the design decision of selecting equipment sizes for batch units in each operation p, among a set $SV_p = \{ v_{p1}, v_{p2}, \ldots, v_{pn_p} \}$ of available discrete sizes. The basic data for representing the operations are the size factors S_{ipt} and processing time t_{ipt} required for each product i at each operation p in every period t.

The duplication in series selected in every operation not only affects itself but also the rest of the operations of the process. In order to maintain a posynomial approach, the yield for all the configurations in series in a given operation is assumed to be constant, through appropriate size factors values. This allows preserving the formulation with fixed size and time factors. In consequence, the size factor for product i at operation p remains equal regardless of the selected configuration. However, each configuration h of units in series now has a different time t_{ipht}. Detailed descriptions of this assumption can be found in Moreno and Montagna (2007).

In every scenario s, production planning decisions allow to determine at each period t and for each product i, the amount to be produced q_{its}, the number of batches n_{its}, and the total time T_{its} to produce product i. Furthermore, at the end of every period t, the levels of both final product IP_{its} and raw material inventories IM_{its} are obtained. Moreover, the total sales QS_{its}, the amount of raw material purchased C_{its}, and the raw material to be used for the production RM_{its} of product i in each period t are determined with this formulation. If time periods are equal, wastes due to the expired product shelf life PW_{its} and due to the limited raw material lifetime RW_{its} are also added in the

formulation. Also, late deliveries ϑ_{its} that take place in each period are determined. The objective is to determine the optimal design and the operation planning of the plant, in the face of uncertainties in market demand for products.

3. Linear Disjunctive Programming Model

Based on the above definitions, the optimization LPDG model is as follows.

$$\max NPV = \sum_s \psi_s \left[\begin{array}{l} \sum_{i,t} np_{it}\ QS_{its} - \sum_{i,t} \kappa_{it}\ C_{its} - \sum_{i,t} \varepsilon_{it} \left(\frac{IM_{i,t-1,s} + IM_{its}}{2} \right) H_t \\ - \sum_{i,t} \sigma_{it} \left(\frac{IP_{i,t-1,s} + IP_{its}}{2} \right) H_t - \sum_{i,t} \left(wp_{it}\ PW_{its} + wr_{it}\ RW_{its} \right) \\ - \sum_{i,t} \left(co_{it}\ q_{its} \right) - \sum_{i,t} \left(cp_{it}\ \vartheta_{its} \right) \end{array} \right] - \sum_t CE_{pt} \quad (1)$$

$$\underset{h \in H_p}{\vee} \left[\begin{array}{c} Z_{ph} \\ \underset{k \in SV_p}{\vee} \left[\begin{array}{c} W_{phk} \\ n_{its} \geq \left(\frac{S_{ipt}}{v_{pk}} \right) \cdot q_{its} \quad \forall\ i,t,s \\ CO_p = h \cdot (\alpha_p \cdot v_{pk}^{\beta_p}) \end{array} \right] \\ \underset{m \in M_p}{\vee} \left[\begin{array}{c} Y_{phmt} \\ N_{pt} = m \\ T_{its} \geq \frac{t_{ipht}}{m} \cdot n_{its} \quad \forall\ i,s \end{array} \right] \quad \forall\ t \\ \underset{g \in G_p}{\vee} \left[\begin{array}{c} X_{phgt} \\ N_{pt} = N_{p,t-1} + g \\ CE_{pt} = g \cdot CO_p \cdot \gamma_{pt} \end{array} \right] \quad \forall\ t \end{array} \right] \quad \forall\ p \quad (2)$$

$$\sum_i T_{its} \leq H_t \quad \forall\ t,s \quad (3)$$

$$IP_{its} = IP_{i,t-1,s} + q_{its} - QS_{its} - PW_{its} \quad \forall\ i,t,s \quad (4)$$

$$IM_{its} = IM_{i,t-1,s} + C_{its} - RM_{its} - RW_{its} \quad \forall\ i,t,s \quad (5)$$

$$IP_{its} \leq \sum_{\tau=t+1}^{t+\chi_i} QS_{i\tau s} \quad \forall\ i,t,s \quad (6)$$

$$IM_{its} \leq \sum_{\tau=t+1}^{t+\zeta_i} RM_{i\tau s} \quad \forall\ i,t,s \quad (7)$$

$$\vartheta_{its} \geq \vartheta_{i,t-1,s} + d_{its}^L - QS_{its} \quad \forall\ i,t,s \quad (8)$$

$$RM_{its} = F_{it}\ q_{its} \quad \forall\ i,t,s \quad (9)$$

The objective function maximizes the expected net present value (*NPV*) over a set of scenarios *S*. This function is the difference between the revenue due product sales and the overall costs, with the latter consisting of the cost of raw materials, inventory costs, operating cost, penalty cost for late delivery, and investment cost. The parameters np_{it},

κ_{it}, ε_{it}, σ_{it}, wp_{it}, wr_{it}, co_{it}, and cp_{it} are the corresponding cost coefficients for every mentioned term in the objective function.

The discrete choices are modeled with the embedded disjunctions (2) defined for each operation p. A Boolean variable Z_{ph} is true when configuration of units in series h is selected in operation p and is false in the opposite case. In the first embedded disjunctions, the Boolean variables W_{phk} is true when discrete size k is selected to carry out operation p with configuration h. Thus, constraints into this disjunction correspond to the sizing equation for the units and the equipment cost for this alternative CO_p. As was previously mentioned, units in parallel are added in each time period, so a set of embedded disjunctions is added where the Boolean variable Y_{phmt} is true when there are m units in parallel in operation p with configuration h at time period t. Each term of these disjunctions includes constraints that determine the number of set of units in parallel N_{pt} at each period t and the total time to produce each product. Finally, in the last embedded disjunctions, if the Boolean variable X_{phgt} is true, g units in parallel are added at time period t in operation p with configuration h. The first constraint in the last disjunctions determines the number of units in each period t considering the number in the previous one plus the units aggregated in the corresponding period. The last constraint determines the expansion cost in each time period, CE_{pt}, where the parameter γ_{pt} is a cost coefficient for operation p taking into account the period t involved.

Constraint (3) ensures that the total time required for producing all products cannot exceed the period length. Eqs. (4) and (5) state the inventory level of each product and raw material at the end of each period, respectively. When the lengths of time periods are equal, Eqs. (6) and (7) guarantee that the stock of both raw materials and products in each period cannot be used after the next ζ_i or χ_i time periods, respectively. Constraint (8) calculates the late delivery. The mass balance (9) determines the amount of raw material necessary for the production of product i in each period where the parameter F_{it} is the process conversion of product i in period t assuming that only one main raw material is used for the production of i.

4. Illustrative Example

To illustrate the proposed approach, a multiproduct batch plant that produce three oleoresins, namely, sweet bay (A), pepper (B) and thyme (C) oleoresins, is considered. All the products are manufactured via the following batch operations: (1) extraction, (2) expression, (3) evaporation, and (4) blending. The process data for this example are shown in Table 1 and prices of raw materials and final products are given in Table 2. A global time horizon of 3 years has been considered, which is divided in 6 equal time periods of six months each (3000 h). There are 4 possible configurations of units in series for the operation 1, with a countercurrent arrangement. Thus, processing times for each product take smaller values as the number of units in series grows (see Table 1). All the operations can be duplicated in parallel up to 3 set of units operating out of phase. A set of 5 discrete sizes is provided for process units. Table 3 shows the available discrete sizes for each operation and cost coefficient associated. Demand for the products is uncertain and 3 forecast scenarios with probabilities $\psi_1 = 0.5$, $\psi_2 = 0.3$, and $\psi_3 = 0.2$ are considered. In the first time period all the scenarios show the same upper demand, i.e., 12000, 9500 and 14000 kg, for product A, B, and C, respectively. It is assumed that product demand for each period will increase in comparison with the present conditions. Thus, scenario 1 expects a 20% growth rate per period, scenario 2 expects a 10% growth rate per period and for scenario 3 no growth rate is expected. Minimum product demands in each period for all scenarios are assumed as a 50% of

product demands. Products and raw materials lifetimes in time periods are 3 and 2, respectively. The problem was solved with GAMS via CPLEX solver with 0% optimality gap.

Table 1. Process data for the example

	Size factors, S_{ipt} (L/kg)				Processing time, t_{ipt} (h)							Conversion Factor
i	1	2	3	4	1 (h_1)	1 (h_2)	1 (h_3)	1 (h_4)	2	3	4	F_{it}
A	20	15	12	1.5	25.95	9.28	5.35	3.47	1	2.5	0.5	11.11
B	23	15	12	1.5	39.46	9.76	5.55	3.59	2.0	1.5	2.0	11.11
C	30	20	17	1.5	27.93	9.41	5.41	3.51	1.0	2.0	1.0	15.87

Table 2. Economic data

	Costs of raw materials, κ_{it} ($/kg)						Prices of products, np_{it} ($/kg)					
Period	1	2	3	4	5	6	1	2	3	4	5	6
A	1.5	2.2	1.5	2.2	1.7	2.4	36	38	36	38	36	38
B	2.5	2.5	2.5	2.5	2.7	2.7	40	40	40	40	40	40
C	1.0	1.0	1.0	1.0	1.0	1.0	37	35	37	35	37	35

Table 3. Available standard sizes

		Discrete volumes, v_{pk} (L) Operation			
Option		1	2	3	4
1		500	500	250	50
2		1000	700	500	100
3		1500	1000	750	150
4		2500	1500	1000	200
5		3000	2000	1500	250
Cost coef	α_p	875	1150	800	700
Cost exp	β_p	0.6	0.6	0.6	0.6

This example has 498 discrete variables, 647 continuous variables, and 4745 constraints and was solved in a CPU time of 170.34 s. The optimal solution yields an expected NPV of $2,546,346.65. The optimal plant structure obtained in every period and the unit sizes are shown in Fig. 1. In this figure, units in dotted line are included in different time periods. As can be seen, in time period 1 there is only one unit in all operations except in operation 1 which has 2 units in series. In period 2, a set of 2 units in series is incorporated in operation 1, working in parallel out of phase. Similarly, in period 3 a new set of units is added in extraction operation and the same structure is maintained until period 6. Production planning decisions for each product are different for each scenario in the optimal solution. In Table 4, the results of planning decisions for scenario 1 are reported.

Table 4. Optimal plan for scenario 1

	A ($\times 10^3$g)					B ($\times 10^3$kg)					C ($\times 10^3$kg)				
t	q_{it}	QS_{it}	IP_{it}	C_{it}	IM_{it}	q_{it}	QS_{it}	IP_{it}	C_{it}	IM_{it}	q_{it}	QS_{it}	IP_{it}	C_{it}	IM_{it}
1	8.6	8.6	0.0	255.6	160	4.7	4.7	0.0	52.8	0.0	7.0	7.0	0.0	111.1	0.0
2	14.4	14.4	0.0	0.0	0.0	8.2	8.2	0.0	90.2	0.0	16.8	16.8	0.0	266.7	0.0
3	34.9	17.3	17.6	422.4	34.8	6.8	6.8	0.0	76.0	0.0	20.2	20.2	0.0	896.3	0.0
4	3.1	20.7	0.0	0.0	0.0	16.4	16.4	0.0	182.4	0.0	34.4	24.2	10.2	546.2	0.0
5	33.8	24.8	8.8	456.4	81.3	9.8	9.8	0.0	109.4	0.0	18.8	29.0	0.0	298.6	0.0
6	7.3	16.2	0.0	0.0	0.0	11.8	11.8	0.0	131.3	0.0	34.8	34.8	0.0	552.9	0.0

Also, each scenario has been considered in isolation by formulating and solving the corresponding deterministic problem. The objective values obtained for scenarios 1, 2, and 3 are \$3,029,459.79, \$2,586,374.59 and \$2,223,378.12, respectively. In the first scenario, the optimal plant structure obtained is the same as Fig.1. On the other hand, for both scenarios 2 and 3, in the first time period, two sets with two units in series each one operate in parallel in operation 1 and a new set is added in period 2 for this operation. The unit sizes selected for each operation are 1000L, 1000L, 750L, and 100L.

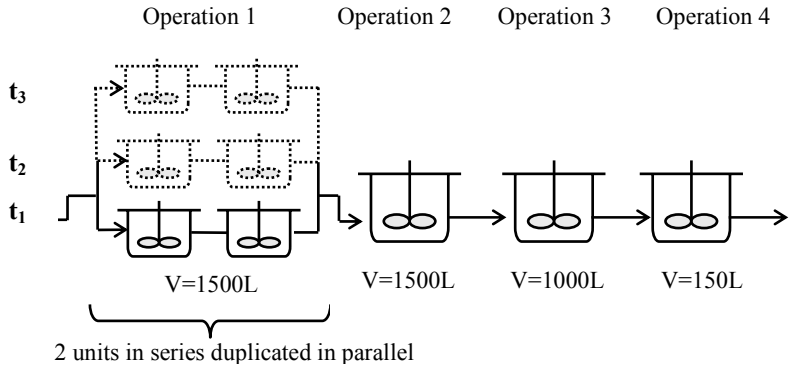


Figure 1. Optimal structure of the plant.

5. Conclusion

This work has presented a LGDP model for the optimal production planning and design of batch plants in the presence of uncertainties in product demands. The model takes into account both structural decisions of duplicating units in series and in parallel selecting the size of the units from a set of available discrete sizes. The proposed approach allows the incorporation of parallel units in different time periods. Design variables such as the selection of equipment of standard size and the addition of new units in parallel in each time period are independent of the scenarios. In contrast, planning variables which include working levels of the plants for each time period are scenario-dependent variables. As an example, the model was applied to the production of vegetable extracts, particularly the production of oleoresins.

References

S. Ahmed, N.V. Sahinidis, 1998, Robust process planning under uncertainty, Ind. Eng. Chem. Res., 37, 1883-1892.

A. Gupta, C. D. Maranas, 2003, Managing demand uncertainty in supply chain planning. Computers Chem. Engng., 27, 1219-1227.

M.S. Moreno, J.M. Montagna, O.A. Iribarren, 2007, Multiperiod optimization for the design and planning of multiproduct batch plants, Computers Chem. Engng., 31, 1159-1173.

M.S. Moreno, J.M. Montagna, 2007, Optimal simultaneous design and operational planning of vegetable extraction processes. Trans IChemE, Part C, Food Bioprod Proc, 85,(C4), 360-371.

J.M. Mulvey, 1996, Generating scenarios for the Towers Perrin Investment System, Interfaces, 26, 1-15.

S. Subrahmanyam, J.F. Pekny, G.V. Reckaitis, 1994, Design of batch chemical plants under market uncertainty. Ind. Eng. Chem. Res., 33, 2688-2701.

J.H.Vanston, W.P. Frisbie, S.C. Lopreato, D.L. Poston, 1977, Alternate scenario planning, Technol. Forecast. Social Change, 10, 159-180.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Combined Approach for Operational Scheduling of a Real-World Pipeline Network

Suelen N. Boschetto,[a] Flávio Neves-Jr,[a] Leandro Magatão,[a] William M. Brondani,[a] Lúcia V.R. Arruda,[a] Paulo C. Ribas[b]

[a]*Federal University of Technology – Paraná (UTFPR / CPGEI)*
Av. Sete de Setembro, 3165, 80230-901, Curitiba - PR, Brazil
[b]*Logistic Division-Research and Development Centre, PETROBRAS-CENPES*
Rio de Janeiro/RJ, Brazil

Abstract

This work addresses the development of an optimization model to aid the operational decision-making of scheduling activities in a real-world pipeline network. In this network, inventory issues should be respected and different products should be transported in order to supply demands. Scheduling details must be given, including pumping sequence in each node, volume of batches, tankage constraints, timing issues, while respecting a series of hard operational constraints. In addition, temporal details must be specified by the system. Thus, the computational burden of determining the short-term scheduling within the considered scenario is a relevant issue. Therefore, a decomposition approach, based on the three key elements of scheduling, is proposed to address such real-world problem. In particular, the focus of this work is to solve the third key element: the timing problem. Such model is based on a combined approach, which uses a Pre-Analysis structure and Mixed Integer Linear Programming (MILP) in a collaborative combination. Many insights have been derived from the obtained solutions, and the proposed approach can aid the decision-making process.

Keywords: Scheduling, Pipeline Network, MILP, Heuristics, Oil Industry

1. Introduction

Nowadays, the majority of the oil derivatives transport is done using pipeline network systems, mainly because of the economical and environmental advantage, when compared to other transport systems. However, the determination of the scheduling of operational activities in pipeline networks is a difficult task and efficient methods to solve such complex problem are required. Mendéz *et al.* (2006) make a review of the optimization methods for the short-term scheduling of batch processes, and some recent papers have also addressed scheduling decisions within pipeline networks (*e.g.* Relvas *et al.* (2007), Moura *et al.* (2008), Cafaro and Cerdá (2008), Rejowski and Pinto (2008)), but the network considered in this paper is particularly complex. Boschetto *et al.* (2008), in a previous work, already introduced the proposed approach but the present paper gives additional information and extends the optimization model in a continuous process of improvement. This work describes an optimization model to aid the operational decision-making of scheduling activities in a real-world pipeline network, in which inventory issues should be respected and different products should be transported in order to supply demands. The complete problem of short-term scheduling was divided into smaller problems, which are based on the three key elements of scheduling: assignment of resources, sequencing of activities, and timing determination for

resources utilization by these activities (Reklaitis, 1992). Each problem is solved by an appropriate technique and integrated by a computational tool. In particular, the focus of this work is to solve the third key element: the timing problem. The assignment of resources and the batch sequencing are previously solved by Yamamoto *et al.* (2008) and the results are used by the timing model. Such model is based on a combined approach, which uses a Pre-Analysis structure and Mixed Integer Linear Programming (MILP) in a collaborative combination. This paper is organized as follows. Section 2 presents some operational details of the considered pipeline network. The problem solution, including a global overview of the proposed optimization approach, is given on Section 3. Section 4 presents the obtained results with conclusions in Section 5.

2. Problem Description

The considered scenario involves 14 areas (nodes), including 4 refineries, 2 harbors, which either receives or sends products, 6 distribution centers, and 2 final clients (Fig. 1). In addition, it includes 29 pipelines, and some of them can have the flow direction reverted, accordingly to operational procedures. Each product presents a specific tank farm, according to the considered node. More than 14 oil derivatives can be transported. During the scheduling horizon, many batches are pumped from (or passing through) different nodes. Scheduling details must be given, including pumping sequence in each node, volume of batches, tankage constraints, timing issues, while respecting a series of hard operational constraints including on-peak demand hours.

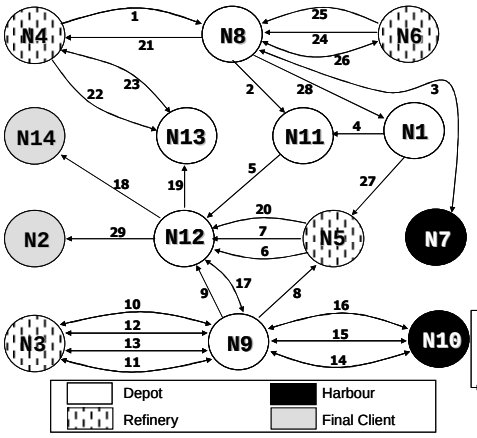


Figure 1. Pipeline Network

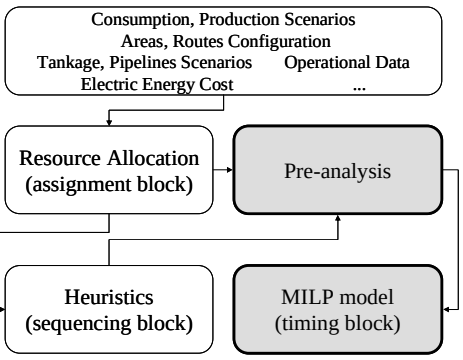


Figure 2. Decomposition Approach

3. Problem Solution

A decomposition approach is proposed to address the scheduling problem (Fig. 2). The input data (consumption/production conditions, areas/routes configuration, inventory management, electrical energy cost, and other operational data) characterizes a "problem instance", or a scenario, to be solved. The assignment block gets part of the information and uses a heuristic procedure to compute batch volumes and to indicate a route for each batch. The choice of the routes considers that incompatible products should be sent by different pipelines. So, undesirable contamination areas are avoided. The assignment takes into account production and consumption information received from the company planning, inventory management information, and available tank capacity for each area. Then, based on the input data, the assignment block calculates

time-windows to send/receive each batch in each origin/destination area, respectively. A list of batches, including the origin/destination area, route, volume, flow rate, and the time-windows for each batch are determined. The heuristic block (sequencing block) can be used or not. It reorganizes the list of batches previously sequenced by the assignment block. This module uses a GA algorithm to optimize the sequencing of batches to be pumped and it takes as inputs the previously calculated time windows, start-up conditions, and inventory levels given by the assignment block. The main goal is to provide one or more sequences of ordered batches that satisfy the consumption/production requirements for each area. To deeply exploit the assignment/sequencing blocks are not within the aim of this article. Further details of assignment/sequencing blocks can be attained in Yamamoto *et al.* (2008). The pre-analysis block evaluates the information derived from assignment/sequencing blocks and calculates temporal and volumetric limits (bounds). These limits present a preliminary indication on the feasibility of the programming. So that, the pre-analysis block provides valuable information, which is used as parameters by the MILP model. Finally, all the parameters calculated by previous blocks are used in a continuous time MILP model. This block determines the operational short-term scheduling for the pipeline network using, for instance, time windows and sequences of batches determined by previous blocks.

3.1. Pre-analysis

The pre-analysis was developed using the languages C++ and PL/SQL. This block allows the identification of conditions in which a batch can remain stopped into the pipeline and also indicates significant flow rate changes due to network constraints. Moreover, the pre-analysis evaluates, step by step, the batch that is being pumped and analyses the influence of this pumping operation on other batches into the pipelines. As a result, the pumping of a batch along the pipeline network is analyzed, and "volumetric parts" (portions of this batch) with similar pumping conditions are identified. The parameters derived from pre-analysis (*e.g.* Table 1) contain information to indicate the receiving of a batch in the end of a pipeline. By Table 1 it is possible to notice that the pumping procedure of the batch 40 (bo_b) was broken in four parts (p_{bo}=1,2,3,4). Each part influences the receiving of one or more batches that will be moved in the pipeline network. For example, the second part (p_{bo}) pumped of the batch 40 with 300 m³ ($part_{bo,p_{bo}}$) influences the receiving of the same volume of the part 2 (p_b) of the batch 14 (b) in the pipeline 18 (d) and of the part 1 (p_b) of the batch 24 (b) in the pipeline 5.

Table 1. Parameters determined by pre-analysis

b	d	bo_b	p_b	p_{bo}	$part_{bo,p_{bo}}$ (m³)	$flw_{bo,p_{bo}}$ (m³/h)
14	18	40	1	1	1700	500
14	18	40	2	2	300	600
24	5	40	1	2	300	600
24	5	40	2	3	1700	500
24	5	40	3	4	350	550
24	5	41	4	1	3350	600

3.2. MILP Model

A continuous time MILP approach is used to create the optimization model. The MILP model considers the route of a batch (origin area, pipelines, and destination area), the volume and the flow rate of each product from an origin area. Special attention is given to the fact that the pipelines have a considerable volume and always operate with full capacity. Also, they can "store" different products. When a new product is pumped from

an origin area, the products previously "stored" into the pipelines are pushed, in accordance with the "new flow rate". In each area, the batches arriving from the pipelines can be sent to tanks or can be deviated to other pipelines. A set of tanks in each area can store different products. The inventory level can increase or decrease according to the volume and the pumping flow rate of each product, or in accordance with the production and the "local" consumption. Moreover, the MILP model must consider the cost of electric energy in on-peak hours and a series of other operational requirements. The optimization model must satisfy a series of constraints. Timing constraints are added in order to make a connection among timing variables (*e.g.* $fb_{b,n,n',d,p}$ and $ib_{b,n,n',d,p}$), calculating the time that each part of a batch will be pushed into the pipelines. Other constraints are able to identify which batch will increase the total pumping time because of on-peak hours. An example of a constraint that uses the information obtained by the pre-analysis is shown in expression (1). The final sending (fb_{b,n,n',d,p_b}) of a batch *b* is calculated considering its start sending (ib_{b,n,n',d,p_b}) in the same pipeline *d*, added to the pumping time ($part_{bo,p_{bo}} / flw_{bo,p_{bo}}$) of another batch (bo_b) that has influenced the movement of the batch *b* (see Tab. 1). The batch bo_b is pumped in its origin area no_{bo} and can remain stopped α hours because of on-peak demand hours, if the binary variable $z_{bo,p_{bo},no_{bo},h}$ = 1. Finally, some constraints describe inventory management issues (*e.g.* the time windows violation). They can be violated, but a penalty is added in the objective function in order to minimize these violations. The objective function of the MILP model also minimizes the total scheduling horizon (the sum of pumping and receiving variables).

$$fb_{b,n,n',d,p_b} \geq ib_{b,n,n',d,p_b} + \left(part_{bo,p_{bo}} / flw_{bo,p_{bo}}\right) + \alpha \cdot \sum_{h \in HP} z_{bo,p_{bo},no_{bo},h} \quad (1)$$

4. Computational Results

The optimization model has been extensively tested in typical operational scenarios. The scenario herein presented, involves 247 batches for a month of planning. This typical instance yielded a large-scale MILP model with 9695 variables (2940 binary) and 38672 constraints. This model was solved to optimality in approximately 60 seconds using ILOG OPL Studio 5.5.1 and the solver CPLEX 11 (ILOG, 2007). The hardware was an Intel Core 2/2.13GHz processor with 2GB of RAM memory. The developed scheduling presents results with a detailed form, indicating various parts for a same batch. Fig. 3 shows the temporization result for the 11 first days of the scheduling horizon. It is presented all the used pipelines, being possible, therefore, to get a general vision on how the management of the pipeline network can be made. It is possible to observe various timing blocks (parts) with the same batch index in Fig. 3. For instance, there are four parts of the same batch 71 in pipelines 24, 28, and 4. So that, the route of batch 71 contemplates the following areas and pipelines: N6 → 24 → N8 → 28 → N1 → 4 → N11 (see Fig. 1). The pre-analysis block divides a batch in various parts and determines the number of parts that will be pumped (or received), supplying the MILP main model through parameters. The difference in size of batches among pipelines is a consequence that a batch is moved by another batch that is pumped in its origin area. So that, the flow rate and the moved volume vary in each pipeline. Moreover, some stoppages can be visualized in Fig. 3 through separations between parts of the same batch in a determined pipeline (*e.g.* batch 71 on pipelines 24, 28, and 4). Based on pumping and receiving times, it is possible to obtain the storage profile in diverse areas

for different products. The initial storage level of each tank, production campaigns, and local market consumption are also considered. In Fig. 4 is shown a typical tankage profile for the product naphtha in different areas. In particular, Fig. 4 presents the naphtha aggregated inventory forecast for a 30-day horizon. Since the area N14 (blue line) does not have a local production, then, an increasing "curve" indicates, in this case, the receiving of a batch of naphtha. In addition, N14 is a consumption area. So, the decreasing curves show the local consumption of naphtha in this area. The different increasing inclinations indicate receiving of (parts of) batches at different flow rates.

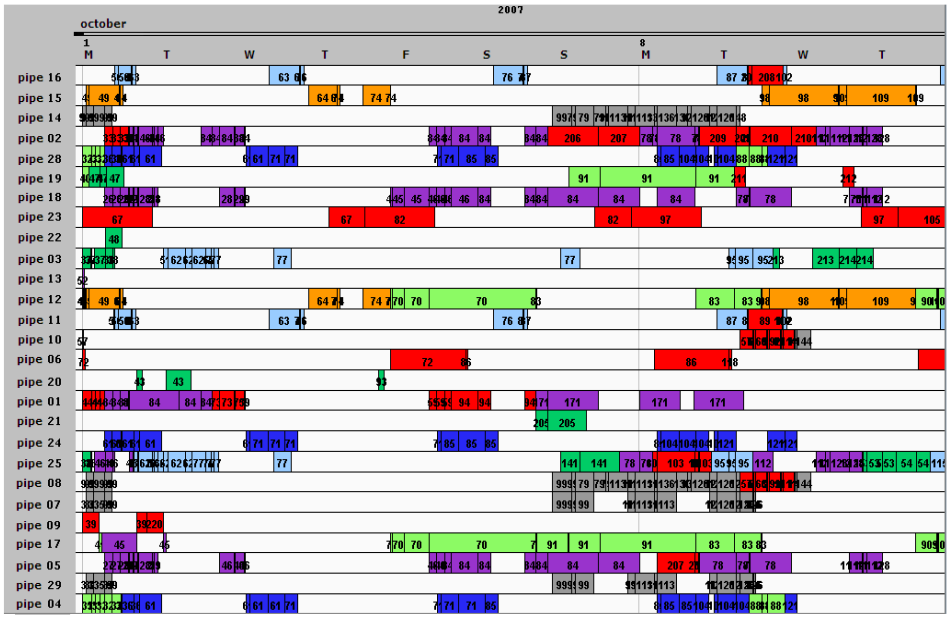


Figure 3. Gantt chart of sending procedure

Figure 4. Naphtha storage at different areas

5. Conclusions

In this paper, a decomposition approach, based on the three key elements of scheduling, is developed to address the scheduling of a real-world pipeline network. In particular, this paper focused on solving the third key element: the timing problem. Such problem is addressed through a Pre-Analysis structure and a MILP model in a collaborative combination. The optimization system has been applied to solve scenarios that represent real operational conditions of a multiproduct pipeline network that transport oil derivatives and organic products. The difficulty of obtaining a feasible scheduling of operational activities within this network is a day-to-day problem faced by the company schedulers, and such framework can be used as a decision support tool. The developed optimization system can attain a computational solution in a reduced time (order of seconds). Moreover, through the development of the Pre-analysis block, and using the parameters obtained by this block, the MILP model addressed in detail some operational conditions that were hard to be identified, even by the specialists. The optimization system is in phase of validation by the oil industry, and its use is expected to provide benefits from an operational/managerial point of view. The optimization of the transference procedures and the storage issues can reduce operational losses.

Acknowledgements

The authors acknowledge financial support from ANP and FINEP (PRH-ANP / MCT PRH10 UTFPR), CENPES (Under grant 0050.0044324.08.4) and CAPES/PDEE (Under grant 3262/08-1).

References

S.N. Boschetto; L.C. Felizari; L. Yamamoto; L. Magatão; S.L. Stebel; F. Neves-Jr; L.V.R. Arruda; R. Lüders; P.C. Ribas; L.F.J. Bernardo, 2008, An Integrated Framework for Operational Scheduling of a Real-World Pipeline Network, In: 18th European Symposium on Computer Aided Process Engineering (ESCAPE 18), Lyon, France, Computer-Aided Chemical Eng., 25, Eds: B. Braunschweig; X. Joulia, Amsterdam, 259-264.

D. Cafaro; J. Cerdá, 2008, Dynamic scheduling of multiproduct pipelines with multiple delivery due dates, Computers and Chemical Engineering, 32, 728-753.

ILOG, 2007, ILOG OPL Development Studio 5.5: User's Manual, ILOG S.A., France.

C.A. Méndez; J. Cerda; I.E. Grossmann; I. Harjunkoski; M. Fahl, 2006, State-of-the-art review of optimization methods for short-term scheduling of batch processes, Computers and Chemical Engineering, 30, 913-946.

A.V. Moura; C.C. Souza; A.A. Cire; T.M.T. Lopes, 2008, Heuristics and Constraint Programming Hybridizations for a Real Pipeline Planning and Scheduling Problem, In 11th IEEE International Conference on Computational Science and Engineering, 445-462.

R. Rejowski; J. Pinto, 2008, A novel continuous time representation for the scheduling of pipeline systems with pumping yield rate constraints, Computers and Chemical Engineering, 32, 1042-1066.

G.V. Reklaitis, 1992, Overview of scheduling and planning of batch process operations, In Proceedings of the NATO Advanced Study Institute on Batch Processing Systems, Antalya, Turkey, 660-705.

S. Relvas; H.A. Matos; A.P.F.D. Barbosa-Póvoa; J. Fialho, 2007, Reactive Scheduling Framework for a Multiproduct Pipeline with Inventory Management, Industrial & Engineering Chemistry Research, 46 (17), 5659-5672.

L. Yamamoto; L.C. Felizari; S.N. Boschetto; L.V.R. Arruda; L. Magatão; S.L. Stebel; F. Never-Jr; R. Lüders; P.C. Ribas, 2008, Using GA and MILP to solve the short-term scheduling of a real multiproduct pipeline network, Submitted to Comp. and Operational Research.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

MODEL FOR THE FIRST-STAGE OF PYGAS UPGRADING: EXPERIMENTAL PROCEDURE AND PARAMETER ESTIMATION

Raquel M. Cavalcante,[a] Rafaella M. B. de Faria, [a] Ofélia de Q. F. Araújo, [a] José Luiz de Medeiros [a]

[a]*Departament of Chemical Engineering, Escola de Química da Universidade Federal do Rio de Janeiro. Sala E-201, Bloco E, Centro de Tecnologia, Ilha do Fundão, CEP: 21949-900 Rio de Janeiro, Brazil*

Abstract

Pyrolysis gasoline (PYGAS) is a by-product of the steam cracking of naphtha with high octane number and high aromatic contents (C5 to C12 range). PYGAS is unstable due to the presence of unsaturated compounds and must be hydrotreated before processing. Two-stage heterogeneous catalytic hydrogenation is commonly used to stabilize PYGAS. The First Stage uses a mild trickle-bed conversion for removing very reactive species (styrene and dienes), prior to the more severe Second Stage where sulfured and remaining olefins are converted in gas phase. This work investigates the First Stage of PYGAS hydrotreatment processes with commercial Pd/Al_2O_3 catalysts. The starting point is a data bank of conversion values for hydrogenating runs of synthetic feeds prepared with the representative reactive species in the First Stage scenario. These data were gathered from experimental runs in a short scale unit at pressures near 30 bar, temperatures between 50 and 70ºC, and weight hourly spatial velocity (*WHSV*) from $2h^{-1}$ to $70h^{-1}$. The data was used for parameter estimation of a kinetic model proposed to the First Stage of PYGAS plants. The obtained model can be employed to develop a computational tool useful in the design and optimization of PYGAS plants.

Keywords: pyrolysis gasoline, hydrotreatment process, parameter estimation

1. Introduction

Naphta steam cracking is the most used process for production of high-value petrochemical olefins. The main liquid byproduct of this process is known as pyrolysis gasoline or PYGAS. PYGAS is rich in benzene and monoaromatics (Table 1), but it also contains very reactive unsaturated species (e.g. styrene and dienes), which demand conversion by hydrogenation prior to the recovery of valuable monoaromatics (de Medeiros et al., 2007). This hydrogenation is normally designed in two stages. In the First Stage, the selective and complete hydrogenation of dienes and styrene is carried out under mild temperature and pressure, using supported Pd catalysts. The partial hydrogenation of mono-olefins and negligible conversion of sulfured species in this stage, demand the subsequent more severe Second Stage where complete vapor-phase destruction of remaining unsaturated and sulfured species takes place. The critical selectiveness issues in the First Stage, demand good kinetic models for designing PYGAS plants. Public studies on kinetic modeling on the First Stage of PYGAS upgrading are rare. Cheng et al. (1986) presented characterization data of typical PYGAS and kinetic data for the hydrogenation of isoprene and styrene. These results

were expressed as simple kinetic models using irreversible first order laws in terms of the reactants dissolved in liquid-phase. A previous study by de Medeiros et al. (2007) presented kinetic data from an experimental procedure reproducing the conditions of the First Stage of PYGAS plants. This study consolidated an experimental data bank of hydrogenating runs of synthetic feeds prepared with a set of reactant and inert model compounds pertinent to the First Stage scenario. These authors also presented a two-phase kinetic model for the First Stage of PYGAS hydrogenation.

Table 1. Typical Weight Composition of PYGAS

Component	Weight Composition (%)
Paraffins + Naphtenics	12
Olefins/Diolefins/Styrene	6 / 18 / 3
Benzene/Toluene/ Xilenes/ C_9+Aromatics	28 / 14 / 7 / 12

In the present work, we employ the experimental data bank and kinetic model of de Medeiros et al (2007) in order to estimate kinetic and adsorption parameters for the First Stage of PYGAS hydrogenation at 60ºC and 70ºC.

2. Experimental Data Base

All available experimental information belongs to the data bank consolidated by de Medeiros et al. (2007) built with 120 hydrogenation runs of synthetic feeds prepared with model compounds and toluene. A continuous bench plant with an up-flow fixed bed isothermal reactor was used. This unit operates under steady conditions via continuous pumping of liquid feed and hydrogen. The reactor temperature, pressure, *H_2/Feed* ratio and *WHSV* were controlled. All detailed techniques used in feed preparation, chemical analysis and process operation are described by de Medeiros et al. (2007). Feed compositions of reactant species (*1,7-octadiene, DCPD, styrene, 1-octene*) were chosen near real PYGAS values. Each feed gives rise to a group of runs with excursions of temperature (*T*) between 35ºC and 100ºC, and *WHSV* between 2 and 70 h^{-1} accordingly to two-dimensional factorial designs (*square-star* design) with 2 replicas in the central point (Montgomery, 2001). Reactor pressure (*P*) was set at 30 bar. *H_2/Feed* ratios were set between 24 Nl/kg and 137 Nl/kg, preferentially near 132 Nl/kg. All runs used a commercial Pd/Al_2O_3 catalyst (0.3wt% Pd). Experiment responses are measured stationary values of weight % of toluene (*TOL*), styrene (*ST*), ethylbenzene (*EB*), 1-octene (*R*=), 1,7-octadiene (*R*==), n-octane (*R*), 1,6 dimethyl-cyclohexane (*CR*), DCPD (*C*==) and adamantane (*C*).

3. Reaction Network and Kinetic Model

Fig. 1 defines the reaction network by the set of irreversible reactions (numbered by (.)) and kinetic rules (numbered by [.]). Reaction rates (mol/s.kgCat) are modeled via a four steps mechanism: (1)equilibrium dissociation adsorption of H_2 [$H_2(g)+2\sigma(s) \leftrightarrow 2H\sigma(s)$]; (2)equilibrium adsorption of hydrocarbon species [$HC(l)+\sigma(s) \leftrightarrow HC\sigma(s)$]; (3)slow hydrogenation of adsorbed species [$HC\sigma(s)+H\sigma(s) \leftrightarrow HC\sigma\text{-}H\sigma(s)$]; (4)fast subsequent hydrogenation [$HC\sigma\text{-}H\sigma(s) + nH\sigma(s) \leftrightarrow HCH\sigma(s) + n\sigma(s)$]. If f_i represents the fugacity of species i, $K_k(T)$ the kinetic constant of k^{th} reaction, and $K_j^{AD}(T)$ the Langmuir adsorption factor for species *j*, the rate expression for k^{th} reaction can be written as (de Medeiros et al., 2007):

$$R_k(T,\underline{f}) = K_k(T) * \Psi(\underline{f}) * K_j^{AD} * f_j \qquad (1)$$

$$\Psi(T,\underline{f}) = \frac{\sqrt{K_{H_2}^{AD} * f_{H_2}}}{\left(1+\sqrt{K_{H_2}^{AD} * f_{H_2}} + \sum_{j\neq H_2}\left(K_j^{AD} * f_j\right)\right)^2} \quad (2)$$

Defining:

$S_{kj} = 1 \Rightarrow k^{th}$ reaction rate is defined by species j, otherwise $S_{kj} = 0$; and

$D_{ki} = 1 \Rightarrow k^{th}$ reaction rate uses i^{th} kinetic rule, otherwise $D_{ki} = 0$;

the vector of reaction rates for all network can be written according to:

$$\underline{R}(T,\underline{f}) = \Psi(\underline{f})\left(\underline{\underline{Diag}}(\underline{\underline{D}}\underline{K}(T))\right)\underline{\underline{S}}\left(\underline{K}^{AD} \bullet \underline{f}\right) \quad (3)$$

Kinetic constants and Langmuir adsorption factors are expressed as function of temperature in Arrhenius representation, as shown below:

$$\underline{K}(T) = \underline{k}_0 \bullet \exp(-\underline{E}_A / T) \quad (4a)$$

$$\underline{K}^{AD}(T) = \underline{k}_0^{AD} \bullet exp(-\underline{E}_A^{AD} / T) \quad (4b)$$

where $\underline{E}_A$, $\underline{E}_A^{AD}$, $\underline{k}_0$ and $\underline{k}_0^{AD}$ express vectors of activation energies (K), reference kinetic constants (mol/s.kg$_{Cat}$) and reference Langmuir factors (bar^{-1}). Symbol • stands for multiplication of corresponding components of vectors. Spatial reactor time is defined – Eq. 5 – from reactor section A (m^2), axial position z (m), bed density ρ^{cat} (kg/m^3), and feed rate F_0 (kg/s). Stationary component material balances are written as in Eq. 6, where $\underline{N}$ and $\underline{\underline{H}}$ express, respectively, the vector of component molar rates and the stoichiometric matrix. Reactor (or experiment) responses are produced by integrating Eq. 6 from *t=0* to *t=3600/WHSV* as in Eq. 7, where $\underline{N}_0$ is the vector of component molar rates at the reactor inlet, defined by F_0, the feed composition and the *H_2/Feed* ratio.

$$t = z * A * \rho^{CAT} / F_0 \quad (5)$$

$$\frac{d}{dt}\underline{N} = F_0\Psi(\underline{f})\underline{\underline{H}}\,\underline{\underline{Diag}}(\underline{\underline{D}}\underline{K}(T))\underline{\underline{S}}(\underline{K}^{AD} \bullet \underline{f}) \quad (6)$$

$$\underline{N} = \underline{N}_0 + \int_{t=0}^{t=3600/WHSV} F_0\Psi(\underline{f})\underline{\underline{H}}\left(\underline{\underline{Diag}}(\underline{\underline{D}}\underline{K}(T))\right)\underline{\underline{S}}(\underline{K}^{AD} \bullet \underline{f})\,dt \quad (7)$$

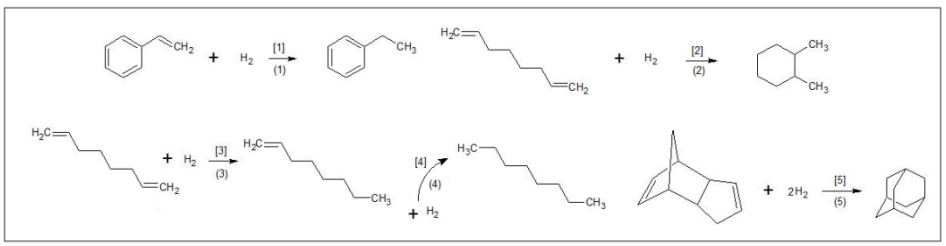


Figure 1: Network of Irreversible Reactions for PYGAS Hydrogenation

Eq. 6 is numerically integrated via adaptive methods suitable to stiff problems. Along the integration path, successive vapor-liquid equilibrium problems with specified $T,P,\underline{N}$ – $flash(T,P,\underline{N})$ – are solved via Newton-Raphson method. Thermodynamic properties are calculated for both phases by Soave-Redlich-Kwong equation of state with classical mixing rules. Critical constants are estimated by Joback Method.

4. Parameter Estimation

The procedure follows closely the proposition given in de Medeiros et al. (2007). At a given temperature, the vector of *np* parameters ($\underline{\theta}$), is estimated with the corresponding isothermal slice of database runs. The estimate of $\underline{\theta}$ is $\underline{\hat{\theta}}$, containing *nk=5* kinetic constants and *nc=10* component Langmuir factors ($\underline{\hat{\theta}}^T = [\,\underline{\hat{K}}^T\ \underline{\hat{K}}^{AD^T}\,]$). Each run *i* produced a *ny x 1* vector of experiment responses with *ny=9*. The prediction of model responses for this run is $\underline{\hat{Y}}_i(\underline{\hat{\theta}})$. If $\underline{\hat{Y}}_i(\underline{\hat{\theta}}) - \underline{Y}_i^{EXP}$ is the vector of response residue for run *i*, and neglecting constant terms, the minimum of the negative of the sum of logarithmic Likelihood Functions for runs is given by Eq. 8, where *m* is the number of runs considered and $\underline{\underline{W}}_i$ is a *ny x ny* diagonal weighting matrix for run *i*:

$$\underset{\{\underline{\theta}\}}{Min}\left(\frac{1}{2}\right)\sum_{i}^{m}(\underline{\hat{Y}}_i(\underline{\hat{\theta}})-\underline{Y}_i^{EXP})^T\,\underline{\underline{W}}_i\,(\underline{\hat{Y}}_i(\underline{\hat{\theta}})-\underline{Y}_i^{EXP}) \qquad (8)$$

This problem was solved numerically via a Nelder-Mead Simplex Method, after a suitable change of variables in order to impose bounded search domains.

5. Statistical Analysis

Statistical entities are used for evaluation of the estimation process. Statistics S_R^2 is an estimator for the intrinsic variance (Eq. 9). The variance-covariance matrix of estimated parameters and correct parameter confidence domain at level (1-α)*100% (α=0.01) are given in Eq. 10, where $\underline{\underline{J}}_i = (\nabla_{\underline{\theta}}\underline{\hat{Y}}_i^{\,T})^T$ is the jacobian matrix of *i* responses with parameters and $\phi_{1-\alpha}$ is the Fisher abscissa with (*np,ny*m-np*) degrees of freedom.

$$S_R^2 = \left(\frac{1}{ny*m-np}\right)\sum_{i}^{m}(\underline{\hat{Y}}_i-\underline{Y}_i^{EXP})^T\,\underline{\underline{W}}_i\,(\underline{\hat{Y}}_i-\underline{Y}_i^{EXP}) \qquad (9)$$

$$\underline{\underline{C\hat{O}V}}(\underline{\hat{\theta}}) = S_R^2 * \left(\sum_{i=1}^{m}\underline{\underline{J}}_i^{\,T}\,\underline{\underline{W}}_i\,\underline{\underline{J}}_i\right)^{-1} \quad , \quad (\underline{\theta}-\underline{\hat{\theta}})^T\left(\sum_{i=1}^{m}\underline{\underline{J}}_i^{\,T}\,\underline{\underline{W}}_i\,\underline{\underline{J}}_i\right)(\underline{\theta}-\underline{\hat{\theta}}) \le np * S_R^2 * \phi_{1-\alpha} \qquad (10)$$

6. Results

Kinetic constants were estimated for data at temperatures of 60ºC and 70ºC. At 60ºC, all parameters $\underline{\hat{\theta}}$ were estimated due to the high number of observed values (59 runs). On the other hand, at 70ºC (19 runs) only the five kinetic constants and the Langmuir factor of H_2 were estimated, whereas the adsorption factors of remaining species were kept at the 60ºC values. Table 2 shows all estimated parameters and respective Arrhenius coefficients (Eqs. 4). Figs. 2A and 2B show the distributions of observed and estimated *wt% styrene* at 60ºC and 70ºC. Fig. 3 shows a 2D projection of the 99% confidence

region for kinetic parameters of styrene (K_1) and 1,7-octadiene (K_2) conversions at 60ºC. Figs. 4 show plots of observed versus predicted responses for 60ºC and 70ºC.

Table 2. Summary of Estimation Results

Adsorption Factors (bar^{-1})			Kinetic Const.($mol/s.kg_{Cat}$)			Arrhenius Coefficients	
	@60ºC	@ 70ºC		@60ºC	@ 70ºC	E_A (K)	k_0
K^{AD}_{ST}	0.0180	---	K_1	1.0396	1.0697	326.01	2.7672
K^{AD}_{EB}	0.0190	---	K_2	0.0576	0.1174	8128.4	$2.2986.10^{9}$
K^{AD}_{TOL}	0.0118	---	K_3	0.1438	0.3291	9455.2	$3.0844.10^{11}$
$K^{AD}_{R=}$	0.0010	---	K_4	0.0223	0.0576	10872	$3.357.10^{12}$
$K^{AD}_{R==}$	0.3461	---	K_5	5.6986	6.5874	1655.5	821.89
K^{AD}_{R}	0.7509	---				---	---
K^{AD}_{CR}	0.1394	---	S^2_R	27.0	23.2	---	---
$K^{AD}_{C==}$	0.0366	---				---	---
K^{AD}_{C}	0.3984	---				---	---
K^{AD}_{H2}	2.7388	0.9191				-12471.6	$1.4868.10^{-16}$

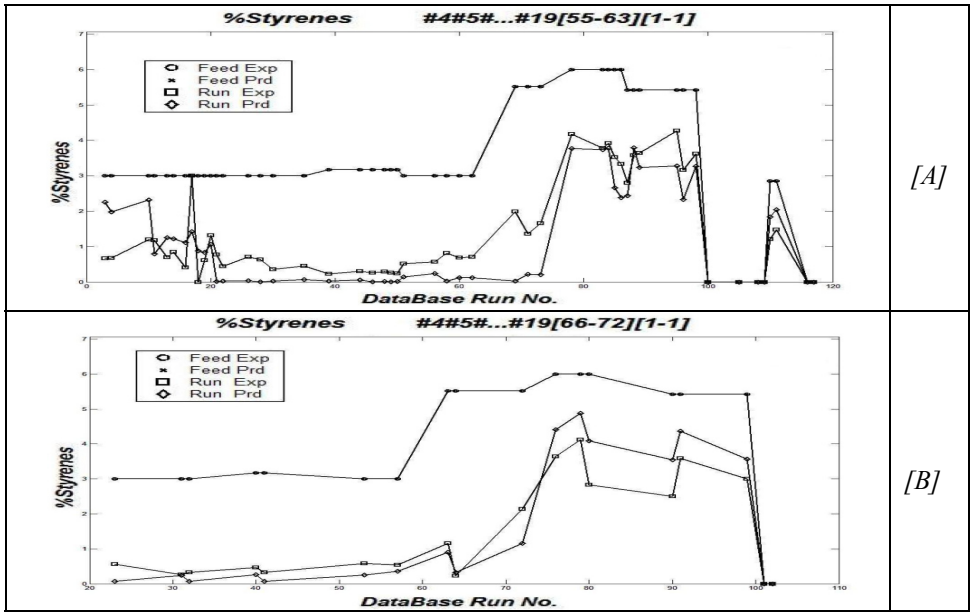


Figure 2: Observed Versus Predicted wt% Styrene *[A]: T=60ºC, [B]: T=70ºC*

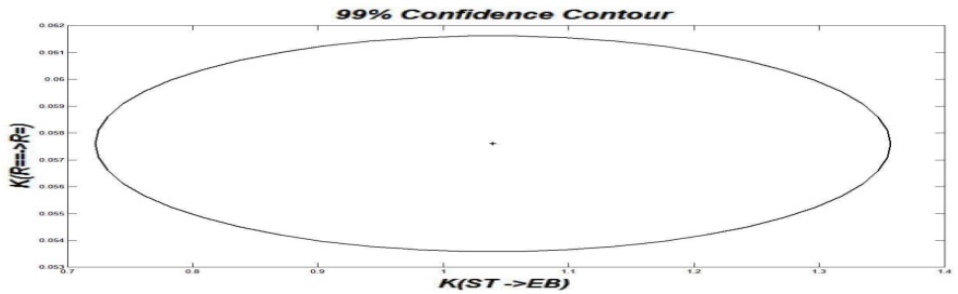


Figure 3: 99% Confidence Region – K_1 versus K_2 at *T=60ºC*

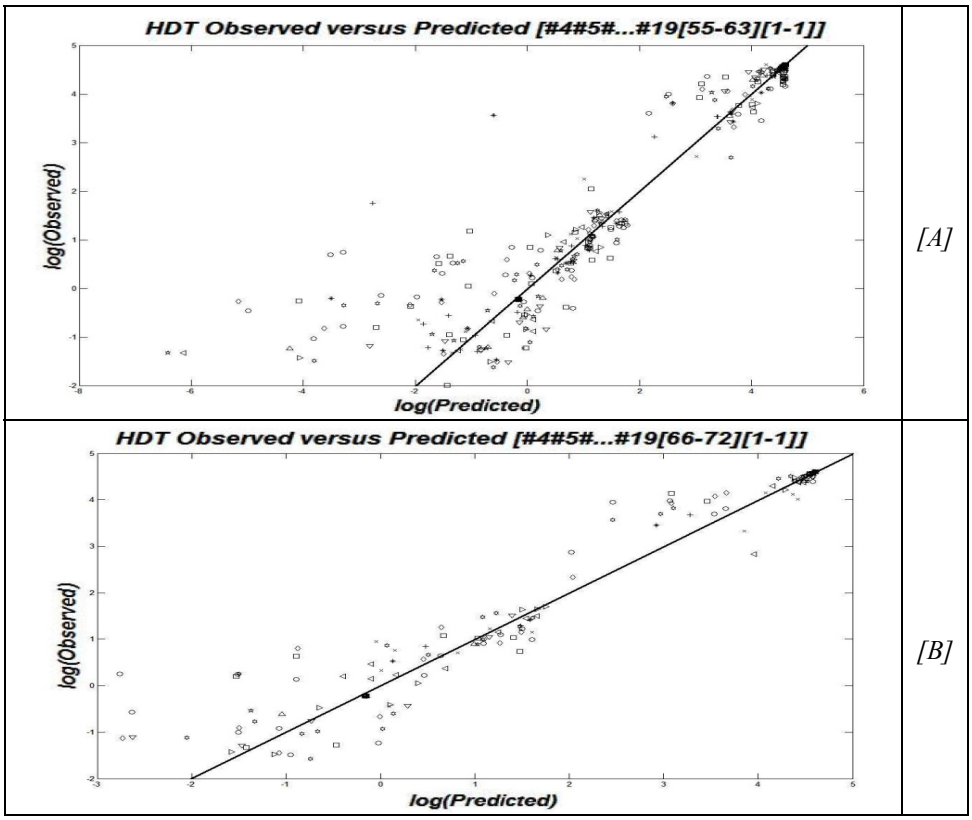


Figure 5: Logarithmic Observed Versus Predicted Responses *[A]:T=60ºC; [B]:T=70ºC*

7. Conclusions

The reactor model of de Medeiros et al. (2007) for the First Stage of PYGAS plants was approached for parameter estimation over 120 hydrogenation experimental runs with synthetic feeds made with model compounds. The model takes into account the two-phase (equilibrated) flow along the reactor and kinetic rates expressed with adsorption effects and component fugacities. Model parameters were estimated at 60 and 70ºC via statistical processing of experimental data. Arrhenius coefficients were generated for these estimations whenever possible. These forms are useful for developing an engineering model of the (adiabatic) industrial reactor for PYGAS upgrading. The adsorption constants were only estimated for 60 ºC, due to the difficulty to estimate them with few experimental data (19 runs at 70 ºC).

References

Y.M. Cheng, J.R. Chang, J.C. Wu,1986, Kinetic Study of Pyrolysis Gasoline Hydrogenation over Supported Palladium Catalyst, Appl. Cat., No. 24, 273

J.L. de Medeiros., O. Q. F. Araújo, A. B. Gaspar, M. A. P. Silva, J. M. Britto, 2007, A Kinetic Model For The First Stage Of PYGAS Upgrading, Braz. J. of Chem. Eng., 24, No. 1, 119-133.

D.C. Montgomery, 2001, Design and Analysis of Experiments, 5th Edition, John Wiley, USA

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling Study on Emulsion Copolymerization of Methyl Methacrylate and Butadiene Systems

Éder D. de Oliveira,[a] Esleide L. Casella,[b] Odair Araujo,[c] Reinaldo Giudici,[d]

[a]*Department of Chemical Engineering, School of Engineering, Federal University of Minas Gerais, Rua Espírito Santo, 35, Belo Horizonte, 30160-030, Brazil*
[b]*Department of Materials Science Engineering, Mackenzie Presbyterian University, Rua da Consolação, 930, São Paulo, 01302-907, Brazil*
[c]*CIBA Specialty Chemicals, 704, Rue Pierre et Marie Curie, Ribécourt, 60772, France*
[d]*LSCP/CESQ - Department of Chemical Engineering, Polytechnic School, University of São Paulo, Av. Prof. Luciano Gualberto, 380, trav. 3, São Paulo, 05508-900, Brazil*

Abstract

A general mathematical model for 1,3-butadiene/methylmethacrylate emulsion copolymerization process in batch reactor is presented. The model is able to explain the effects of simultaneous changes in emulsifier concentration, initiator concentration, and monomer feed composition on conversion, average particle size, and pressure evolution histories. The main features of the system, such as the increase of the rate of polymerization as emulsifier and initiator concentrations increase and the dependence of the reaction on the number of particles, are correctly represented by the model. The model accounts for the basic features of the process and may be useful for practical applications, despite its simplicity and reduced number of adjustable parameters.

Keywords: emulsion copolymerization, mathematical modeling, methyl methacrylate, 1,3-butadiene.

1. Introduction

Emulsion polymerization is an important industrial process for manufacturing commercial products such as rubbers, paints, and adhesives. It is a very complex heterogeneous process involving several chemical and physical phenomena, many of which have not yet been completely elucidated. Because of its industrial relevance, it is important to develop mathematical models that can represent the mechanisms involved in these processes. This is an essential task to better understand a given system, reflecting the knowledge on the interactions between all the factors that affect it. As opposed to many vinyl copolymer systems, the emulsion polymerization of methyl methacrylate (M) and 1,3-butadiene (B) has received very little attention in the literature. To the authors' knowledge, only the works by Shapiro *et al*. (1981) and Araujo *et al*. (2001) have focused on experimental investigations on this system. Similarly, modeling of this process seems to have received attention only by Saldívar *et al*. (2002) so far. In this work, a general mathematical model has been developed and validated by comparison with the experimental data of Araujo *et al*. (2001), aiming at getting an insight on the mechanisms governing emulsion copolymerization processes of this copolymerization system. The influence of several process variables, such as initiator concentration, emulsifier concentration, and feed composition has been examined.

2. Theory

The mathematical model used in this work is based on the approaches of Casella *et al*. (2003), Zubitur *et al*. (2004), and Pinelli Filho *et al*. (2006). Following Casella *et al*. (2003), the main idea is to provide and test a general common model with a reduced number of adjustable parameters applicable to many copolymerization systems. The main model equations are detailed in the next sections.

2.1. Monomer Balances

The emulsion copolymerization process begins with the thermal decomposition of an initiator in the aqueous phase. The initiator decomposition is represented by a first-order reaction (Casella *et al*., 2003). Polymer chains start to grow from the reaction of these radicals with the monomers. The conversion of monomer i (i = 1 or 2), x_i, is calculated from the monomer balance, written as (Casella *et al*., 2003):

$$M_{io}\frac{dx_i}{dt} = R_{pi} \tag{1}$$

where M_{io} is the initial ratio of monomer i to volume of water, and R_{pi} is the molar rate of polymerization of monomer i, accounting for the polymerization in both the aqueous and polymer particle phases. Assuming a pseudo-homopolymerization approach and terminal model (Gilbert, 1995; Casella *et al*., 2003), it is calculated by

$$R_{pi} = \left(k_{p1i}P_1^p + k_{p2i}P_2^p\right)[M_i]_p\frac{N_T\bar{n}}{N_A} + \left(k_{p1i}P_1^{aq} + k_{p2i}P_2^{aq}\right)[M_i]_{aq}[R\cdot]\phi_w^{-1} \tag{2}$$

where k_{pij} is the propagation rate constant between radical type j and monomer type i, $[M_i]_p$ is the concentration of monomer in the particles, $[M_i]_{aq}$ is the concentration of monomer in aqueous phase, $[R^-]$ is the concentration of hydrophilic sulfated radicals in the aqueous phase, N_T is the total number of particles per volume of water in the reactor, N_A is Avogadro's number, ϕ_w is the volume fraction of water in the aqueous phase, and $\bar{n}$ is the average number of radicals per particle. The probability of finding an active chain with ultimate unit (1 or 2) in the polymer particles (p) or aqueous phase (aq), P_i^k (k = p or aq), can be calculated from k_{pij} and the concentration of monomer in the respective phase (Casella *et al*., 2003).

2.2. Concentration of Free Radicals in the Aqueous Phase

For the balance of free radicals in the aqueous phase, it was considered that the radicals containing hydrophilic sulfate groups, $[R^-]$, with size larger than a certain length, z, can be captured by micelles or particles or can be consumed by propagation and termination reactions. However, the radicals with size ranging from 1 to z – 1 cannot be absorbed by micelles or particles (Gilbert, 1995). They can grow until the reach of a critical size, jcr, from which they precipitate out from the aqueous phase and form new particles by homogeneous nucleation. For homopolymers, jcr and z can be estimated from the correlations proposed by Gilbert (1995). For copolymer radicals, following a suggestion from Dubé *et al*. (1997), jcr and z were obtained from the average of the values for the homopolymers, weighted by the instantaneous composition of the copolymer. The radical balance for the radicals produced in the aqueous phase can be represented by equations similar to those described by Araújo and Giudici (2003). Radicals can also be present in aqueous phase due to desorption from particles. The generation of these is usually attributed to chain transfer reactions to monomers inside the particles (Dubé *et al*., 1997; Asua, 2003). They have a distinct chemical nature from the previous radicals, being much less hydrophilic. Therefore, they can readily be reabsorbed by the particles. Their concentration, $[M^-]$, can be obtained by a balance similar to that proposed by Casella *et al*. (2003) for the free radicals in the aqueous phase.

2.3. Number of Polymer Particles

The total number of particles is calculated considering that they are formed from homogeneous and micellar nucleation (rates: R_{hom} and R_{mic}, respectively).

$$\frac{dN_T}{dt} = R_{mic} + R_{hom} \quad (3)$$

$$R_{mic} = k_{am}\left\{\left[\frac{\alpha_2 - \alpha_2^{jcr-z+1}}{1-\alpha_2}\right](\alpha_1)^{(z-2)}\alpha_3^{-1}[R\cdot] + [M\cdot]\right\}N_m \equiv k_{am}\sigma_1 N_m \quad (4)$$

$$R_{hom} = (k_p[M]_{aq})(\alpha_2)^{(jcr-z)}(\alpha_1)^{(z-2)}\alpha_3^{-1}[R\cdot]N_A/\phi_w \quad (5)$$

$$\alpha_3 \equiv \left[(1-\alpha_1^{z-1})/(1-\alpha_1) + (\alpha_2 - \alpha_2^{jcr-z+1})/(1-\alpha_2)\right](\alpha_1)^{(z-2)} \quad (6)$$

where k_{am} is the rate constant of radical capture by micelles, ($k_p[M]_{aq}$) is the average propagation rate coefficient (Casella *et al.*, 2003), α_1 and α_2 represent the probability that radicals of sizes smaller than z and from z or higher to propagate in the aqueous phase, respectively (Casella *et al.*, 2003; Araújo and Giudici, 2003). N_m is the concentration of micelles and can be determined from the emulsifier balance.

2.4. Concentration of Free Radicals in the Polymer Particles

The average number of radicals per particle, $\bar{n}$, depends on various phenomena such as absorption and desorption of radicals and termination within the particles (Casella *et al.*, 2003; Asua, 2003) and is calculated from the following differential equation:

$$\frac{d\bar{n}}{dt} = k_{ap}\sigma_1 - k_{des}\bar{n} - \frac{2(2k_{ap}\sigma_1 + k_{des})k_{tp}}{(2k_{ap}\sigma_1 + k_{des} + k_{tp}/v_p N_A)v_p N_A}\bar{n}^2 \quad (7)$$

where k_{ap} is the rate constant of radical capture by particles, σ_1 is defined in Equation (4), k_{des} is the desorption constant, k_{tp} is the termination rate constant in the particles, and v_p is the volume of a polymer particle. Values of k_{tp} were determined according to Casella *et al*. (2003) while those of k_{des} were calculated using the simplified version of Asua's model (Asua, 2003).

2.5. Concentration of Monomers and Other Components

Monomer concentration in the dense phases (aqueous, monomer droplets, and polymer particles) were evaluated by assuming phase equilibrium at each instant of the copolymerization process and employing the partition coefficient method as described by Pinelli Filho *et al*. (2006). From the method it is also possible to determine the volume of a polymer particle (v_p).

Following Pinelli Filho *et al*. (2006), the number of mols of water and less volatile monomers in the vapor phase can be determined from the balance equation:

$$\frac{dN_i^v}{dt} = -\frac{k_L a_T}{PM_i}\left(Py_i - P_i^{sat}x_i^{aq}\gamma_i^{aq}\right) \quad (8)$$

in which N_i^v and PM_i are, respectively, the number of mols in the vapor and molecular weight of water or monomer, k_L is the mass transfer coefficient, a_T the gas-liquid interfacial area, y_i and x_i^{aq} are the mole fractions of water or the monomer in the vapor and aqueous phase, respectively. γ_i^{aq} is the activity coefficient of these compounds in the aqueous phase, P_i^{sat} is the vapor pressure of the compounds at the reaction temperature, and P is the pressure inside the reactor, which was calculated assuming ideal gas behavior (Pinelli Filho *et al*., 2006). Adapting the approach of Zubitur *et al*. (2004) to the more volatile monomer (1,3-butadiene), Equation (8) becomes

$$\frac{dN_i^v}{dt} = -\frac{k_L a_T}{PM_i}\left(Py_i - H_i x_i^{aq}\right) \quad (9)$$

where H_i is the Henry's law constant for 1,3-butadiene in water at the reactor pressure and temperature.

3. Methods

Details of the experimental procedure are described elsewhere (Araujo *et al*., 2001). The experiments that were simulated and compared with the model are shown in Table 1.

Table 1. 1,3-butadiene/methylmethacrylate experiments at 70ºC (Araujo *et al*., 2001).

Test	B/M (mol/mol)	B (g)	M (g)	Initiator (g)	Emulsifier (g)	Water (g)
BM-1	30/70	109.6	474.1	0.78	6.93	1717
BM-5	30/70	109.6	474.1	1.57	13.86	1717
BM-15	30/70	109.6	474.1	0.78	13.86	1717
BM-2	70/30	310.5	246.1	0.75	6.61	1637
BM-6	70/30	310.5	246.1	1.49	6.61	1637
BM-14	70/30	310.5	246.1	1.49	13.22	1637

Three parameters (k_{am}, k_{ap}, $k_L a_T$) were estimated by minimizing a weighted nonlinear least squares criterion involving monomer conversion, particle diameter, and pressure in the reactor using Marquardt and Simplex algorithms (Bevington, 1969; Rowan, 1990; Casella *et al*., 2003). The results were very similar. One set of parameter values was fitted for systems with a monomer molar ratio of 30/70 B/M and another to those displaying molar ratio of 70/30 (Table 1). The system of differential equations present in the model was solved numerically using Gear's method.

4. Results and Discussion

Comparison of experimental and model predictions for conversion, average particle size, and pressure is made in figures 1 to 3 for batch 1,3-butadiene/methylmethacrylate emulsion copolymerization reactions at 70ºC. Results for batches with less butadiene in the feed are shown on the left and those more concentrated on the right side.

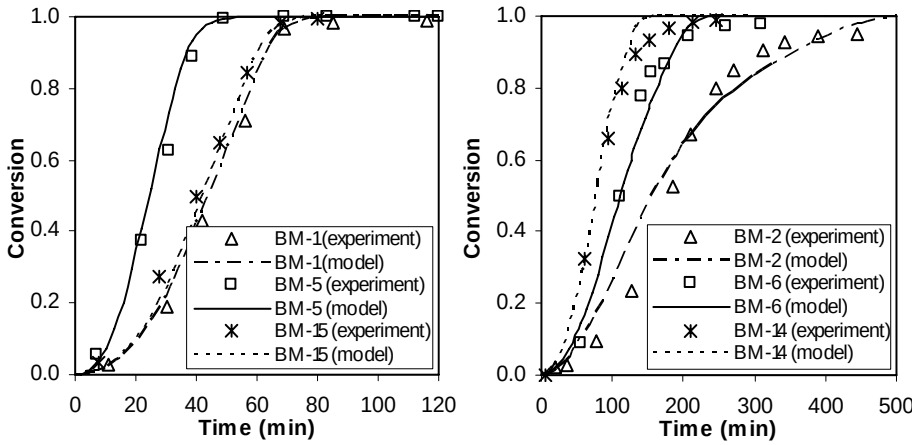


Figure 1. Evolution of conversion during B/M batch emulsion copolymerizations.

Polymerization is faster as initiator concentration increases and raises slightly by increasing the amount of emulsifier. This behavior can be related to an increase in particle nucleation that, in turn, leads to a more rapid conversion of the monomers to polymers (Gilbert, 1995). As the quantity of methyl methacrylate increases, the reaction rate also becomes faster since this monomer reacts more rapidly than butadiene (Araujo *et al*., 2001). The congruence of simulated and experimental results is quite good, being similar to that of the model proposed by Saldívar *et al*. (2002) and employing only half the number of adjustable parameters.

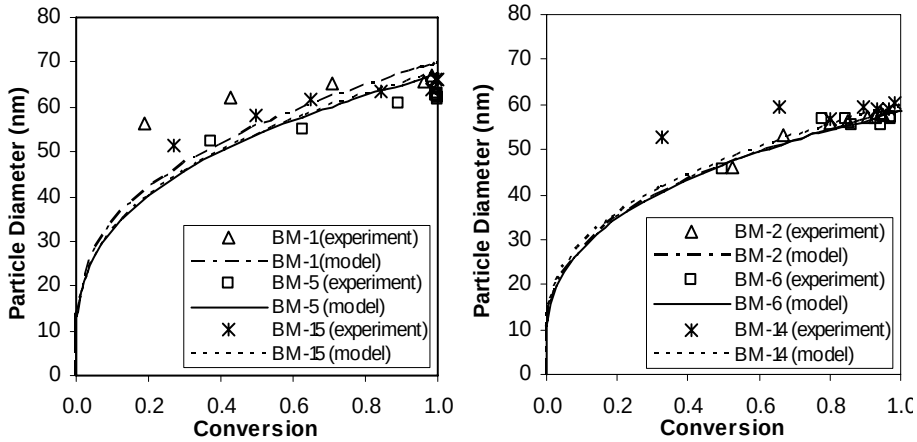


Figure 2. Evolution of particle diameter during B/M batch emulsion copolymerizations.

It is generally expected that the higher the initiator and emulsifier concentrations the faster the reaction rate and the smaller the particle size (Casella *et al*., 2003). This trend can be observed in batches with a monomer molar ratio of 30/70 B/M. However, results of particle diameter for systems richer in butadiene are barely distinguishable. The model correctly predicts the trends shown in Figure 2; however, the evolution of the particle diameter is described only in a semi-quantitative way as also reported by Saldívar *et al*. (2002). Problems in measuring this property (Casella *et al*., 2003)

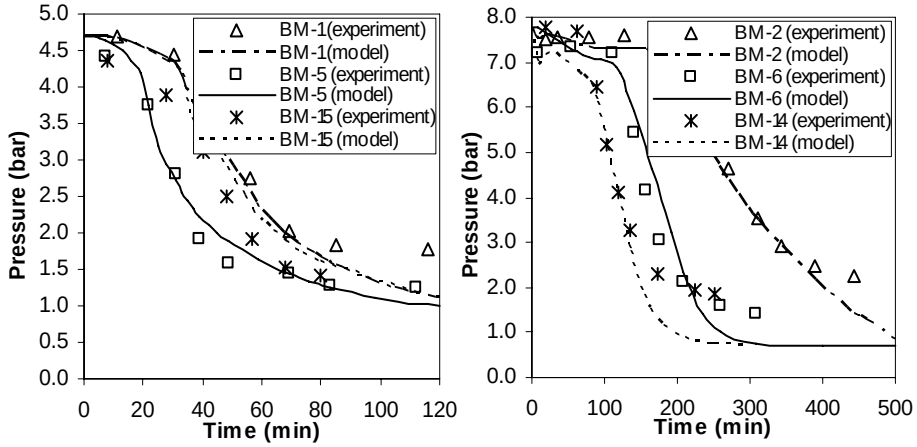


Figure 3. Evolution of reactor pressure during B/M batch emulsion copolymerizations.

and difficulties in improving the theoretical description of particle nucleation phenomena (Saldívar *et al*., 2002) contribute to this.

An additional contribution of the present model, not found in most approaches in the literature, is the prediction of the pressure-time history in the reactor. Pressure measurements are useful to identify some changes during polymerization reactions involving volatile monomers such as 1,3-butadiene. For instance, the inflection points observed in pressure evolutions (Figure 3) coincide with the instant in which the monomer droplets disappear (Pinelli Filho *et al*., 2006). Comparison of experiments and predictions for the pressure (Figure 3) indicate that the model is able to correctly describe the variation of pressure inside the reactor as the formation of polymer occurs. As there is a correlation between the variations of pressure and monomer conversion during the batch (figures 1 and 3), this points to the possibility of using the model as a state estimator for monitoring the monomer conversion from readily available, simple pressure measurements in industrial 1,3-butadiene-methylmethacrylate emulsion copolymerization processes.

Acknowledgments

The authors are grateful to Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq/BRAZIL) for supporting this work.

References

O. Araujo, R. Giudici, E. Saldívar, W. H. Ray, 2001, Modeling and Experimental Studies of Emulsion Copolymerization Systems. I. Experimental Results, *J. Appl. Polym. Sci.*, 79, 2360-2378.

P. H. H. Araújo, R. Giudici, 2003, Optimization of Semicontinuous Emulsion Polymerization Reactions by IDP Procedure with Variable Time Intervals*, Computer and Chemical Engineering*, 27, 1345-1360.

J. M. Asua, 2003, A New Model for Radical Desorption in Emulsion Polymerization, *Macromolecules*, 36, 6245-6251.

P. R. Benvington, 1969, Least Squares Fit to an Arbitrary Function, *Data Reduction and Error Analysis for the Physical Sciences*, 235-236.

E. L. Casella, O. Araujo, R. Giudici, 2003, Mathematical Modeling of Batch Emulsion Copolymerization Processes, *Polymer Reaction Engineering*, 11, 4, 869–910.

M. Dubé, J.B.P. Soares, A. Penlidis, A. E. Hamielec, 1997, Mathematical Modeling of Multicomponent Chain-Growth Polymerization in Batch, Semi-Batch, and Continuous Reactors: A Review, *Ind. Eng. Chem. Res.*, 36, 966–1015.

R. G. Gilbert, 1995, Particle Formation, *Emulsion Polymerization. A Mechanistic Approach.*, 304.

A. Pinelli Filho, O. Araujo, R. Giudici, C. Sayer, 2006, Batch and Semicontinuous Styrene-Butadiene Emulsion Copolymerization Reactions, *Macromol. Symp.*, 243, 114-122.

T. Rowan, 1990, Functional Stability Analysis of Numerical Algorithms, Ph.D. Thesis, University of Texas at Austin, Austin, Texas, USA.

E. Saldívar, O. Araujo, R. Giudici, C. Guerrero-Sánchez, 2002, Modeling and Experimental Studies of Emulsion Copolymerization Systems. III. Acrylics, *J. Appl. Polym. Sci.*, 84, 1320-1338.

Y. Y. Shapiro, N. P. Dorozova, N. M. Mironova, T. G. Balyberdina, 1981, Influence of an Emulsifier on the Microstructure of Co-Polymers of Butadiene with Methylmethacrylate, *J. Polym. Sci. USSR*, 23, 6, 1522-1529.

M. Zubitur, S. Ben Amor, C. Bauer, B. Amram, M. Agnely, J. R. Leiza, J. M. Asua, 2004, Multimonomer Emulsion Copolymerization in Presence of Inhibitors, *Chem. Eng. J.*, 98, 183–198.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Dynamic modeling and simulation of dividing-wall distillation column systems

Maria A. Vargas,[a] Gerit Niggemann,[a] Georg Fieg[a]

[a] *Institute of Process and Plant Engineering, Hamburg University of Technology Schwarzenbergstr. 95c, Hamburg 21073, Germany*

Abstract

Possible configurations of two dividing wall column systems are investigated for the separation of a five component feed mixture. Hence, a rigorous mathematical model is develop, which allows the stationary and dynamic simulation of the process. Steady state optimization is performed for the design of the column system and the selection of the configuration with the minimum total annualized costs. The selected configuration for our study case is subjected to numerous steady state sensitivity analyses. Furthermore, dynamic simulations in open loop mode are performed to analyze the transient behavior of the column system.

Keyword: dividing-wall column, distillation, column system, optimization

1. Introduction

Distillation remains the most commonly used separation process in the chemical industry. The separation of a homogenous multicomponent mixture into more than two products is usually realized in sequences of conventional distillation columns. This sequential separation does not only require high capital cost but also high amounts of energy. Complex column arrangements, such as dividing-wall columns, have been proposed as alternative to significantly reduce distillation costs. Theoretical studies (e.g., *Petlyuk et al. 1965* and *Kaibel 1988*) have revealed that in the separation of ternary mixtures, dividing-wall columns are capable of achieving 30% of energy savings.

However, in the industrial practice it is usually necessary to separate mixtures into more than three pure fractions. In that respect, dividing-wall column systems arise as an alternative. *Shah et al. 2002*, determined that these systems can reduce the energy costs in more than 20 % compared with traditional arrangements for the separation of five components mixtures. Operability issues and lack of knowledge that arise from a more complex dynamic behavior have discouraged wider industrial acceptance of these systems.

The separation of a five components mixture A,B,C,D,E (A is the lightest and E the heaviest boiling component) in a dividing-wall column system can be accomplished in three different configurations (Figure 1). The connection between the two columns could be either the bottom, the side or the distillate stream. The best sequence depends on the characteristics of the feed mixture. The complexity in the dynamic behavior of this system lies in the variable conditions in flow, temperature and composition of the connecting stream between the two columns in case of disturbances in the first column.

The purpose of this work is to perform a detailed analysis of dividing-wall column systems for the separation of five components feed mixtures. In earlier studies a rigorous process model for one dividing-wall column has been developed (Niggemann

et al., 2006). The mathematical model was subjected to an extensive validation with experimental data of our pilot plant. Based on this model, a new rigorous model for the simulation of a two dividing-wall columns system is developed for a theoretical examination. In this paper, simulation studies include steady state sensitivity analysis and analysis of the transient behavior of the system subjected to feed disturbances are conducted.

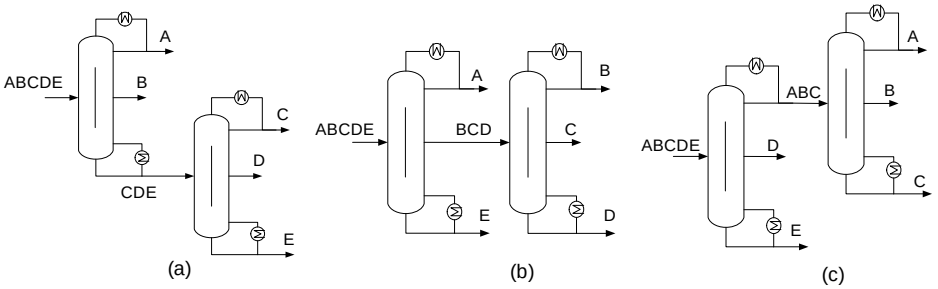


Figure 1. Configurations of dividing-wall distillation column systems for the separation of a five components mixture (a) bottom-feed BF, (b) side-feed SF and (c) distillate-feed DF.

2. Modeling

In previous studies, a rigorous mathematical model was developed for the detailed analysis of dividing-wall columns (Niggemann *et al.*, 2006). The model consists of ordinary differential equations (Eq. 1) obtained from mass and energy balances around each theoretical stage and a set of algebraic equations (Eq. 2) which are used to predict the physical properties, the vapour-liquid equilibrium and the column hydrodynamic.

$$\frac{d\dot{y}}{dt} = f(\dot{y}, y, u_d, u_o) \tag{1}$$

$$0 = g(\dot{y}, y, u_d, u_o) \tag{2}$$

The variable y denotes a vector of process state variables; u_d and u_o are the design and operating parameters. The highly nonlinear differential algebraic equation system is implemented in the commercial software tool Aspen Custom Modeler^{TM}. Important characteristics of dividing-wall columns, as the self adjusting vapour split, are considered in the model.

An extensive validation with experimental data of our pilot plant demonstrated the capability of the model to predict the steady state and transient behavior for the separation of a ternary mixture of linear alcohols.

Based on this model, a new model for the simulation of two dividing-wall columns system is developed. In our case study, a five component feed mixture, consisting of linear fatty alcohols (n-hexanol C6, n-octanol C8, n-decanol C10, n-dodecanol C12 and n-tetradecanol C-14) is separated into pure products with concentrations higher than 0,991 kg/kg. The components in the feed mixture have equal mass concentration and a feed flow rate of 5000 kg/h is used. The two dividing-wall column system configuration with the lowest total annualized costs (TAC) is selected to perform the task. Hence, the design and operating parameters of the three different arrangements DF, SF and BF (Figure 1) were optimized to minimize the objective function value TAC.

$$\min_{u_d, u_o} TAC \tag{3}$$

$$TAC = \frac{k \cdot \sum_i E_i}{n} + \sum_i U_i \tag{4}$$

The first part of Eq. 4 indicates the annualized investment cost and the second part the production costs. The investment costs are determined as the summation of column cost and periphery equipments such as condenser and reboiler. The factor *k* considers the installation costs and *n* is the plant life time. The operation variables that affect the operating costs are the reboiler and condenser duty, which values are used to determined the utility costs (*U*).

The steady state optimization problem is constrained by the mathematical process model (Eq. 1-2), the product purities (Eq. 5), liquid load (Eq. 6), the optimization bonds of the design parameters (Eq. 7) and the operating control variables (Eq. 8).

$$x_i \geq x_{i,\min} \tag{5}$$

$$B \geq B_{\min} \tag{6}$$

$$u_{d,\min} \leq u_d \leq u_{d,\max} \tag{7}$$

$$u_{o,\min} \leq u_o \leq u_{o,\max} \tag{8}$$

In this work, the set of operating variables includes the reboiler heat duty (QR) and the liquid split (Ls). The design variables include the number of theoretical stages in all segments of the dividing wall column. The feed and side streams are always located in the middle of the column. The diameter of the columns are determined to satisfied the hydraulic constrains of the column packing (F-factor ≤ 2,5 $Pa^{0,5}$). For the solution of the optimization problem was used the deterministic solution method SRQP (Successive Reduced Quadratic Programming).

The configuration BF resulted to have the lowest TAC. Figure 2 depicts the optimized operating and design parameters. The first column of the arrangement operates with a reflux ratio of 1,8 and heat duty in the reboiler of 698 kW. The column has 26 theoretical stages and diameter of 1,32 m. The second column operates with a reflux ratio of 2 and heat duty in reboiler of 402 kW. The column has 34 theoretical stages and diameter of 1,13 m.

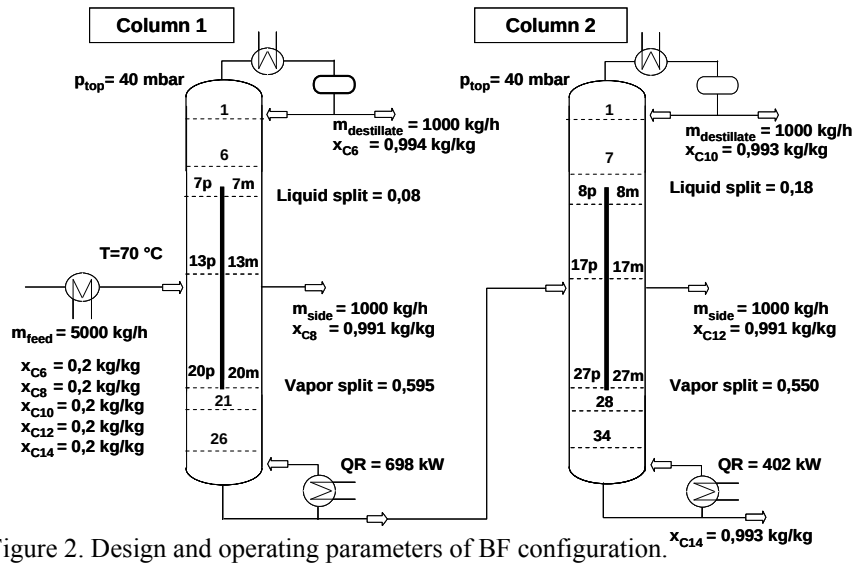


Figure 2. Design and operating parameters of BF configuration.

3. Steady state simulation

The dividing-wall column system BF selected for our study case was subjected to numerous sensitivity analyses. The high complexity of this system lies in the variable conditions in flow, temperature and composition of the connecting stream between the two columns, in case of disturbances in the first column. These disturbances can be produced by variations of process parameters in column 1 during the operation. Among these parameters are the liquid split (Ls), heat duty (QR) and the product streams flow rate. In the sensitivity analysis one operating parameter is changed while the rest parameters are kept constant.

Increments of the heat duty in column 1 result in an increase in the concentration of the key components (Figure 3). The connecting stream between the two columns experiences a reduction of impurities. These impurities, which are composed by C8, contaminate the distillate stream of column 2. In this case the maximum concentration of impurities in the feed stream of column 2 must be lower than 0,3 % w/w, otherwise the desired purity of top product in column 2 cannot be achieved. On the other hand, reduction of the heat duty in column 1, does not affect the concentration of key components in the side and bottom stream of column 2. In column 2, only the concentration of the top product is affected, due to the higher amount of C8 in the connecting stream.

Another important parameter considered in the sensitivity analysis is the liquid split. Its variation implies also changes in the vapor split (Vs), which is determined in the model by self adjusting to equalize the pressure drop in both sides of the dividing wall column. Figures 4 (a) shows that increasing Ls in column 1, produces the reduction of the key components concentration in column 1 and only affects the concentration of the key component of the top product in column 2. The explanation is similar to the case of heat duty variation. Unexpectedly, the vapor split in column 2 shows a slight increase (Figure 4 (b)). The higher amount of C8 in column 2 reduces the density of the vapor in the prefractionator, resulting in an increase of Vs. The small change does not affect the concentration of the key components in side and bottom of column 2.

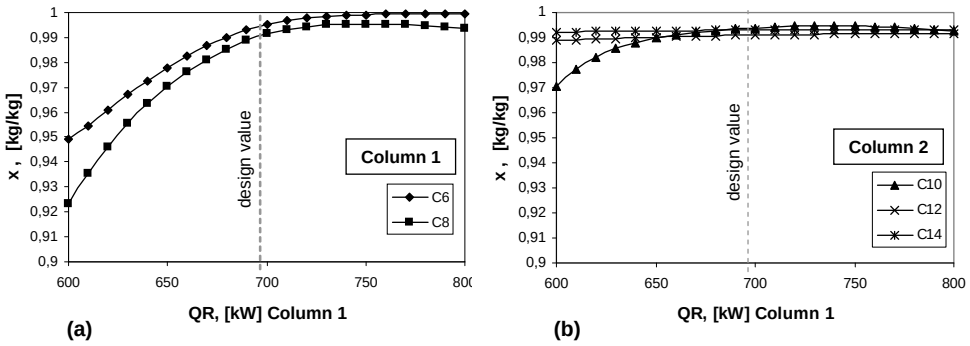


Figure 3. Variation of key components concentration in dividing-wall (a) column 1 and (b) column 2, as function of reboiler heat duty in column 1.

From the previous analysis it can be concluded that changes in the Ls and QR of column 1, can increase the concentration of C8 in the connecting stream. This produces the increment of impurities in the top stream of column 2. The concentrations of key components of side and bottom stream are not affected by changes in Ls and QR of column 1.

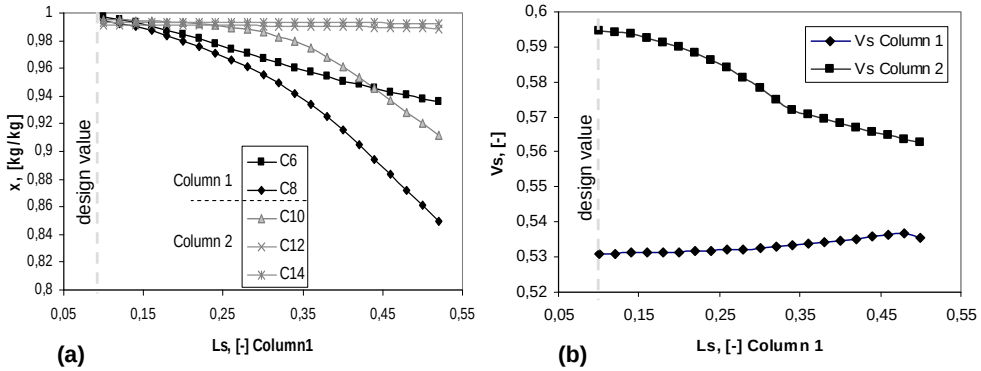


Figure 4. Variation of (a) key components concentration and (b) vapor split in dividing-wall column 1 and 2, as function of liquid split in column 1.

The performed sensitivity analyses provide important information about the interaction between the two dividing-wall columns in the system configuration BF. They quantitatively depict the variation of usual controlled variables such as temperature and concentration of key components as function of changes in process variables. These results facilitate the selection of appropriate variables for the control of the dividing-wall column system.

4. Dynamic simulation

Theoretical examinations, of the transient behaviour of a dividing-wall column system are one part of the integral analysis of this contribution. The developed mathematical model is tested and verified for dynamic simulations. The dividing-wall column system BF is inspected. Hence, perturbations in the composition of components in the feed of ± 3% are assigned. The transient behaviour of the system is analysed in open loop mode.

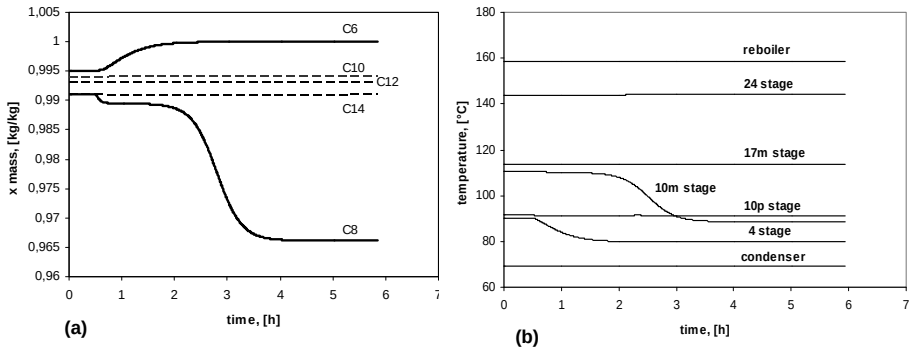


Figure 5. Transient behavior of the BF configuration in open loop mode with feed concentration disturbance, x_{C6} = + 3 % and x_{C8} = – 3 %, (a) concentration of key components and (b) temperature profiles in column 1.

In the first case, Figure 5, the concentration of C6 is incremented in 3 %, while the concentration of C8 is reduced in 3 %. The perturbation is introduced at t = 0,5 h. First, the temperature in the segment above the dividing-wall (4 stage) of column 1 decreases

and reaches a new steady state at t = 2 h. At the same time the temperature in the dividing-wall segment above the side stream withdrawal (10s stage) starts to decrease. It coincides with the high reduction in concentration of C8 in side stream column 1 (Figure 5 (a)). The temperature profile of the column 2 is not affected by the disturbances, which can be seen in the invariability of the key components concentration in column 2 (Figure 5 (a)). After around t = 4 h the system reaches a new steady state.

In the second case analyzed, the concentration of C6 is reduced in 3 % and the concentration of C14 is incremented in 3 %. Figure 6 depicts the transient behavior of the temperatures in column 1 and column 2. In both columns the disturbances increases temperature, which are stabilized again at t = 3,5 h.

The performed dynamic simulations give a first insight of the transient behavior in open loop mode of a two dividing –wall column system.

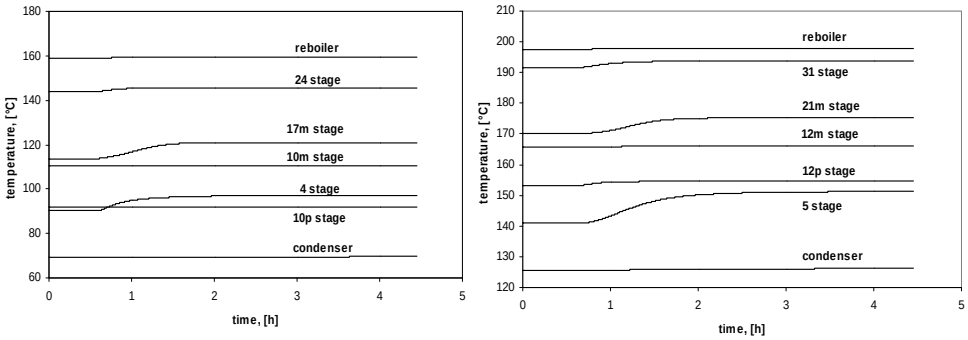


Figure 6. Transient behavior of the BF configuration in open loop mode with feed concentration disturbance, x_{C6} = - 3 % and x_{C14} = + 3 %, (a) temperature profile column 1 and (b) temperature profile column 2.

5. Conclusion

A rigorous mathematical model, developed for steady state and dynamic simulation of a two dividing-wall column system have been tested and verified. The results showed the robustness and stability of the numerical solution. In our study case, the design and operating parameters of the possible column configurations were optimized to minimize the TAC. The configuration with the minimum TAC (BF configuration) was investigated. Steady state sensitivity analyses revealed the influence of changes in process variables in the purity of the products. Furthermore, a first insight of the transient behavior in open loop mode of a two dividing–wall column system presented. The results of this contribution serve as a promising basis for the development of control concepts in dividing-wall column systems.

References

Petlyuk, F.B., V.M. Platonov and D.M. Slavinskii (1965). Thermodynamically Optimal Method for Separating Multicomponent Mixtures. Int. Chem. Eng., 5, 555 - 561.

G. Kaibel, (1988), Distillation Column Arrangement with Low Energy Consumption, IChemE Symposium Series, 109, 43-59

P. B. Shah and A. C. Kokossis, (2002). New Synthesis Framework for the Optimization of Complex Distillation System, AIChE Journal, 48, 3, 527 – 550

Niggemann, G., S. Gruetzmann and G. Fieg (2006). Distillation startup of fully thermally coupled distillation columns: theoretical examinations. Symposium Series, 152, 800 - 808.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Singularity theory based stability analysis of reacting systems

WANG Hangzhou, CHEN Bingzhen*, HE Xiaorong, QIU Tong, ZHAO Jinsong

Dept. of Chemical Eng. , Tsinghua University, Beijing, 100084, China
**Corresponding author. Tel.: +86 10 62784572; fax: +86 10 62770304.*
E-mail address: dcecbz@tsinghua.edu.cn (B. Chen).

Abstract

Recently, how to design reacting systems that are inherently safer is a problem of common interest. Clearly, instability limits the degree of inherent safety of a reacting system, but so far much more attention has been paid to the analysis of the multiplicity features of reacting systems while only a little works has been reported on the stability analysis of reacting systems under open-loop conditions. This paper presents a singularity theory based stability analysis method that could characterize regions in parameter space over which different kinds of stability characteristics may exist. Thus, the stability characteristics of different scenarios could be quickly identified under the open-loop condition, based on which the inherently safer operational space could be located at the conceptual design stage. A case is studied to demonstrate the effectiveness of the proposed method.
Keywords: Steady state solutions; Stability analysis; Singularity points

1. Introduction

As modern chemical plants are becoming more complex and bigger in scale, the associated chance of things going wrong is also increasing rapidly. Due to the flammable, explosive, toxic and corrosive nature of chemical process, any single accident may trigger a major catastrophe that brings tremendous environmental, social and economical loss. For insurance of running chemical process safely, how to design reacting systems that are inherently safer--with the focus on disturbances having the potential for hazardous responses--is a problem of common interest.

As we all know, chemical process are strongly non-linear systems [1, 2], usually there are multiple steady state solutions [3-8]. Much attention has been paid to these solutions. Uppal [9] studied the dynamic behavior of continuous stirred tank reactors, Balakotaiah [10], YUAN Qi-peng [11], Xu [12] using bifurcation Theoretical Analysis the multiple steady state solution of reactors, Razon [13, 14] reviewed multiplicities and instabilities in chemically reacting systems, Seider [1, 15, 16] also studied multiple steady-state solutions of reactors. Some of these solutions are instable.

Clearly, instability limits the degree of inherent safety of a reacting system, but so far much more attention has been paid to the analysis of the multiplicity features of reacting systems while only a little works has been reported on the stability analysis of reacting systems under open-loop conditions. This paper presents a singularity theory based stability analysis method that could characterize regions in parameter space over which different kinds of stability characteristics may exist. Thus, the stability characteristics of different scenarios could be quickly identified under the open-loop condition, based on

which the inherently safer operational space could be located at the conceptual design stage.

2. Algorithm

The method is described as follows. First of all, the homotopy-continuation method which is effective in achieving global convergence is applied to get the all steady state solutions of a reacting system in its parameter space. Then, the singularity theory is introduced to analyze these steady state solutions and to identify their stability characteristics. The stability characteristics of the reacting system will change along with the system's parameter variation. So it is important to identify singularity points where the change of a reacting system's stability characteristics occurs. As a result, the parameter space can be divided into several subspaces over which different characteristics in terms of stability exist.

Assume the chemical process $\frac{dX}{dt} = F(X,\lambda)$, where X is the state variables, $X \in R^n$, λ is the parameter. $F_X(X,\lambda)$ is the Jacobian matrix of $F(X,\lambda)$. X_0 is the solution of $F_X(X,\lambda)=0$ when $\lambda=\lambda_0$. The stability of system is determined by the eigenvalues of Jacobian matrix, if all eigenvalues are in left panel the system is stable, otherwise the system is instable.

The algorithm is described as follows:

1. Computing all steady state solutions of chemical systems.

Computing the solutions of $F_X(X,\lambda)=0$ under different values of λ.

2. Computing singular values of chemical systems' Jacobian matrix.

Computing the rank of matrix $F_X(X,\lambda)$ under certain λ and the related solution of X. If $r(F_X)=n-1$, the λ and related X are singularity point. The stability of system would change a lot between two sides of the singularity point. Record the singularity point (X_s,λ_s).

3. Determine the stability status at two sides of the singularity points.

Get one point at each side of singularity point in solution curve, computing all eigenvalues of $F_X(X,\lambda)$, if all the eigenvalues are at left panel, the system at the operation point is stable, otherwise the system is instable. The stability status is continuous, often changed pass through the singularity points. In this way only few points are calculated to determine the stability status.

4. Repeat the above process to calculate all the singular points and determine the stability of singular points on both sides.

There are 3 kinds of subspaces: the stable subspace, the unstable subspace and the mixed subspace. Within the stable subspace the system will run smoothly still, even if some variable disturbance is encountered, within the unstable subspace the system may change a lot or run away even under very little disturbance. This subspace is dangerous, so effort should be made at the conceptual design of a reacting system to avoid selecting process operating region in this subspace. The mixed subspace is the one within that the system may change a lot or not change at all when disturbance occurs. This situation is also should be avoided in order to guarantee the designed reacting system is an inherently safer one. Thus, the elimination of instability could be achieved through changes in the operating conditions, such as the inlet changes involving feed concentration and/or the cooling water temperatures.

3. Example

In this example, an anaerobic fermentor is designed for high product quality. A continuous culture of K. pneumoniae is grown on glycerol. The fermentation products

are ethanol, acetic acid, and 1,3-propanediol, with 1,3-propanediol being the main product.[15, 17]

$$\frac{dX}{dt} = X(\mu - D)$$

$$\frac{dC_s}{dt} = D(C_{s0} - C_s) - Xq_s$$

$$\frac{dC_p}{dt} = Xq_p - DC_p$$

Where

$$\mu = \mu_{\max}\frac{C_s}{C_s + K_s}\left(1 - \frac{C_s}{C_s^*}\right)\left(1 - \frac{C_p}{C_p^*}\right)$$

$$q_s = m_s + \frac{\mu}{Y_s^m} + \Delta q_s^m \frac{C_s}{C_s + K_s^*}$$

$$q_p = m_p + Y_p^m \mu + \Delta q_p^m \frac{C_s}{C_s + K_p^*}$$

Here X is the biomass concentration, C_s is the substrate concentration in reactor, C_{s0} is the substrate concentration in feed medium, C_p is the product concentration, D is the dilution rate.

For Convince, the dimensionless model as follows.

$$\frac{dx}{d\tau} = x(u - d)$$

$$\frac{dy}{d\tau} = d(y_o - y) - x\phi_1(\beta_1 + \gamma_1 u + \frac{y}{y + \alpha_1})$$

$$\frac{dz}{d\tau} = x\phi_2(\beta_2 + \gamma_2 u + \frac{y}{y + \alpha_2}) - zd$$

$$u = \frac{y}{y + \alpha_3}(1 - y)(1 - z)$$

Here x is the dimensionless biomass concentration, y is the dimensionless biomass concentration in reactor, z is the dimensionless product concentration, d is the dimensionless dilution rate and y_0 is the dimensionless biomass concentration in feed medium.

When $d=0.2$ $d$imensionless product concentration z various with y_0. Figure 1 shows the result, there two singularity in the system, marked as LP.

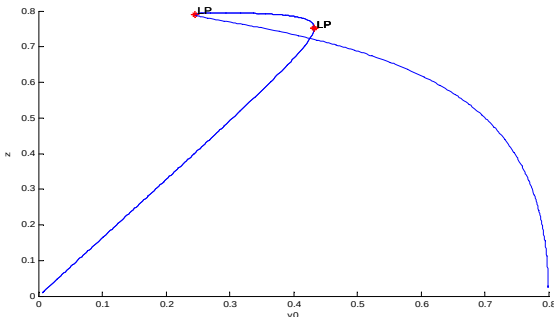


Fig 1. when d=0.2 the dimensionless product concentration varies with dimensionless biomass concentration in feed medium

Calculate the Jacobian matrix to determine the stable status. As we know there are there eigenvalues of Jacobian matrix in this reactor. Here we focus on the value of eigenvalue varies with the dimensionless biomass concentration in feed medium.

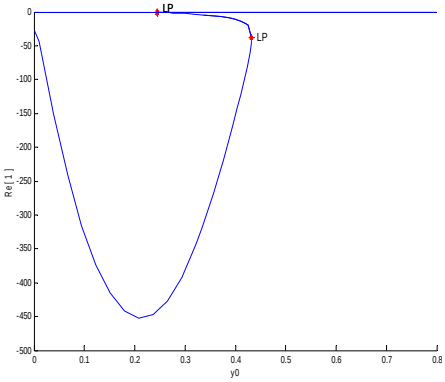


Fig 2. the first eigenvalue's real part's curve

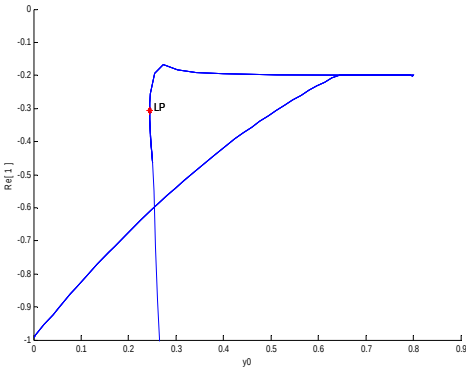


Fig 3. the first eigenvalue's real part's curve (partly)

Fig 2 and Fig 3 show that the first eigenvalue's real part is always below zero, so the stability status is determined by the other eigenvalues.

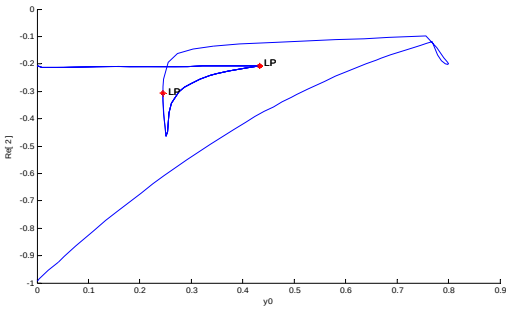


Fig 4. the second eigenvalue's real part's curve

Fig 4 shows that the second eigenvalue's real part is also always below zero.

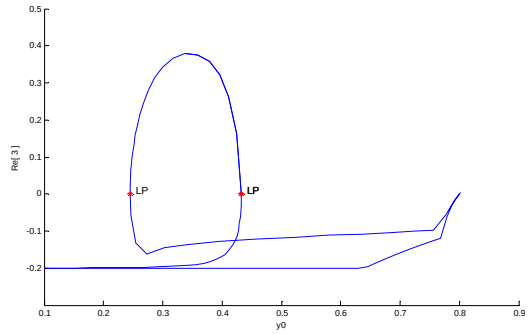


Fig 5. the third eigenvalue's real part's curve

Fig 5 shows that the third eigenvalue's real part changes pass zero just at the singularity points, where system's stability status changed a lot.

According to the character of stability, the stability of system in Fig 1 is showed as Fig 6. There are three areas, two of them are stable and one of them is unstable.

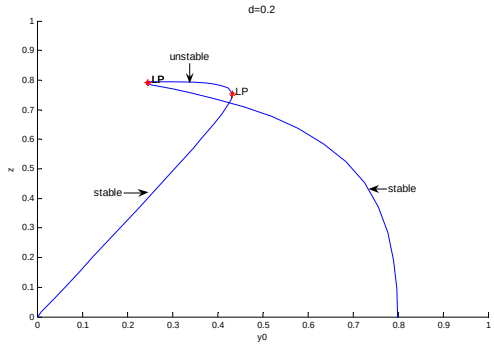


Fig 6. the different stable areas separated by singularity points

The results are showed in Fig 7-10 when d changes. Singularity points are marked as LP. The dimensionless dilution rate varies from 0.3 to 0.6, the systems stability status varies. There are three different stability statuses in each scenario. It is a convenient method that using singularity point theory to determine the system's stability status.

4. Conclusion

The proposed singularity theory based stability analysis method provides a guidance on how to make a change of the operating conditions so that to achieve an elimination of instability. Also, this method reduces a lot of works for the identification of the stability characteristics of a reacting system to compare with the existing approaches. In this paper, a case is shown to illustrate the effectiveness of the proposed method.

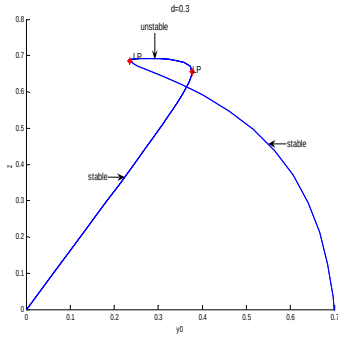


Fig 7 the stable areas when d=0.3

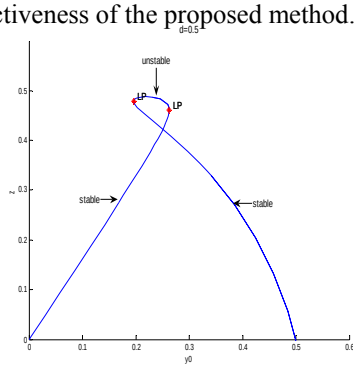


Fig 8 the stable areas when d=0.5

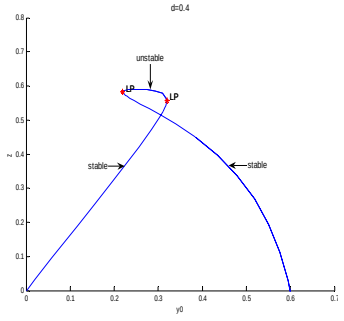


Fig 9 the stable areas when d=0.4

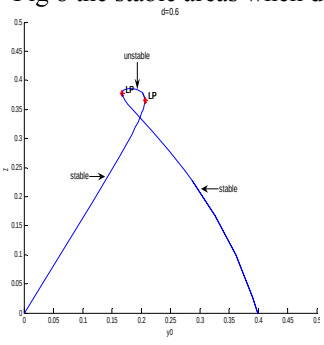


Fig 10 the stable areas when d=0.6

References

[1] W. D. Seider, D. D. Brengel, A. M. Provost, and S. Widagdo, "Nonlinear analysis in process design. Why overdesign to avoid complex nonlinearities?," Industrial & Engineering Chemistry Research, vol. 29, pp. 805-818, 1990.

[2] A. Kuhlmann and D. Bogle, "Study on nonminimum phase behaviour and optimal operation," Computers & Chemical Engineering, vol. 21, pp. S397-S402, 1997.

[3] K. S. Gritton, J. D. Seader, and W.-J. Lin, "Global homotopy continuation procedures for seeking all roots of a nonlinear equation," Computers & Chemical Engineering, vol. 25, pp. 1003-1019, 2001.

[4] F. Jalali, J. D. Seader, and S. Khaleghi, "Global solution approaches in equilibrium and stability analysis using homotopy continuation in the complex domain," Computers & Chemical Engineering, vol. 32, pp. 2333-2345, 2008.

[5] T. L. Wayburn and J. D. Seader, "Homotopy continuation methods for computer-aided process design," Computers & Chemical Engineering, vol. 11, pp. 7-25, 1987.

[6] V. N. Kiva and B. M. Alukhanova, "Multiple steady states of distillation and its realisation," Computers & Chemical Engineering, vol. 21, pp. S541-S546, 1997.

[7] C. J. Wang, D. S. H. Wong, I. L. Chien, R. F. Shih, S. J. Wang, and C. S. Tsai, "Experimental investigation of multiple steady states and parametric sensitivity in azeotropic distillation," Computers & Chemical Engineering, vol. 21, pp. S535-S540, 1997.

[8] S. S. E. H. Elnashaie and J. R. Grace, "Complexity, bifurcation and chaos in natural and man-made lumped and distributed systems," Chemical Engineering Science, vol. 62, pp. 3295-3325, 2007.

[9] A. Uppal, W. H. Ray, and A. B. Poore, "On the dynamic behavior of continuous stirred tank reactors," Chemical Engineering Science, vol. 29, pp. 967-985, 1974.

[10] V. Balakotaiah and D. Luss, "Global analysis of the multiplicity features of multi-reaction lumped-parameter systems," chemical Engineering Science, vol. 39, pp. 865-881, 1984.

[11] YU Qipeng and QIAN Zhongming, " The Study on Static Bifurcation Behavior of the Immobilized Yeast Spherical Beads for Ethanol Production," JOURNAL OF CHEMICAL ENGINEERING OF CHINESE UNIVERSITIES, vol. 17, pp. 527-533, 2003.

[12] X. Yan, M. Xinbin, and X. Genhui, "The Bifurcation Behavior of CO Coupling Reactor " CHINESE JOURNAL OF CHEMICAL ENGINEERING vol. 13, pp. 56-61, 2005.

[13] L. F. Razon and R. A. Schmitz, "Multiplicities and instabilities in chemically reacting systems: a review," chemical Engineering Science, vol. 42, pp. 1005-1047, 1987.

[14] L. F. Razon, "Stabilization of a CSTR in an Oscillatory State by Varying the Thermal Characteristics of the Reactor Vessel," International Journal of Chemical and Reactor Engineering, vol. 4, p. 1320, 2006.

[15] A. Meel and S. W. D, "Game theoretic approach to multiobjective designs: Focus on inherent safety," AIChE Journal, vol. 52, pp. 228-246, 2006.

[16] W. D. Seider and S. Widagdo, "Multiphase equilibria of reactive systems," Fluid Phase Equilibria, vol. 123, pp. 283-303, 1996.

[17] Z. L. Xiu, A. P. Zeng, and W. D. Deckwer, "Multiplicity and stability analysis of microorganisms in continuous culture: Effects of metabolic overflow and growth inhibition," Biotechnology and bioengineering, vol. 57, pp. 251-261, 1998.

[18] R. Antonelli and A. Astolfi, "Continuous stirred tank reactors: easy to stabilise?," Automatica, vol. 39, pp. 1817-1827, 2003.

[19] L. F. Razon, "Effect of reactor heat capacity on the stability and start-up time of a diabatic controlled-cycled stirred tank reactor," vol. 192, pp. 1194 - 1203, 2005.

[20] W. D. S. M. S. Anjana Meel, "Game theoretic approach to multiobjective designs: Focus on inherent safety," AIChE Journal, vol. 52, pp. 228-246, 2006.

[21] A. E. Gamboa-Torres and A. Flores-Tlacuahuac, "Effect of process modeling on the nonlinear behaviour of a CSTR Reactions A--> B--> C," Chemical Engineering Journal, vol. 77, pp. 153-164, 2000.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Novel Approach to Mechanism Recognition in Escherichia coli Fed-Batch Fermentations

M. N. Cruz Bournazou [a], H. Arellano-Garcia [a], J. C. Schöneberger [a], S. Junne [b], P. Neubauer [b], G. Woznya

[a] *Chair of Process Dynamics and Operation, Sekr. KWT-9*
Technische Universität Berlin, Str. des 17. Juni 135, D-10623 Berlin, Germany
[b] *Chair of Bioprocess Technology, Sekr. ACK24, Technische Universität Berlin, Ackerstr. 71-76, D-13355 Berlin, Germany*

Abstract

In this work, a novel systematic approach to achieve an efficient mechanistic modeling and simulation of fed-batch fermentations is presented. In order to show the efficiency of the developed simulation framework, data of *Escherichia coli* fed-batch fermentations are used. Fermentation processes are characterized by its dynamic behavior described by parameters such as growth rate, substrate concentration and cellular metabolic activity. Although there are models able to describe individual fed-batch fermentations, they become unreliable when fitted to new fermentations. To overcome this drawback, in this work different models are used at different optimal time points enabling not only a better description of the process, but also a better understanding of non measurable characteristics. By these means, three models compete in different intervals of the process. The candidate models are: an Overflow metabolism model (OF), a Citric Acid Cycle model (CAC) and a Survival or Maintenance model (M). Using an adequate model sequence, acetate formation, substrate consumption and cell growth are predicted with high accuracy. Moreover, the data needed to fit the models are reduced and a standardization of the model to be applied in different process states is enabled. Besides, with the development of a robust and effective model, the possibility of an online implementation for monitoring and control of the fermentation is exhibited. The results show that an efficient process monitoring based on the Dissolved Oxygen Tension and the Mechanistic Recognition is only limited by the convergence velocity of the algorithm.

Keywords: Mechanistic Recognition, Model Discrimination, *Escherichia coli*, Fed-Batch Fermentation.

1. Introduction

1.1. Process simulation and control in biologic systems

The importance of real-time measurement and parameters determination in bioreactor processes is rapidly increasing. Adapted regulatory guidelines encourage the biotechnological industry to apply innovative process development, manufacturing technology, and quality system approaches (fda.gov). The purpose is to accomplish a quality control based on the process instead of the product by means of Process

Analytical Technology (PAT) (Pillai 2005). This enhances the benefits of an accurate process description through simulation.
It is well known that mechanistic (rigorous or first principle) models offer a number of advantages over "black-box" modeling e.g. a higher process comprehension and a more accurate scale-up capability. By this means, rigorous models provide the basis needed for an efficient quality control based on the process. Nevertheless, the simulation of dynamic processes based on rigorous modeling is confronted with an important problem. Due to the physical foundation of the model, the choice of the mechanistic model to be used for the simulation is based on the dominant physical phenomena of the process. These phenomena, which dictate the process behavior, change over the process run. Hence, the appropriate approach would be to simulate the process with various models changing over time based on how and when these phenomena change. Most of these processes can be simulated effectively for short time periods but not for the complete process.

1.2. Mechanistic recognition in biologic systems

Every cell is adapted to different conditions which give them an advantage on the competition against other organisms. Factors like temperature, pH, salinity, dissolved oxygen (for aerobes), substrate availability and coexistence with other organisms are major aspects for bacteria survival. Models proposing to explain biological systems depend on many parameters, which have to be determined each time the system passes slight variations. E. *coli* fermentations are a significant example. The preculture conditions have an enormous influence on the behavior of devolution of cultivation parameters. This leads to models that are usually over-parameterized and have to be fitted with heuristic approaches. Instead, a number of simpler models, which can only predict specific stages of the process, can be applied. By this means it is possible to reduce the number of parameters, increase model flexibility and model validity with less experimental information. Further development could present important support to the theory of Nonlinear State Observer (Soroush 1997). A reliable mechanistic recognition will also allow process monitoring of non-measurable variables through the determination of the prevailing physical phenomenon. The simplicity of the models applied, the robustness and a better evaluation of the changing states of the process through are potential contributions of the mechanism recognition to the control of dynamic systems..
Many indirect methods have been studied with respect to the determination of cell activity and product concentration in fermentation processes. Some examples are dissolved oxygen tension (dot) (Whiffin 2004), on-line high performance liquid chromatography (*on-line* HPLC) (Turner 1994) and *at-line* and *in-situ* near-infrared spectroscopy (NIR) (Arnold 2002). The proposed mechanistic recognition has an advantage over them, because it evaluates the run of the curve instead of the data points separately. In other words, it enables the accumulation of more information without increasing the recollected data.

1.3. Principle of mechanistic recognition

The mechanistic recognition is an effort to overcome the problems confronted when modeling dynamic systems. In a previous work, a first attempt towards mechanism identification was proposed for a heterogeneous catalytic reaction (Schöneberger, Arellano-Garcia et al. 2006). A similar concept was published by Keneth D. Forbus (Forbus 1984). Forbus proposes the Qualitative Process theory (QP), a model discrimination based on qualitative information. A disadvantage of this method is the implementation of quantitative analysis. Most of the processes can not be described

with such simple concepts as positive-negative slope or open-closed valve. Neither semi-quantitative programing can be applied in most cases. A common problem is that quantitative variables or changing parameters can not be measured neither directly nor indirectly with the precision needed, e.g. catalyst fouling or bacterial stress.
In this work, the proposed approach substitutes the qualitative concept of QP theory with a quantitative model discrimination as used in the Optimal Experiment Design (OED) (Box and Hill 1967) in order to select the appropriate model. OED is a statistical method used to reduce the costs of experimental work needed so as to select the model which best describes the investigated topic. An analysis of the correlations and variances of the model parameters determines how many experiments have to be realized, and which inputs and initial conditions will provide the most information (Bauer, Bock et al. 2000).

2. Simulation of E. *coli* Fed-Batch Fermentations

2.1. Application

In this work, the mechanistic recognition approach is applied to the model developed by Lin *et al* based on a dynamic maximal uptake rate of glucose and oxygen. This model successfully described the process run of ten fed-batch fermentations of *E. coli* by considering that both, the glucose and oxygen uptake, depend on the growth rate. Nevertheless, the model contains an excess of parameters and is very difficult to fit without using heuristic rules as well as an important quantity of experimental data.
The model prediction relies on the coordination of three sub-models (reduced models): an overflow metabolism (OF) at substrate excess where the intermediate acetate is accumulated and the process is limited by the oxygen uptake; a citric cycle metabolism (CAC) where the process is limited by the substrate uptake; and a survival stage (M), at which the cell employs all the substrate and acetate available to survive. A sensitivity study verifies that the model is highly unstable (data not shown). Minor changes in the parameters hinder the ODE-Solvers to converge. The reason for this instability is that the intervals, where each submodel is active: Overflow, Citric Acid Cycle and Survival, is drastically affected by changes in the parameters. This produces discrete changes during the numerical integration. When these changes exceed a certain value, the ODE-Solver is not able to continue with the integration.
This example is perfectly suited for the application of the mechanistic recognition. First, the general model is divided into three simplified ones, based on the sub-models obtained from Lin *et al*. Then a model discrimination is performed in each interval throughout the process. Finally, it is determined which model fits which interval with the highest accuracy and is applied.

2.2. Material and Methods

All the information regarding the experimental procedure can be obtained from (Lin, Mathiszik *et al*. 2001). The original model written in SciLab was implemented into Matlab 7.7.0. The routine of SNOPT with the help of the TomLab interface was applied to estimate the parameters with a sequential optimization.
The parameters considered for the model fit are:

K_A:	Saturation constant for acetate [g/L]
K_La:	volumetric oxygen transfer coefficient [1/h]
K_s q:	Saturation constant for substrate uptake [g/L]
q_Esd, q_Eod:	specific death rates for the uptake enzymes of substrate and oxygen respectively [g/(gL)]

q_Smax:	Maximal specific substrate uptake rate [g/(gh)]
q_Omax:	Maximal specific oxygen uptake rate [g/(gh)]
Y_EOX,Y_ESX:	Yield of the enzyme for substrate uptake and respiration from the biomass growth, respectively [-]
Y_XA:	Yield of the enzyme for acetate from the biomass growth [-]
Y_XSof Y_XSox:	Yield of the enzyme for overflow and oxidative energy production from the biomass growth, respectively [-]

The solution strategy to generate the model sequence is briefly described in fig.1.

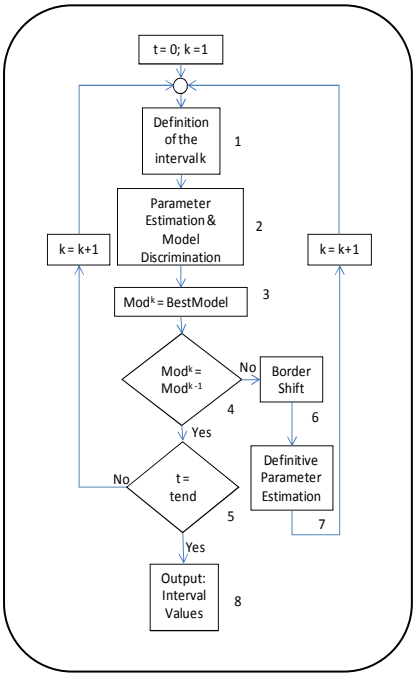


Figure 1: Flow chart of the algorithm for mechanism recognition

1. Definition of the interval k (between bounds k-1 and k). The length of the interval depends on the differentiability and identifiability of the process as well as the models considered (Franceschini and Macchietto 2007). All the mechanistic models, which are proposed for the process, are taken into account and analyzed.
2. A model discrimination is carried out. The parameters of each model are estimated and the models are compared in order to detect the model that best fits the process in the present interval.
3. The model selected by the model discrimination as the "best" model is applied to the interval.
4. If the "best" model of the interval k corresponds to the "best" model of the interval k-1 $Mod^{k-1}=Mod^{k}$, then the bound is considered to be an "invalid" bound and the algorithmproceeds to step 5. Invalid bounds are bounds located between intervals where the "best" model is the same. This means no change on the process behaviour is to be seen. Otherwise, if $Mod^{k-1}\neq Mod^{k}$, the bound is considered to be a representative bound (RB) and its exact location is searched in step 6.
5. If the interval k is the last interval of the process and the final point t = tend is reached, the program proceeds to step 8. Otherwise k is set to k+1 and the loop is restarted.
6. The representative bound (RB) is shifted step by step throughout the time interval between the bounds k-1 and k. The model k-1 is applied between k-1 and RB. The best model k is applied between RB and k. The Maximum Likelihood (MXL) is calculated in every new location of the RB. Finally the position of RB which shows the smaller MXL is set as the final location of the bound and though as the location where the model-switch takes place.
7. A final parameter estimation of the active model, between the new representative bounds and the last representative bound registered, is calculated to determine the definitive value of the parameters.

8. The program terminates the loop and prints a report with the representative bounds, the active model in each interval and the corresponding parameters estimated in each interval.

An example should illustrate the concept of the mechanistic recognition. In this example, a fictitious experimental data is simulated. Three equations are used during the process at different intervals. Each equation is considered to be a mechanistic model describing a certain phenomenon in the process under study.

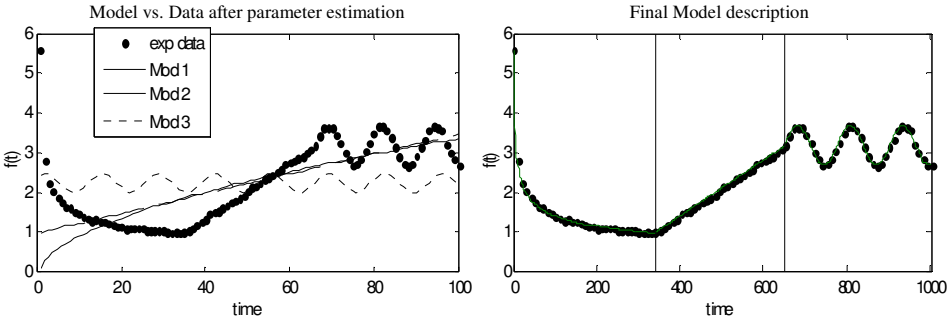


Figure2: *left:* Simulated "Sample Data" and prediction of the three models after a parameter estimation. *Right:* the mechanistic recognition enables a better prediction of the process as well as a deeper understanding.

The responses of the three different models (equations 1, 2 and 3) after parameter estimation are compared in fig.2 left). As it can be seen, none of the models is able to describe the process. A conventional model discrimination would discard all models and no solution for the simulation of this process would be found. This result is actually correct, though none of these three models is able to describe the complete process.
Nevertheless, without any further change on the models, a smart combination of them is able to reproduce the experimental data perfectly, as it can be seen in fig.2 right). In addition the time periods, where each phenomenon occurs, are detected.

3. Results

The following model sequence was detected by the mechanistic recognition: OF, CAC, OF, M fig.3.

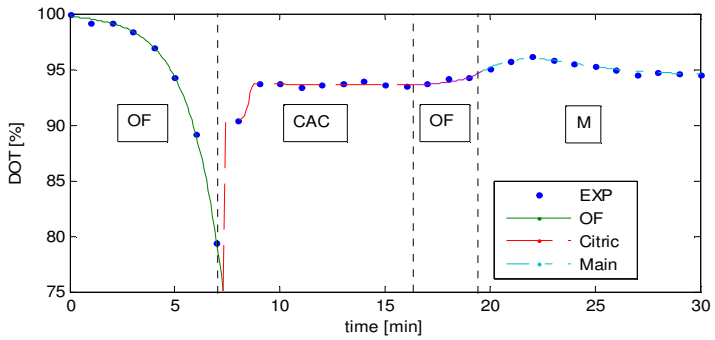


Figure 3: Validation of the data simulated with mechanistic recognition. The experimental results of the dissolved oxygen tension (dot) and the simulation results obtained with mechanistic recognition are compared.

The result was compared with the original model. It shows an acceptable similarity in the outputs. The first advantage is that the model sequence allows fitting to different data sets automatically and a reduction on the number of parameters to be fitted simultaneously. Second, an indicator of the current cell state is obtained. The mechanistic recognition is able to report whether the metabolism is dominated by overflow (acetate production), conversion of metabolits in the citric acid cycle or by maintenance. Finally, the new model sequence can be fitted with less experimental data. One measured variable, the dot, leads to a satisfying model fit. This suggests the possibility to adopt the mechanistic recognition technique to monitor and control the fed batch fermentation of *E. coli* based on the model developed by Lin *et al*.

4. Conclusions

The proposed approach to mechanistic recognition proved to be able to transform the model originally proposed by Lin *et al* 2001 into one suitable for parameter estimation and *on-line* applications. This was achieved by dividing the original model in three simpler ones and applying them in the precise process interval. The evidence provided in this project suggests a robust and efficient model sequence which can be applied in different fermentations and can also be effectively fitted.

A sophistication of the model through the evaluation of the "model health", based on the analysis of the parameter changes and its position with respect to the parameter bounds, should allow further information recollection as predicting a change in the bacteria state with anticipation, and speed up the interval shift procedure. In order to enable the application of the mechanistic recognition as an *on-line* tool, a compromise between conversion velocity, robustness and accuracy must be evaluated.

References

Arnold, A. G., R. Harvey, L. M. McNeil, B. (2002). "Use of at-line and in-situ near-infrared spectroscopy to monitor biomass in an industrial fed-batch Escherichia coli process " Biotechnology and Bioengineering 80(4): 405-413.

Bauer, I., H. G. Bock, et al. (2000). "Numerical methods for optimum experimental design in DAE systems." Journal of Computational and Applied Mathematics 120(1-2): 1-25.

Box, G. E. P. and W. J. Hill (1967). "Discrimination among mechanistic models." Technometrics 9(1): 57-71.

fda.gov http://www.fda.gov/Cder/OPS/PAT.htm#Introduction.

Forbus, K. D. (1984). "Qualitative Process Theory." Artificial Intelligence 24(1-3): 85-168.

Franceschini, G. and S. Macchietto (2007). "Model-based design of experiments for parameter precision: State of the art." Chemical Engineering Science.

Lin, H. Y., B. Mathiszik, et al. (2001). "Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited fed-batch cultivations of Escherichia *coli*." Biotechnol Bioeng 73(5): 347-57.

Pillai, V. W., M. (2005). "The impact of PAT on information architecture and infrastructure." Process Analytical Technology 2(6):16-19.

Schöneberger, J., H. Arellano-Garcia, et al. (2007). "A Novel Approach to Automated Mechanism Recognition based on Model Discrimination." Chemical Engineering 12, 471-477.

Soroush, M. (1997). "Nonlinear state-observer design with application to reactors." Chemical Engineering Science **52**(3): 387-404.

Turner, C. G., M. E. Thornhill, N. F. (1994). "Closed-loop control of fed-batch cultures of recombinant Escherichia *coli* using on-line HPLC." Biotechnology and Bioengineering 44(7): 819-829

Whiffin, V. S. C., M. J. Cord-Ruwisch, R. (2004). "On-line detection of feed demand in high cell density cultures of Escherichia *coli* by measurement of changes in dissolved oxygen transients in complex media " Biotechnology and Bioengineering 85(4): 422-433

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling of biomass gasification applied to a combined gasifier-combustor unit: Equilibrium and kinetic approaches

Rodolfo Rodrigues,[a,b] Argimiro R. Secchi,[a,c] Nilson R. Marcílio,[b] Marcelo Godinho[b]

[a]*GIMSCOP - Group of Integration, Modeling, Simulation, Control, and Optimization of Processes;* [b]*LPR - Laboratory of Residues Processing*
Department of Chemical Engineering, Federal University of Rio Grande do Sul. Rua Eng. Luis Englert, s/n - Campus Central. CEP 90040-040. Porto Alegre (RS), Brazil. E-mail: rodolfo@enq.ufrgs.br, nilson@enq.ufrgs.br, godinho@enq.ufrgs.br
[c]*COPPE/PEQ/LMSCP - UFRJ, Av. Horácio Macedo, 2030 - Centro de Tecnologia. CEP 21941-914. Rio de Janeiro (RJ), Brazil. E-mail: arge@peq.coppe.ufrj.br*

Abstract

This study proposes an analysis of biomass gasification using two mathematical models: one equilibrium model and one kinetic model based on the level of details and input data. After validation with data from literature, the models were applied for analysis of a combined gasifier-combustor unit for processing of solid wastes (biomass) of footwear industries. The gas product predicted by the models was indirectly validated with experimental data. For this, the adiabatic flame temperature estimated from the gas product was confronted to experimental temperature of the combustor unit, showing the accuracy of each model. Sensitivity analyses of the models were carried out regarding to the feed air flow rate. The responses to parametric changes were observed in major output parameters: low heating value (LHV) of gas product, and cold gas efficiency, which are important measures to quantify the performance of the system.

Keywords: biomass gasification, combined gasifier-combustor unit, equilibrium model, kinetic model.

1. Introduction

An alternative destination to solid wastes in general is the thermal treatment technologies. By this is possible to reduce filled volume and to stabilize dangerous components. One of these technologies is the gasification which generates a useful gas fraction under moderate conditions. According to Higman and van der Burgt [1], the gasification is a thermochemical process of conversion of any carbonaceous fuel to gaseous product with a useable heating value. As a result it can be to use like fuel gas or syngas to after using.

This study proposes an analysis of biomass gasification through mathematical modeling. However it is well-known that practical data (geometric measures, identification of intermediary compounds, etc.) are oftentimes available on high or low detailed levels to satisfactory process simulation. In this sense, two mathematical models are presented: one equilibrium model and one kinetic model based on the level of details and input data. After validation with data from literature, the models were applied for analysis of a pilot combined gasifier-combustor unit for processing of solid

wastes (biomass) of footwear industries. As a renewable resource, the processing of that abundant biomass is enclosed inside the natural carbon cycle unlike fossil-derived resources such as oil and coal.

2. Modeling of the process

Biomass gasification models are predominantly separated in two groups: equilibrium approach and kinetic approach [2]. Kinetic models take into account the chemical kinetics of the main reactions and the transfer phenomena among the phases, estimating the composition of each species on any point of space and time of a system. These models are specific in general for each process, providing important considerations related to chemical mechanisms and ways to increase reaction rates and process performance. On the other hand, equilibrium models provide the greatest possible conversion of each species regardless the system size and the time needed to reach equilibrium. These models do not require details of system geometry neither estimate the necessary time to reach that equilibrium.

2.1. Equilibrium model

Among the equilibrium modeling, two approaches can be cited: stoichiometric and non-stoichiometric [2]. Although equivalent in essence, the stoichiometric approach applies the equilibrium constants from related chemical reactions [3-4], while the non-stoichiometric approach minimizes the Gibbs free energy subject to mass balance and non-negative constraints [2, 5].

A stoichiometric model was developed in this work, assuming as major hypotheses: all carbon of biomass is gasified, making char formation negligible; the system is adiabatic and isobaric; and all oxygen is consumed on the process. In Table 1, the main equations of the equilibrium model are depicted.

Table 1. Main equations of the equilibrium model.

Overall mass balance:	Equilibrium constant of reaction *j*:
$F_{in}\sum_{in} n_{k,i}x_i = F_{out}\sum_{out} n_{k,i}x_i$ (1)	$K_j = \prod_i (P_i/P_o)^{\nu_{i,j}}$ (3)
where $n_{k,i}$ is the number of atoms k of a molecule i, and x_i is the molar fraction of a component i.	which is related to temperature by:
Overall energy balance:	$-RT\ln K_j = \Delta G_j^o$ (4)
$F_{in}\sum_{in} x_i H_i(T_{in}, P) = F_{out}\sum_{out} x_i H_i(T_{out}, P)$ (2)	where ΔG_j^o is the variation of standard Gibbs free energy of reaction j as function of temperature.

In this case, two equilibrium reactions are considered: water-gas shift reaction (5) and metanation (6):

$$CO + H_2O \leftrightarrow CO_2 + H_2 \quad (5)$$

$$C + 2H_2 \leftrightarrow CH_4 \quad (6)$$

By means of the energy balance is possible to estimate the final temperature for a specific kind of biomass and fraction of air stream. As a result, the flue gas composition is estimated as function of temperature.

2.2. Kinetic model

The important features of the kinetic model are the chemical kinetic approach and transfer phenomena relations. In this sense, in literature, only few models were found for concurrent biomass gasification [6-7]. The proposed model is based on mass and

energy balances for solid- and gas-phases, described by a dynamic system and one-dimensional equations, and involving the following components: *B* (biomass), *M* (moisture), and *C* (char), for solid-phase and *T* (tar), O_2, CO_2, H_2, H_2O (steam), CH_4, CO, and N_2, for gas-phase. The following processes are considered in the model: moisture evaporation (m), biomass pyrolysis (p_1), thermal cracking of the tars (p_2), combustion of the volatiles (c_1-c_4), and char combustion and gasification (c_5, g_1-g_3).
The main model assumptions are: single size of the particles; no momentum transfer; constant bed porosity; heat and mass transfer across the bed resulting from macroscopic (convection) and molecular (diffusion and conduction) exchanges; extra-particle mass transfer resistance; and solid- and gas-phase heat transfer with the reactor walls. In Table 2, the main equations of the kinetic model are depicted.

Table 2. Main equations of the kinetic model.

Mass balance for solid-phase:

$$\frac{\partial \rho_i}{\partial t} + \frac{\partial (\rho_i U_s)}{\partial z} = \sum_j \nu_{i,j} R_j \quad (7)$$

where i = *B*, *M*, *C*; $j = m, p_1$.
Overall mass balance for solid-phase:

$$\rho_s \frac{\partial U_s}{\partial z} = \sum_i \sum_j \nu_{i,j} R_j \quad (8)$$

where i = *B*, *M*, *C*; $j = m, p_1, c_5, g_1$-g_3.
Mass balance for gas-phase:

$$\varepsilon \frac{\partial \rho_i}{\partial t} + \frac{\partial (\rho_i U_g)}{\partial z} = \frac{\partial}{\partial z}\left(D_i \rho_g \frac{\partial x_i}{\partial z}\right) + Mw_i \sum_j \nu_{i,j} R_j + \omega_i \quad (9)$$

where i = *T*, O_2, CO_2, H_2, H_2O, CH_4, CO; $j = c_1$-c_5, g_1-g_3, wg, and ω_i is the rate of species i generated in solid-phase.
Overall mass balance for gas-phase:

$$\varepsilon \frac{\partial \rho_g}{\partial t} + \frac{\partial (\rho_g U_g)}{\partial z} = \sum_i \sum_j \nu_{i,j} Mw_i R_j + (1 - \nu_{C,p1}) R_{p1} + \sum_i \nu_{i,p2} R_{p2} \quad (10)$$

where i = *T*, O_2, CO_2, H_2, H_2O, CH_4, CO, N_2 and $j = m$, c_1-c_5, g_1-g_3, wg.
Energy balance for solid-phase:

$$\frac{\partial (\rho_s H_s)}{\partial t} = \frac{\partial}{\partial z}\left(\lambda_s^* \frac{\partial T_s}{\partial z}\right) + \frac{\partial (U_s \rho_s H_s)}{\partial z} - \sum_j R_j \Delta H_j - Q_{sg} - Q_{sw} \quad (11)$$

where $j = c_5$, g_1-g_3, p_1; Q_{sg} is the heat exchange between solid- and gas-phase; and Q_{sw} is the heat exchange between solid-phase and walls.
Energy balance for gas-phase:

$$\varepsilon \frac{\partial (\rho_g H_g)}{\partial t} = \frac{\partial}{\partial z}\left(\lambda_g^* \frac{\partial T_g}{\partial z}\right) + \frac{\partial (U_g \rho_g H_g)}{\partial z} - \sum_j R_j \Delta H_j + Q_{sg} - Q_{gw} \quad (12)$$

where $j = c_1$-c_4, wg, p_2; and Q_{gw} is the heat exchange between gas-phase and walls.
Pressure drop (modified Darcy law):

$$\frac{K}{\mu} \frac{\partial P}{\partial z} = U_s - U_g \quad (13)$$

where *K* is the permeability to gas flow.

The coefficients of heat and mass transfer are calculated according to Hobbs, Radulovic and Smoot [8]. The gas thermal conductivity and viscosity are described as in Purnomo, Aerts and Ragland [9]. The specific heats (available at 1000 K) and diffusivity are considered constants.

2.3. Implementation and validation of the models

The models were implemented in the equation-based process simulator EMSO (Environment for Modeling, Simulation and Optimization [10]). The validation of the equilibrium model was made in two steps: (1) comparing equilibrium constants of considered reactions with tabulated and predicted data from Sharma [4]; (2) comparing composition of gaseous product with experimental and predicted data from Melgar *et al.* [3]. For the kinetic model, it was validated by predictions of Di Blasi's [6] model.

3. Application to a combined gasifier-combustor unit

3.1. Description of the pilot unit

The pilot unit (Fig. 1) is basically formed by one combined gasifier-combustor and one air pollution control system. The gasifier is a downdraft type so that feedstock is

concurrent to gas flow. There are two incoming air points along gasifier height. Other three air injections are fed in the combustor to oxidize the fuel gas coming from the gasifier.

The input solid wastes are leather strips with ultimate analysis of 49.31%C, 8.52%H, 24.7%O, 12.42%N, and 1.83%S, and a HHV of 18,448 kJ/kg (d.b.). And their proximate analysis afford 77.3% volatile matter, 5.8% ash, and 16.9% carbon fixed (as received with 14.1% moisture content) [11]. A specific parameter related to this waste is the mass relation of O_2 to complete oxidation of biomass: 1.57 kg_{O2}/kg_B. The gasifier is a reactor of 1.465 m height and 0.611 m equivalent diameter, where the bed density is 400 kg/m^3 and the void fraction is 0.5. A typical operation involves 60 kg/h feedstock that is gasified by two air injection at 300 K of 102.4 Nm^3/h and 55.7 Nm^3/h, respectively. The acquisition of experimental data is made by 5 thermocouples (T1, T2, T11, T12, and T13) along the gasifier and 1 thermocouple (T5) in the final of the combustion chamber.

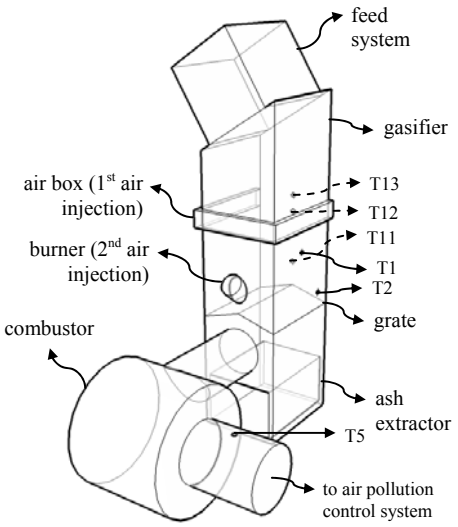


Figure 1. Combined gasifier-combustor unit.

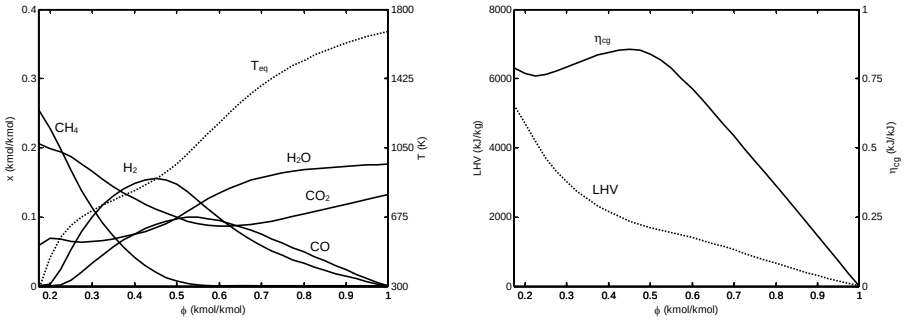
3.2. Simulation of the pilot unit

Sensitivity analyses of the process were carried out regarding to the feed air flow rate as an operating parameter. The responses to parametric changes were observed in output parameters: low heating value (LHV) of the energetic gas product (CO, H_2, and CH_4) and cold gas efficiency (η_{cg}), according to the equilibrium model. These output parameters are important to quantify the performance of the process. In this study, the feed air flow rate is related to the feedstock by a parameter called equivalence ratio (ϕ), which indicates the oxygen used relative to that required for complete combustion.

The predictions of the equilibrium model for typical operating conditions from measured and predicted data by Godinho [11] have shown good accuracy. Divergences could be justified by assuming an adiabatic system. Figure 2a shows the molar fractions of gaseous products, beyond adiabatic reaction temperature, related to the feeding air amount (equivalence ratio). Taking into account only components of energetic gaseous product, it is possible to observe that CH_4 increases with low equivalence ratios (ϕ<0.5) and CO and H_2 is maximum within $0.3 < \phi < 0.7$. The equilibrium temperature is minimum (T = 300 K, ambient temperature) at ϕ = 0.175 and maximum (T = 1680 K) at ϕ = 1, where CO_2 and H_2O (inconvenient to the process) are maximum and CO, CH_4, and H_2 are negligible. Cold gas efficiency (η_{cg}) and low heating value (LHV) to flue gas, related to conditions of Fig. 2a are shown in Fig. 2b. The maximum efficiency (85.5%) is reached at ϕ = 0.45. Therefore the optimal range for gasification of leather strips is $0.34 < \phi < 0.53$ (±5% maximum efficiency range).

The predicted gas composition at average conditions was also compared to experimental data, indirectly. It is possible to relate the measured temperatures in the reduction zone of the gasifier (T11) and the combustion chamber (T5) estimating the gaseous product at measured average temperature (906 K, ϕ = 0.467) and, from this fraction, the

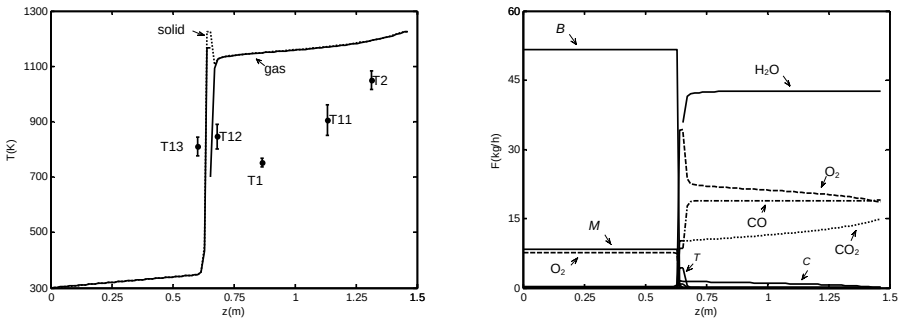
adiabatic flame temperature at a range of feeding O_2. At the last step of combustion, there is an average amount of 11% O_2 (v/v, d.b.) in excess [11] which corresponds to an adiabatic flame temperature of 1174 K. This value is close to the average value measured in T5: 1133 K.



(a) (b)

Figure 2. Properties to $0.175 < \phi < 1$ by the equilibrium model: (a) Gaseous product and adiabatic reaction temperature. (b) Cold gas efficiency and low heating value to energetic gaseous product.

Likewise, the same conditions of the process were simulated using the kinetic model at steady-state. In this case, the reactor was divided in three sections defined by the top-gasifier (0 m), first air injection (0.655 m), second air injection (1.315 m), and bottom-gasifier (1.465 m). The results present here were applied to the first two sections. In Fig. 3a, the predicted profiles have qualitative agreement with respect to experimental data.



(a) (b)

Figure 3. Results of kinetic model: (a) Temperature profiles of solid- and gas-phase, and measured data. (b) Mass flow rate of solid- and gas-phase species along gasifier before second air injection.

The inflection point in 0.655 m means the cold feeding air mixed to the preceding hot gas stream from first section. In Fig. 3b, the fast consumption of O_2 and formation of H_2O close to the air injection are observed. After that point, all small amounts of volatile materials are quickly consumed and, as result, there is an excess of O_2 to oxidize. For solid-phase (Fig. 3b), despite the fast conversion of biomass, small amounts of char are created. The inconsistencies associated to low release of volatiles and low formation of char are justified by the quality of kinetic parameters for pyrolysis reactions (p_1-p_2) adapted from literature data to leather wastes [12-13]. The uncertainties in the heat transfer coefficients also have significant role in the results. Another important factor to take into account, not observed in literature [6-7], is the reflux of air in two directions for one-dimensional representation by injecting air in intermediate

heights. In this case, the model considers that first feeding air (102.4 Nm^3/h) is portioned between the top (25%) and the first air injection point (75%).

4. Conclusion

The proposed models, equilibrium and kinetic, presented compatible results compared with literature data. The equilibrium model reached good accuracy to predict the steady-state behavior of the gasifier. With this approach it was possible to identify an optimum range of operation within $0.34 < \phi < 0.53$ and to confront measured temperature from the combustor with predicted data from the gasifier. Besides, due to poor quality of kinetic parameters and heat transfer coefficients, which are not full available for the specific case, the kinetic model could only predict qualitative steady-state behavior of the gasifier. Also, the representation of a three-dimensional system into one-dimensional model has limited prediction capability.

Acknowledgement

The authors would like to acknowledge the Brazilian agencies of research support: FAPERGS, CNPq, and FINEP; and the companies: Luftech Soluções Ambientais and Preservar Tratamento e Reciclagem de Resíduos.

References

1. C. Higman and M. van der Burgt, 2003, *Gasification*, Gulf Professional Publishing, Oxford.
2. X. Li, J. R. Grace, A. P. Watkinson, C. J. Lim, and A. Ergüdenler, 2001, Equilibrium modeling of gasification: a free energy minimization approach and its application to a circulating fluidized bed coal gasifier, *Fuel*, 80, 2, 195-207.
3. A. Melgar, J. F. Pérez, H. Laget, and A. Horillo, 2007, Thermochemical equilibrium modelling of a gasifying process, *Energy Conversion and Management*, 48, 1, 59-67.
4. A. K. Sharma, 2008, Equilibrium modeling of global reduction reactions for a downdraft (biomass) gasifier, *Energy Conversion and Management*, 49, 4, 832-842.
5. M. Baratieri, P. Baggio, L. Fiori, and M. Grigiante, 2008, Biomass as an energy source: thermodynamic constraints on the performance of the conversion process, *Bioresource Technology*, 99, 15, 7063-7073.
6. C. Di Blasi, 2000, Dynamic behaviour of stratified downdraft gasifiers, *Chemical Engineering Science*, 55, 15, 2931-2944.
7. F. V. Tinaut, A. Melgar, J. F. Pérez, and A. Horillo, 2008, Effect of biomass particle size and air superficial velocity on the gasification process in a downdraft fixed bed gasifier. An experimental and modelling study, *Fuel Processing Technology*, 89, 11, 1076-1089.
8. M. L. Hobbs, P. T. Radulovic, and L. D. Smoot, 1992, Modeling fixed-bed coal gasifiers, *AIChE Journal*, 38, 5, 681-702.
9. D. Purnomo, D. J. Aerts, and K. W. Ragland, 1990, *Pressurized downdraft combustion of woodchips*, 23th Symposium (International) on Combustion, 1025-1032.
10. R. P. Soares and A. R. Secchi, 2003, *EMSO: a new environment for modelling, simulation and optimization*, ESCAPE 13, Elsevier Science Publishers, 947-952.
11. M. Godinho, N. R. Marcílio, A. C. F. Vilela, L. Masotti, and C. B. Martins, 2007, Gasification and combustion of the footwear leather wastes, *Journal of the American Leather Chemists Association*, 102, 6, 182-190.
12. J. A. Caballero, R. Font, and M. M. Esperanza, 1998, Kinetics of the thermal decomposition of tannery waste, *Journal of Analytical and Applied Pyrolysis*, 47, 2, 165-181.
13. R. Font, J. A. Caballero, M. M. Esperanza, and A. Fullana, 1999, Pyrolytic products from tannery wastes, *Journal of Analytical and Applied Pyrolysis*, 49, 1-2, 243-256.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Tool for Optimizing the Design and Operation of Reverse Electro-Enhanced Dialysis of Monoprotic Carboxylic Acids

Oscar A. Prado-Rubio, Sten B. Jørgensen, Gunnar E. Jonsson

CAPEC, Depertment of Chemical and Biochemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark

Abstract

A dynamic model is derived to describe the simultaneous transport of multiple ions through anion exchange membranes and Nernst diffusion layers in a cell of the Reverse Electro-Enhanced Dialysis (REED) module. The model is based on first principles for dissociation, diffusion, convection and migration of species commonly found in a fermentation broth. The approach leads a system of multiregion partial differential equations that are solved numerically. The developed tool provides insight into the transport phenomena in the electrochemical system when it is operated with or without imposing an electrical field. Potentially, it can be applied to optimize the design and the operation of the REED module for different production scenarios.

***Keywords**: Reverse Electro-Enhanced Dialysis (REED), Donnan dialysis (DD), lactate recovery, ions transport modelling.*

1. Introduction

The removal of carboxylic acids from diluted solutions has been a complex separation problem. Presently, economical and sustainability issues are the driving competitive forces behind improving process design and operation. Particularly, improvements in sustainable production of commodity chemicals, such lactic acid, are desirable. Lactic acid is attractive since it is widely used in industry and recently the most interesting industrial application is as feedstock for the production of Poly-Lactic Acid (PLA). Derivates of PLA are sustainable polymers which have potential to substitute hydrocarbon based polymers in several applications. Even though PLA polymers can not be considered completely biodegradable anymore, it will definitely reduce our dependency on fossil feedstock.

It is well known that the bioproduction of lactic acid by lactic acid bacteria is impaired by lactate inhibition, like many other fermentation processes at a certain concentration level of the product or one of the bi-products. From this point of view, reducing the concentration of the biotoxic lactate in the culture solution will enhance the productivity and product yield, under controlled pH environment. Integrating separation and fermentation will also enable operation at higher cell densities, thereby providing additional enhanced productivity potential.

Since 1960's, membrane separation processes have been suggested as an alternative for lactic acid removal and biomass confinement, their application is well documented. The purpose of this paper is to depict a model which describes the multiple ion transport through anion exchange membranes under current reversal conditions, when the separation is carried out by Reverse Electro-Enhanced Dialysis. This device was

recently designed, its performance suggests it to be a promising alternative for continuous removal of lactic acid during fermentation (Rype, 2003). This paper is organized as follows. In section 2 the system is described and the assumptions are depicted. Afterwards, the model is presented. Section 3 addresses how the model is solved and a brief description of the parameter estimation procedure. In addition, some simulation results of ion transport under current reversal conditions are discussed. Finally, the conclusions are drawn.

2. Modelling Reverse Electro-Enhanced Dialysis

Since electrically driven membrane separation processes have been proposed as an *in situ* alternative for removal of carboxylic acids in 1980's, several problems have been pointed out: membrane fouling due to the biomaterial content of a fermentation broth and the presence of divalent ions generates scaling and, when bipolar membranes are used for a further concentration and recovery of the lactic acid, those ions can impair the membrane (Hongo *et al*., 1986; Rype, 2003).

The REED design and operation minimize the impact for the above mentioned aspects. The REED can be described as a Donnan dialysis (DD) device which is operated as Electrodialysis Reversal (EDR) module. The REED stack is composed by several cells in parallel. A cell consists of two feed channels with one dialysate channel in between, separated by anion exchange membranes (AEM). Using only AEM, scaling and bipolar membrane degradation problems are avoided due to Donnan exclusion. When no currrent is applied, fouling problem is reduced employing high flow velocities and a destabilization mechanism, the last is generated by hydroxyl flux through the membrane in opposite direction of the fouling layer formation. However, the main drawback using no current is the rather low anion flux. REED design emerges as a potential method to enhance the lactate fluxes in conventional Donnan dialysis operation, this is done by imposing an external electrical field. The external potential gradient changes the transport mechanism from counter ion transfer to a competitive ion transport. In addition, the unfavorable fouling influence is reduced further by periodically reversing the current (Rype, 2003).

Even though ion exchange membranes have been used widely in industry, the transport mechanism has not yet been completely understood. Thus, it is desirable to develop reliable models which can provide a better understanding of the transport phenomena in such electrochemical systems. Furthermore, models will enable optimization of the design and operation of the modules.

It is challenging to model this system, since several phenomena must be taken into consideration such as diffusion, the influence of electrical potential, migration effects, ion dissociation degree and membrane water content. A dynamic model in two spatial dimensions is derived from first principles for dissociation, diffusion, convection and migration of the species for a cell of the REED stack, by employing an irreversible thermodynamic approach. The modelled cell is depicted in Fig. 1, the system is divided in regions: three the bulk solutions, two membranes and four boundary layers. The species found in each section are also shown in the figure, for the lactic acid recovery case.

2.1. Model assumptions

Due to the complexity of real REED system, any theoretical analysis is of necessity based on a simplified model of the actual process. The assumptions taken are listed.

- General: electroneutrality condition at any location in the system, current is carried by ions, constant temperature, ideal solution and the process is carried out at sub-limiting current densities. Species included in the model are: carboxylic anion, hydroxyl, sodium, dissociated protein, carboxylic acid and undissociated protein. A protein specie is included to account for the buffer effect commonly found in cultivation solutions.
- Membrane: there are diffusive and electrophoretic transport in the *x*-direction, transport of water by osmosis and electro-osmosis is neglected, there is no transport of uncharged or large molecules through the membrane, equilibrium at membrane surface and constant membrane dimensions.
- Boundary layer: diffusive and electrophoretic transport in the *x*-direction are investigated, convective transport is neglected and the thickness of the boundary layers are constant for a given flow condition.
- Bulk channels: there is convective transport in *y*-direction in the bulk channels and tank is series model is used.

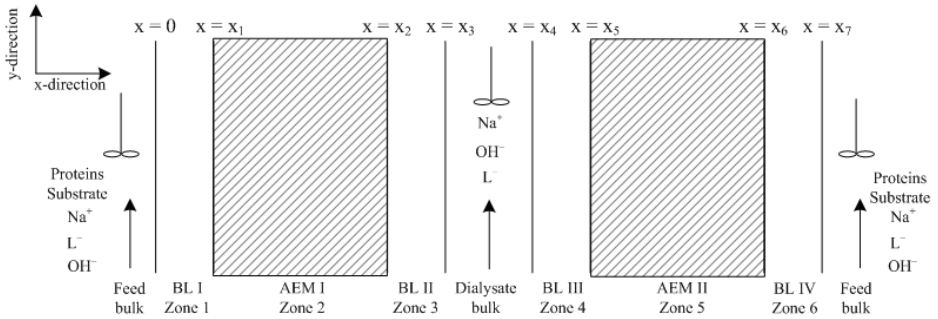


Figure 1: sketch of a cell in the REED module for lactate recovery. BL: boundary layer, AEM: anion exchange membrane, x_j: interface *j*

2.2. Nernst layer and membrane transport

Mass balances are formulated for transport through boundary layers and membranes in *x*-direction. Substances and phases are denoted by the subscripts *k* and *p*, respectively.

$$\frac{\partial C_{k,p}}{\partial t} + \nabla J_{k,p} - \Delta R_{k,p} = 0 \tag{1}$$

The species dissociation and buffer effect are introduced into the model employing the reaction rate term ($\Delta R_{k,p}$). Nernst-Plank equation for ideal solution is used to estimate the fluxes $J_{k,p}$, this equation describes diffusion and migration through the section. The convective transport is neglected in *x*-direction (Strathmann, 2004):

$$J_{k,p} = -D_{k,p}\left(\frac{\partial C_{k,p}}{\partial x} + \frac{z_k F C_{k,p}}{RT}\frac{\partial \psi}{\partial x}\right) \tag{2}$$

Where $D_{k,p}$ is the diffusion coefficient, z_k the valence, F is the Faraday number, R is the ideal gas constant, T is temperature and ψ is the electrical potential. For undissociated species, the terms involving potential gradients are eliminated since the current is exclusively carried by ions. The required first and second derivatives of the potential can be calculated using the assumption that the entire current I_d is carried by ions, substituting the Nernst-Plank equation into Eq. 3 .

$$I_d = \sum_k z_k F J_{k,p} \tag{3}$$

2.3. Bulk channel model

Mass balances are approximated using tanks in series approach in the *y*-direction. Each tank has mass exchange with the adjacent membranes. The dissociation reactions are present as well. The mass balance for each tank in the feed channel is depicted in equation (4), where q is the flow rate and h, L, W are the height, length and width of the channel, respectively. The model for the dialysate channel is completely analogous.

$$\frac{dC_k^{feed}}{dt} = \frac{q_{feed}}{h_{feed} L W}\left(C_k^{feed,in} - C_k^{feed}\right) + \frac{1}{h_{feed}}\left(J_k\big|_{x=x_7} - J_k\big|_{x=x_o}\right) + \Delta R_{k,p} \tag{4}$$

2.4. Boundary conditions at the interfaces

At the membrane surface, the electrochemical potential in the solution and on the membrane side are identical, this condition is denoted as Donnan equilibrium. This condition defines the potential build up at the membrane interface and determines the transport of charged molecules. The following relation arises for an ideal solution (Strathmann, 2004). Where $\Delta\psi_{Don}$ is the Donnan potential and superscripts s and m correspond to solution and membrane sides of the interface, respectively.

$$\Delta\psi_{Don} = \frac{RT}{z_k F} \ln\left(\frac{C_k^s}{C_k^m}\right) \tag{5}$$

The fluxes are continuous at the interfaces, it means there is no accumulation. In equation (6), the points x_j^- and x_j^+ correspond to left and right hand side of the interface located at x_j, respectively (Fig.1).

$$J_k\big|_{x=x_j^-} = J_k\big|_{x=x_j^+} \tag{6}$$

Electroneutrality condition defines the concentration distribution in both membrane and solution zones. For the solution and membrane sides the following algebraic relations arise:

$$\sum_k z_k C_k^s = 0 \quad \text{and} \quad \sum_k z_k C_k^m + z_{fix} C_{fix}^m = 0 \tag{7}$$

Where the subscript *fix* is related to the fixed charge in the membrane. Finally, the assumption that all current is carried by ions remains.

3. Results and discussion

3.1. Model Solution

The model consists of a system of multiregion partial differential equations. The method of lines is employed to discretize the spatial *x*-dimension, resulting in a system of differential and algebraic equations (DAEs). Sixth order Taylor expansion was used to achieve the desired accuracy using a reasonable computational time. Besides, asymmetric centered differences were employed to deal with the equilibrium boundary conditions. A feasible set of initial conditions must be known in order to ensure convergence. An inconsistent set of initial conditions leads to numerical problems (index of the DAE's is larger than one). Møllerhøj (2006) proposed an initialization procedure in order to guarantee, to some extend, convergence.

3.2. Parameter estimation for model tuning

The model is tuned for neutralization dialysis of some monoprotic carboxylic acids, when no current is imposed to the system. The experimental data are taken from Zheleznov (1998). From a sensitivity analysis, the thickness of the boundary layers and the ion exchange capacity of the membrane were fixed. Employing the interior reflective Newton method for non linear minimization subject to bounds (Coleman and Li, 1994), a black box swelling model was tuned for three different monoprotic carboxylic acids: acetate, lactate and propionate. Details about the model tuning are illustrated by Prado Rubio *et al.*, 2009.

3.3. Ions transport under current reversal conditions

Membrane fouling is an important factor since carboxylic acids are recovered from fermentation broths, which are characterized by high biomaterial content that easily can be adsorbed at the membrane surface. The adverse fouling influence in REED is diminished by periodically reversing the current density. REED is operated using a constant absolute value of current density, this implies that the strength of the external potential gradient must be enlarged with time to counteract the increasing electrical resistance. Through simulations, the potential average lactate recovery is estimated under current reversal conditions, for pseudo steady state operation. For this simulation case, the imposed current is $\pm 100\ A/m^2$ and the inlet lactate and protein concentrations in feed channel are *100*, *10* mol/m^3, respectively. Besides, inlet hydroxyl concentration in dialysate channel is *50* mol/m^3. The results are depicted in Fig. 2. It is expected to obtain higher recoveries by increasing the reversal time as the concentration profiles develop. The recovery achieved without imposing an external potential gradient is located at zero reversal time (Donnan dialysis conditions). It can be seen that for low reversal times the average lactate recovery increases linearly, this condition remains up to a reversal time of 20 min. For higher reversal times, lactate recovery reaches a limit since the maximum fluxes are practically achieved before the current is reversed. The maximum average lactate recovery is approximately 58% higher that under Donnan dialysis conditions, and it is obtained for $t_{rev} \approx 76.67\ min$. On the other hand, average lactate recovery obtained with Donnan dialysis is higher than expected. The reason being that the experiments used pure carboxylic acid solutions, therefore the fluxes were not affected by fouling (Zheleznov, 1998).

In spite of the maximum lactate recoveries are obtained at long time reversal times, the main drawback is the energy requirements to keep constant current density operation. For that reason is important to identify a potential REED operating window. In previous experimental work in a REED cell, the potential build up during lactic acid recovery was measured (Rype, 2003). From there, the potential gradient can be correlated with reversal time (shown in Fig. 3). Prolonged operation at constant current requires a higher potential gradient and therefore the energy consumption increases as well. The regressed model predicts a continuous linear increment of the potential gradient. However, there is a maximum potential gradient allowed by power source which defines the operative window for REED unit. Once the current source reaches its limit, operation at constant current conditions is not longer feasible. This means the operation mode switches to constant voltage conditions, implicating a reduction in the ion fluxes and lactate recovery. It must be stressed that the presented model was not developed to predict operation under constant voltage conditions. The optimal operating point for the equipment represents a trade off between the carboxylic anion recovery and the energy consumption, subject to the operation constraints.

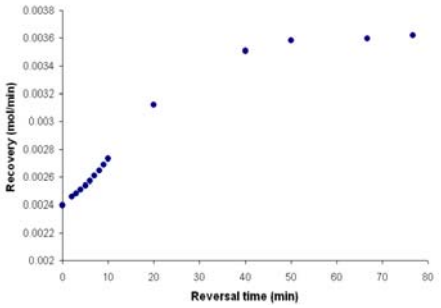


Figure 2: average lactate recovery as a function of the reversal time for a current of I_d=±*100 A/m*^{2} and $C_{dbin,NaOH}$=*50 mol/m*^{3}

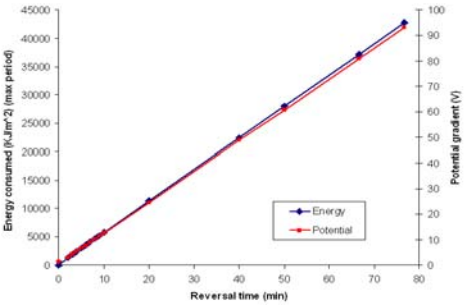


Figure 3: maximum potential and energy requirements for REED

4. Conclusions

Using first principles, simultaneous transport of multiple ions across anion exchange membranes in a REED cell is modelled. The solution of the system of multiregion partial differential equations is approximated numerically, employing sixth order Taylor expansion with asymmetric centered differences. The tuned model is used to investigate the average lactate recovery as a function of the reversal time, under pseudo steady state operation. Maximum recovery is basically achieved for t_{rev}>*50* min. However, there is a maximum potential gradient allowed by the power source that ultimately defines the operative window. Finally, this model is implemented as a tool which can be used to study the competitive ion transport through ion selective membranes under current load conditions, and therefore it is useful for development of an optimal design and operation of REED in dependence of different objective functions.

Acknowledgments

This project is carried out within the Bioproduction project which is financed by the 6th Framework Programme, EU.

References

Coleman, T. and Li, Y. (1994). On the Convergence of Interior-Reflective Newton Methods for Nonlinear Minimization Subject to Bounds. *Mathematical Programming*, 67, 189–224.

Fila, V. and Bouzek, K. (2003). A Mathematical Model of Multiple Ion Transport Across an Ion-Selective Membrane under Current Load Conditions. Journal of Applied Electrochemistry, 33, 675–684.

Hongo, M.; Nomura, Y. and Iwahara, M. (1986). Novel Method of Lactic Acid Production by Electrodialysis Fermentation. *Applied and Environmental Microbiology*, 52(2), 314–319.

Møllerhøj, M. (2006). Modeling the REED Process. Master's thesis, Technical University of Denmark.

Prado Rubio, O.A.; Jørgensen, S.B. and Jonsson, G. (2009) Lactic Acid Recovery in Electro-Enhanced Dialysis: Modelling and Validation. *Accepted for publication at ESCAPE19.*

Rype, J. (2003). Modelling of Electrically Driven Processes. Ph.D. thesis, Technical University of Denmark.

Strathmann, H. (2004). Ion-Exchange Membrane Separation Processes. Membrane Science and Technology Series, 9. Elsevier.

Zheleznov, A. (1998). Dialytic Transport of Carboxylic Acids through an Anion Exchange Membrane. *Journal of Membrane Science*, 139, 137–143.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Process Structure Optimization using a Hybrid Disjunctive-Genetic Programming Approach

Wei Yuan[a], Andrew Odjo[b], Norman E. Sammons, Jr.[a], Jose Caballero[b] and Mario R. Eden[a]

[a]*Department of Chemical Engineering, Auburn University, Auburn and 36830, USA,*
[b]*Chemical Engineering Department, University of Alicante, Spain*

Abstract

Discrete optimization problems, which give rise to the conditional modelling of equations through representations as logic based disjunctions, are very important and often appear in all scales of chemical engineering process network design and synthesis. Disjunctive-Genetic Programming (D-GP), based on the integration of Genetic Algorithm (GA) with the disjunctive formulations of the Generalized Disjunctive Programming (GDP) for the optimization of process networks, has been proposed in this work. With the increase in the problem scale, dealing with such alternating routes becomes difficult due to increased computational load and possible entanglement of the results in sub-optimal solutions due to infeasibilities in the MILP space. In this work, the genetic algorithm (GA) has been used as a jumping operator to the different terms of the discrete search space and for the generation of different feasible fixed configurations. This proposed approach eliminates the need for the reformulation of the discrete /discontinuous optimization problems into direct MINLP problems, thus allowing for the solution of the original problem as a continuous optimization problem but only at each individual discrete and reduced search space.

Keywords: Process Optimization, Genetic Algorithm, Generalized Disjunctive Programming

1. Introduction

Processes with nonlinear functions and discontinuities in the objective and/or constraint space are found in a large number of synthesis problems. It has been shown that using disjunctions for expressing the discrete decisions, which conditions the selection of process units among various alternatives, can be very beneficial in handling of discontinuities in chemical process systems (Vecchietti et al., 2003; Turkay and Grossmann, 1996). The Generalized Disjunctive Programming (GDP) is an efficient method for Mixed Integer Non-Linear Programming (MINLP) for discrete/continuous optimization problems (Turkay and Grossman, 1996; Lee and Grossman, 2000). In GDP, problems are modeled with Boolean and continuous variables for the optimization of a given objective function subject to different types of constraints. GDP represents discrete decisions in the continuous space with disjunctions, and constraints in the discrete space with logic propositions.

Evolutionary search methods, of which the most widely used is the Genetic Algorithm, originally initiated by Holland (1975), has been known to be less susceptible to the existence of locally optimal solutions in optimization problems, yielding very good

solutions even with discontinuous objective and/or constraint functions. The evolutionary search method solutions have been found in some cases, to outperform those obtained through traditional deterministic approaches, by exhibiting robustness through the use of the objective function information and not derivatives, as well as by the ease at which they handle discrete and integer variables as well as non-smooth and non-continuous functions (Androulakis and Venkatasubramanian, 1991). Yet, the overall applicability of these algorithms to constrained problems remains an active research focus, especially taking into consideration the fact that they may exhibit slow convergence and may have difficulties finding the optimal solution to a problem with very small feasibility space. However, promising results are being obtained in joint heuristic and evolutionary (deterministic and stochastic) approaches to the efficient solution of engineering processes (Leboreiro and Acevedo, 2004).

This work is based on the integration of GA with the disjunctive representation of discrete/continuous optimization problems. The strategy involves the decoupling of the disjunctions and the propositional logics from the overall disjunctive formulation to the GA space where these constraints are treated before being returned as active terms with satisfied logics to the resulting NLP space. The population of disjunctive terms, corresponding to the different process superstructures or configurations are manipulated by the genetic operators in an evolutionary manner in order to obtain the best NLP solution. The implication of this is the final solution of an optimization problem with smooth nonlinear constraints and objectives confined within a reduced search space corresponding to specific active terms of a set of disjunctions, as determined by the GA. An advantage of this approach is that, apart from the fact that the dimensionality of the optimization problem can be considerably reduced due to the reduction in the total number of constraint equations (only common and active disjunctive terms equations are analyzed at each NLP call), only feasible superstructures are generated.

2. Disjunctive-Genetic Programming Method

The basic idea behind the D-GP approach involves the decoupling of the disjunctive terms represented by Boolean variables as well as the propositional logic constraints these variables form, to the genetic algorithmic space where they are encoded into chromosomes with specific structures corresponding to potential superstructure solution alternatives. In a typical solution framework such as the Big M - OA approach (Williams, 1999), the reformulation of optimization problems with discontinuous functions modeled with disjunctions involves conversion of the Boolean variables to binary and the proposition logics into simple linear equations, leading to MINLP problems, which in turn are solved using an iterative MILP-NLP scheme (Fig. 1).

An important feature of our proposed approach is that no MINLP reformulation of the disjunctive representation is needed. Terms of the disjunctions only need to be identified, well defined and coded into strings or chromosomes. For the adaptation of GA to the efficient handling of the generated chromosomes, we propose a segment-based crossover and mutation strategy. Furthermore, we apply the all-feasible populations approach, consisting of the creation and use of chromosomes that always satisfy the logic constraints (Odjo et al., 2008). The basic steps of the D-GP framework are depicted in Fig. 2.

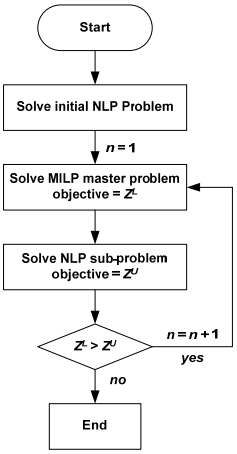


Figure 1. *MILP-NLP iterative solution*

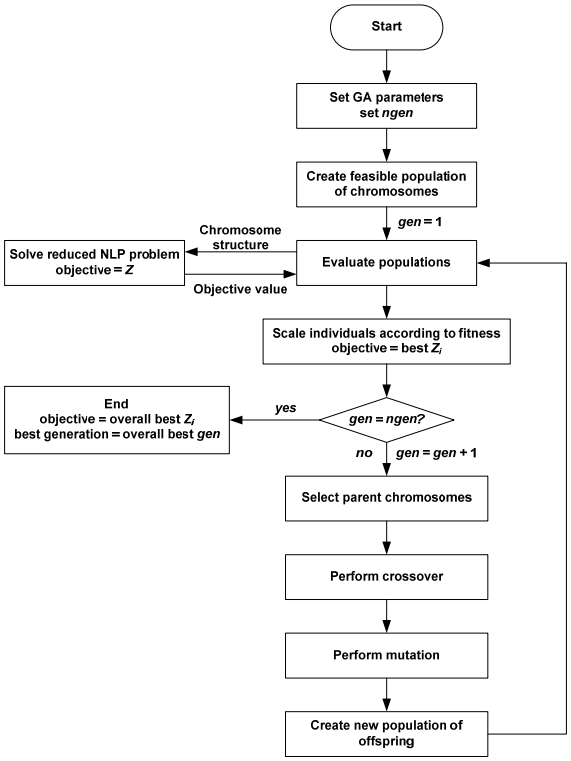


Figure 2. *The D-GP solution framework*

An important step in the application of GA to the solution of optimization problems is the creation of random individuals of a population with a fixed number of members. This population typically corresponds to the search space of the optimization problem, for example randomly generated temperature or pressure values between a specified lower and upper limit. However, in large synthesis problems with large numbers of variables, this gives rise to huge combinatorial problems such as possible different flowsheet configurations of a process with large number of units and process streams. Coupled with the existence of a narrow feasibility region due to different specifications and restrictions, the traditional approach for the generation of individuals within the optimization search space might lead to slow convergence and even sub-optimal solutions. In this work, a population of feasible chromosomes approach is adopted. The important aspect of this approach is that infeasible chromosomes have been eliminated from the initial population before the GA fitness function evaluation stage, as opposed to the traditional approach where feasibility constraints defined by the logic propositions and embedded as linear algebraic equations were satisfied during the GA fitness function evaluation.

3. Segment-Based Cross-over and Mutation Strategy

In order to conserve parts of the feasible configuration from the parents to the offspring, and also to achieve less disruptive chromosomes, the crossover operation strategy

employed in this work involved the selection of two strings (parent chromosomes) with the subsequent identification and exchange of identical segments of the parents. The number of identical segments in a parent string or chromosome is independent of the size of the optimization problem, but depends on the number of disjunctions in the synthesis problem. The genes in each segment represent the terms of the disjunctions. The rules governing the cross-over of identical segments, involve the selection and direct inter-exchange of the bulk of those segments, but not individual genes in them.

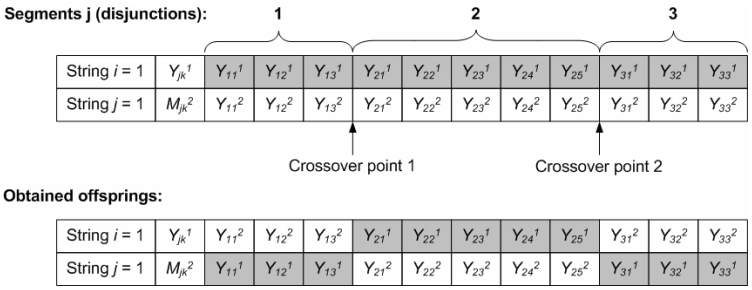


Figure 3. *Adopted crossover operation strategy for the D-GP approach*

Mutation, carried out at random on any bit of a string (gene) of the population, often leads to infeasible configurations. This can be improved, however, if the operator is modified by incorporating specific information on desired mutation routes. Rather than carry out mutations on individual genes, mutations are performed on segments of specific size and position in the string, corresponding to individual disjunctions and the order in which they are placed, respectively. It is worth noting, that the length of each segment (the number of genes in it) corresponds to the number of terms in each disjunction. In this work, this leads to the creation of mutant genes, which basically can be considered as sets of disjunctions of the current problem with different configurations of the propositional logics. In this way, diversity of the next population is guaranteed without loss of feasibility (Odjo et al., 2008).

4. Case Study – Process Network

This process network problem was taken from the works of Lee and Grossmann (2000). It was sought to determine the most profitable flowsheet configuration involving the possible selection of eight processes from the network as represented in Fig. 4.

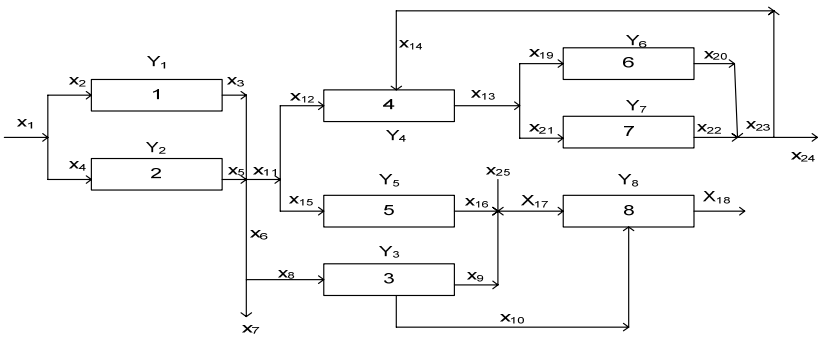


Figure 4. *Superstructure for process network*

This model was originally formulated as a MINLP and later as a GDP problem by Lee and Grossmann (2000). The problem formulated as a D-GP problem involves disjunctions for the selection of units, and propositional logic for the relationship of these units. Each disjunction contains the equations for the given unit. The model is shown in Eqn 1.

$$\min Z = \sum_{k=1}^{8} c_k + a^T x + 122$$

s.t.

$$x_1 - x_2 - x_4 = 0,\quad x_6 - x_7 - x_8 = 0,\quad x_3 + x_5 - x_6 - x_{11} = 0,\quad x_{13} - x_{19} - x_{21} = 0,\quad x_{17} - x_9 - x_{16} - x_{25} = 0,\quad x_{11} - x_{12} - x_{15} = 0,$$

$$x_{23} - x_{20} - x_{22} = 0,\quad x_{23} - x_{14} - x_{24} = 0,\quad x_{10} - 0.8x_{17} \le 0,\quad x_{10} - 0.4x_{17} \ge 0,\quad x_{12} - 5x_{14} \le 0,\quad x_{10} - 2x_{14} \ge 0,$$

$$\begin{bmatrix} Y_1 \\ \exp(x_3) - 1 - x_2 = 0 \\ c_1 = 5 \end{bmatrix} \vee \begin{bmatrix} \neg Y_1 \\ x_3 = x_2 = 0 \\ c_1 = 0 \end{bmatrix} \qquad \begin{bmatrix} Y_2 \\ \exp(x_5/1.2) - 1 - x_4 = 0 \\ c_2 = 8 \end{bmatrix} \vee \begin{bmatrix} \neg Y_2 \\ x_4 = x_5 = 0 \\ c_2 = 0 \end{bmatrix} \quad \begin{bmatrix} Y_3 \\ 1.5x_9 - x_8 + x_{10} = 0 \\ c_3 = 6 \end{bmatrix} \vee \begin{bmatrix} \neg Y_3 \\ x_9 = x_8 = x_{10} = 0 \\ c_3 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_4 \\ 1.5(x_{12} + x_{14}) - x_{13} = 0 \\ c_4 = 10 \end{bmatrix} \vee \begin{bmatrix} \neg Y_4 \\ x_{12} = x_{13} = x_{14} = 0 \\ c_4 = 0 \end{bmatrix} \quad \begin{bmatrix} Y_5 \\ x_{15} - 2x_{16} = 0 \\ c_5 = 6 \end{bmatrix} \vee \begin{bmatrix} \neg Y_5 \\ x_{15} = x_{16} = 0 \\ c_5 = 0 \end{bmatrix} \qquad \begin{bmatrix} Y_6 \\ \exp(x_{20}/1.5) - 1 - x_{19} = 0 \\ c_6 = 7 \end{bmatrix} \vee \begin{bmatrix} \neg Y_1 \\ x_{19} = x_{20} = 0 \\ c_6 = 0 \end{bmatrix}$$

$$\begin{bmatrix} Y_7 \\ \exp(x_{22}) - 1 - x_{21} = 0 \\ c_7 = 4 \end{bmatrix} \vee \begin{bmatrix} \neg Y_7 \\ x_{21} = x_{22} = 0 \\ c_7 = 0 \end{bmatrix} \qquad \begin{bmatrix} Y_8 \\ \exp(x_{18}) - 1 - x_{10} - x_{17} = 0 \\ c_8 = 5 \end{bmatrix} \vee \begin{bmatrix} \neg Y_8 \\ x_{10} = x_{17} = x_{18} = 0 \\ c_8 = 0 \end{bmatrix}$$

$$a^T = [0,1,-10,1,-15,0,0,0,-40,15,0,0,0,-40,15,0,0,0,15,0,0,80,-65,25,-60,35,-80,0,0,-35]$$

$$x_j, c_k \ge 0, Y_k \in \{true, false\}, k = 1,2,...,8; j = 1,2,...,25.$$

propersitional $\log ic[\Omega = (Y_k)]$ $Y_1 \Rightarrow Y_3 \vee Y_4 \vee Y_5$ $Y_2 \Rightarrow Y_3 \vee Y_4 \vee Y_5$ $Y_3 \Rightarrow Y_1 \vee Y_2$ $Y_3 \Rightarrow Y_8$ $Y_4 \Rightarrow Y_1 \vee Y_2$ $Y_4 \Rightarrow Y_6 \vee Y_7$

$Y_5 \Rightarrow Y_1 \vee Y_2$ $Y_5 \Rightarrow Y_8$ $Y_6 \Rightarrow Y_4$ $Y_7 \Rightarrow Y_4$ $Y_8 \Rightarrow Y_3 \vee Y_5 \vee (\neg Y_3 \wedge \neg Y_5)$

specifications $Y_1 \forall Y_2$ $Y_4 \forall Y_5$ $Y_6 \forall Y_7$ (1)

As can be seen in Eqn. 1, each chromosome contained 10 genes, with the first eight containing the information on the activation or non-activation of a certain term of each disjunction. The GA returned a chromosome with an optimal objective value of 68.01 at the optimal flowsheet configuration that involves selection of units 2, 4, 6 and 8. It implies the feasible chromosome was generated during the GA first population creation (with the structure 01010101(+1)). Table 1 shows the comparison of the problem reformulated from the D-GP to the MINLP and iteratively solved between the NLP sub-problem and MINLP master problems (Lee and Grossmann, 2000). It can be seen that the reduced NLP problem of the D-GP approach is more simplified by the existence of fewer nonlinear constrains.

Table1. Comparison of NLP model parameters of G-DP with GAMS SNOPT7 output.

Model parameters	Formulations	
	Reduced NLP (D-GP)	NLP
Objective value	68.01	68.01
Number of variables	33	34
Total number of equations	38	61
Constraints: Linear	10	10
Nonlinear	28	51
Number of bonds	33	34

5. Conclusions

In this paper a joint genetic algorithm-disjunctive representation approach to synthesis of process networks involving discrete/discontinuous functions has been presented. Special modified GA operators were presented to better handle the proposed solution approach, and it was found that the generation of feasible populations yield better results as compared to the case when the population is composed of randomly generated individuals that underwent feasibility test at the fitness function evaluation stage. An interesting feature of the D-GP approach is that it eliminates the need for the reformulation of the GDP problems into a direct MINLP problem, thus allowing the solution of the original problem as a continuous optimization problem but only at each individual discrete and reduced search space. The effectiveness of the D-GP algorithm has been illustrated using a benchmark case study of process network design. An important aspect of this work is the use of an evolutionary algorithm for the efficient handling of discontinuities in optimization problems, while simultaneously applying deterministic approaches to handle the continuous functions. Reformulation of the benchmark case showed significant reductions in the number of constraints/equations compared to the standard formulation. The objective function value in this case remained unchanged, but other examples have been solved where improved objective function values have been obtained. No comparison of CPU time is included at this stage, as the software implementation is not yet seamless and as a result, there is some lag time when transferring information between the GA solver in MATLAB and the NLP solver in GAMS. In future works, this lag time will be eliminated to enable direct comparison of solution time.

References

I.P. Androulakis, V. A. Venkatasubramanian, 1991, Genetic algorithmic framework for process design and optimization. Computers and Chemical Engineering, 15, 4, 217–228.

J. H. Holland, 1975, Adaptation in Natural and Artificial Systems, Ann Arbor: University of Michigan Press.

J. Leboreiro, J. Acevedo, 2004, Process synthesis and design of distillation sequences using modular simulators: a genetic algorithm framework, Computers and Chemical Engineering, 28, 1223-1236.

S. Lee, I.E. Grossmann, 2000, New algorithm for Nonlinear Generalized Disjunctive Programming, Computers and Chemical Engineering, 24, 9-10, 2125-2141.

S. Lee, I.E. Grossmann, 2003, Global Optimization of Nonlinear Generalized Disjunctive Programming with Bilinear Equality Constraints: Applications to Process Networks. Computers and Chemical Engineering, 27, 1557.

A.O. Odjo, N.E. Sammons Jr., A. Marcilla, M.R. Eden, J. Caballero, 2008, A Disjunctive-Genetic Programming Approach to Synthesis of Process Networks, Proceedings of 18th International Congress of Chemical and Process Engineering (CHISA).

R. Raman, I. E. Grossmann, 1994, Modeling and Computational Techniques for Logic Based Integer Programming. Computers and Chemical Engineering, 18, 563.

M. Turkay, I.E. Grossmann, 1996, Logic-Based MINLP Algorithms for the Optimal Synthesis of Process Networks. Computers and Chemical Engineering, 20, 959.

A. Vecchietti, S. Lee, E. I. Grossmann, 2003, Modeling of discrete/continuous optimization problems: characterization and formulation of disjunctions and their relaxations. Computers and Chemical Engineering, 27(3), 433-448.

H.P. Williams, 1999, Model Building in Mathematical Programming, John Wiley and Sons, Ltd, Chichester.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

PETROX – PETROBRAS Technology in Process Simulation

Jacques Niederberger[a], Márcia S. Gama[a], Letícia C dos Santos[a], José A. da Silva[b], Carmen E. Vargas[c], Victor R. R. Ahón[c], Ellen P. Silva[c], Domingos F. de S. Souza[c], Cláudio A. dos S. Aquino[a], Joyce S. S. Aires[a]

[a]*Petrobras - Petróleo Brasileiro S.A., Research and Development Center (CENPES), Av. Horácio Macedo, n. 950, Rio de Janeiro - RJ, 21941-915, Brazil.*
[b]*SPASSU Tecnologia e Serviços, Avenida Princesa Isabel, 629, 6° andar, Centro - Vitória – ES, 29010-904, Brazil.*
[c]*Projectus Consultoria Ltda, Rua General Jardim, 770 - 4º andar, São Paulo – SP, 01223-010, Brazil.*

Abstract

Petrobras – Petroleo Brasileiro S.A., is an active player in research, oil and gas exploration and processing, petrochemical and fertilizers industries as well as in bio fuels production. As the process designs are elaborated with a high degree of innovation, the utilization of an in-house process simulator has strategic advantages.
PETROX – Petrobras' Process Simulator comes from this conception. The simulator is being developed in Petrobras' Development & Research Center since the late 80's and is now on its 20 year birthday. PETROX' success over 20 years, coexisting with commercial simulators, is due to its insertion in Petrobras' Technology System and to the importance that the company gives to innovation. This work will discuss general and specific features of the simulator, structure and its utilization in research, development and design of production facilities and processing units of oil and gas.

Keywords: oil refining, modeling, simulation, design.

1. Introduction

Petrobras – Petroleo Brasileiro S.A., is an integrated energy company which plays in oil and gas exploration and processing, petrochemical and fertilizers industries as well as in bio fuels production. The company researches and develops technology on deep water oil and gas production, as well as refining technologies, in CENPES, its research and development center. These new/improved technologies are applied in the design of production facilities and refining units. The concretization of the new technologies in engineering designs, in the RD&E concept (research, development and engineering), passes through the elaboration of mathematical models of the new processes and the implementation of these models in computer programs for process simulation. Most part of the oil refined by Petrobras is domestic, so it is necessary to research and develop novel refining technologies and process schemes, as well as catalysts suited to processing national feed stocks. As the process designs are elaborated with a high degree of innovation, the utilization of an in-house process simulator has strategic advantages. In this scenario, building process models and implementing them in simulators is a very important part in the technology development.

PETROX – Petrobras' Process Simulator is fruit of this conception. The simulator is being developed in Petrobras' Development & Research Center since the late 80's and is now on its 20 year birthday. PETROX' success, coexisting with commercial simulators, is due to its insertion in Petrobras' Technology System and to the importance that the company gives to innovation. These factors are responsible for the characteristics that make PETROX unique, when compared to commercial process simulators.

Examples of these unique characteristics are the models of proprietary conversion process like fluid catalytic cracking, delayed coking and hydro treating; rigorous heat exchanger models from Heat Transfer Research Inc (HTRI) and the integration with software from Fractionation Research Inc (FRI) and the Centre for Process Integration (CPI), from the University of Manchester.

Petrobras' affiliations to consortiums like HTRI, FRI, CPI, Center for Phase Equilibria and Separation Processes (IVC-SEP) and Computer Aided Process-Product Engineering Center (CAPEC), as well as collaboration with Brazilian universities, like UFRJ, USP and PUC-RJ, and also Brazilian companies, potentialized the development of specific models and correlations for domestic oils.

2. History

Process simulation in Petrobras using computers started in the late 1960s, when independent programs for vapor-liquid equilibrium and thermal properties calculations were developed and used for design and education purposes (Niederberger et al, 2005). These programs were unsuitable for large plant simulations as the results had to be transferred manually between each module. Nevertheless, a gas processing unit was designed in 1976 using this approach. In 1979 the first commercial simulator (Process) was used by process design groups. Hysim followed shortly after as the choice of most of our refineries staff.

With the main purpose of acquiring know-how, a small group of engineers started the development of PETROX, using an executive program bought from Rio de Janeiro Federal University (Castier, 1985). The academic program was thoroughly modified keeping only the basic structure. In 1993, after the implementation of the distillation module, the simulator was ready to be used by the whole company (Niederberger et al, 2005).

Due to a strong research investment on exploration and production, national crude feed has been continuously increased by Petrobras' refineries, hence, considerable changes in the feed stocks composition and characteristics have been taking place. Brazilian crude oils use to be much heavier than the average crude imported, therefore, new correlations had to be developed in order to best simulate oil blends properties and processing units. Furthermore, PETROX has been an important tool which includes these original simulation models of specific oil blends in Brazil.

Development of proprietary conversion process started with the FCC technology transfer from Pullman Kellog, including a simulation tool for this process. At the same time, some models for Delayed Coked yields and Hydrotreating were developed at Petrobras' Development & Research Center.

A brief timeline of PETROX development is as follows:

- 1987: Acquisition of PSPE – Simulation Program for Stationary Processes from Rio de Janeiro Federal University (COPPE/UFRJ)
- 1989: Starting of PETROX development.

- 1993: Release of 1.0 version (DOS)
- 1999: Release of 2.0 version (Windows)
- april/2008 - Version 2.9
- **september/2008 - Version 3.0**

3. Structure

PETROX is a process simulator with a graphical user interface written in C++, C and C#.NET and a Simulation Engine written in C and FORTRAN. Figure 1, as follow, shows this structure.

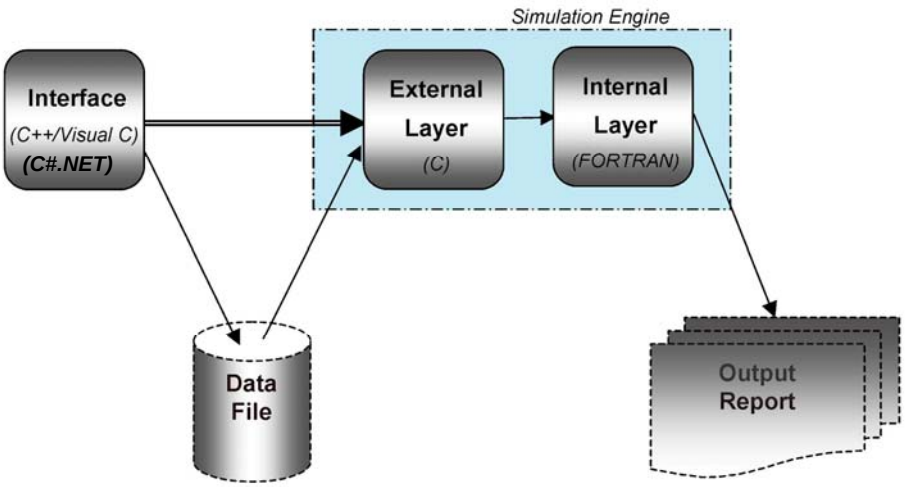


Figure 1 – PETROX Structure

The infrastructure of CENPES also has a virtualization software which allows any PETROX user runs simulations from all machines connected to the intranet.

4. Features

PETROX has several thermodynamic models and property methods required for simulating crude refining, gas processing, offshore production facilities and basic petrochemical units.
The process flow sheet visualization became available in its 3.0 version, released in September 2008, but there are some special features that make PETROX unique:

- **Crude Oils Data Bank.** It contains information on 378 crude oils. The data includes complete TBP and API gravity curves, light ends contents and composition, viscosity data for the crudes, distillates and residues, crude oil sulfur content, among others. Most of these properties were measured in our laboratories.

- **Conversion Processes Models.** Proprietary models for fluid catalytic cracking, delayed coking, mild thermal cracking, hydrotreating/hydrocracking and sulfur recovery units (SRU) are available.
- **HTRI's Rigorous Heat Exchanger Models.** The rigorous models developed by HTRI for shell and tube, air cooler, plate and frame and economizer heat exchangers are available in PETROX.
- **Integration to other process design and process analysis tools.** PETROX is capable of exchanging data with software from FRI and CPI, as well as the system COMOS® and other proprietary refining process design tools.
- **Unit operations.** Some of the available unit operations are: Flash drum, Rigorous Distillation (three different models: Inside-out, Sour-water(Wilson, 1978) and Newton), Shortcut distillation column, Mixer, Splitter, Separator, Valve, Pump, Turbine, Reciprocating compressor (uses adiabatic paths and efficiencies, with or without aftercoolers), Centrifugal compressor (uses the polytropic analysis made by Schultz (1962), with or without aftercoolers), Turboexpander, Depressurizer, Simplified heat exchanger, LNG heat exchanger, Rigorous heat exchanger (uses HTRI (2004) program to simulate equipment behavior), Feedback controller, Reactors (Shift, Stoichiometric, Gibbs and Hydrotreater), Material recycles (uses methods of successive substitution, Wegstein or dominant eigenvalue) and Calculator.
- **Proprietary correlations for Brazilian heavy oils.** As the major part of the oil refined in our refineries is domestic, and as the major part of the Brazilian oils are heavy, it was necessary to adapt literature methods for a better description of heavy oil fractions phase equilibrium and transport properties.

5. Applications

PETROX has been successfully used in revamps and design of new refining units such as atmospheric and vacuum distillation, fluid catalytic cracking, delayed coking, hydrotreating / hydrocracking, natural gas processing, basic petrochemicals, offshore production facilities and treating units like sour water treatment and sulfur recovery units. It is also used for product quality prediction and other process analysis tasks. Some applications are:

5.1. Process design of hydrodesulfurization unit of naphta catalytic cracked

Due to environmental rules, the allowed sulfur content of fuels is decreasing. In 2008, our first proprietary design of a hydrodesulfurization unit of catalytic cracked naphtha was delivered. This process uses hydrogen in order to reduce the sulfur content in this kind of stream without reducing the content of olefins and, as a consequence, the octane number of the final product. Figure 2 shows a simplified process flowchart.

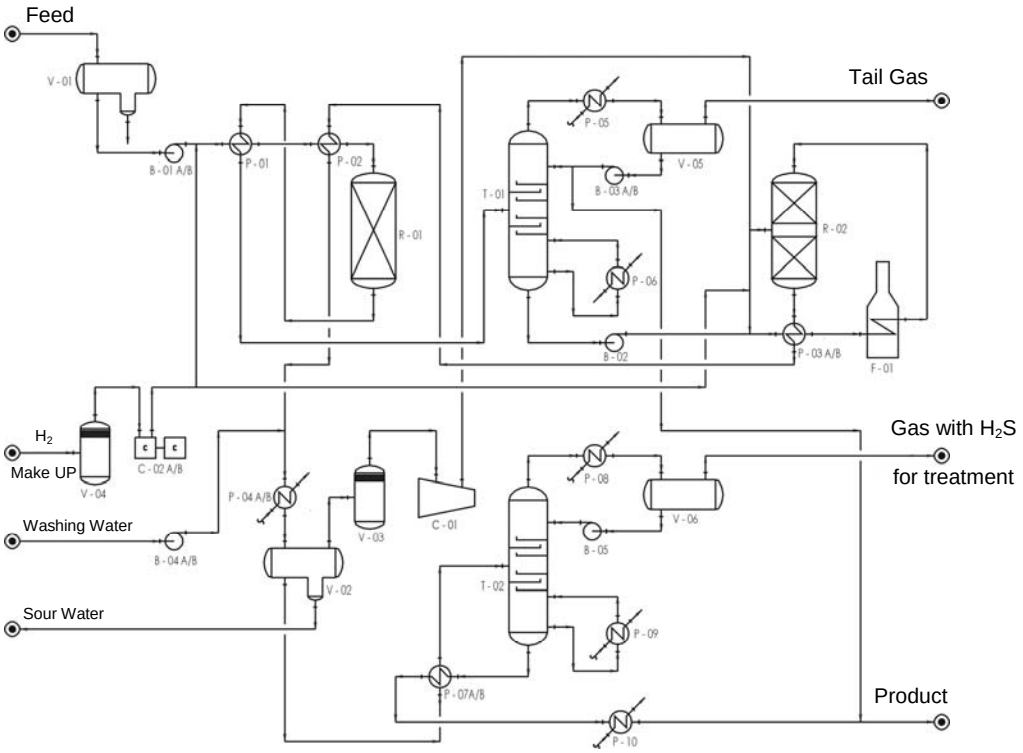


Figure 2 – Flowchart

5.2. Off-line Optimization

The sulfur recovery in industrial units is enhanced by the removal of intermediate product, by decreasing temperatures in the reactors and the number of reactors. Therefore, to optimize the performance, the design of a sulfur plant must take into account all these variables. The Sulfur Recovery Unit (SRU) is one of the Conversion Process Models available in PETROX. A thermodynamic model, based on the equilibrium of the reaction, was developed and successfully applied to the Modified Claus Process.

As environmental laws regulating the emission of sulfur are becoming more stringent, this process is increasingly important in Petrobras' refineries. Some of them use the simulator to evaluate critical data and optimize operating conditions, helping to reduce utilities comsuption and maximizing the sulfur removal.

5.3. On-line Optimization

Another Conversion Process Model is the Fluid Catalytic Cracking (FCC). It is a proprietary model that consolidates all existing knowledge in the company from the FCC area, incorporating technological innovations such as cracking of residues and others.

Currently there is an ongoing project that intends to integrate this model with a mathematical solver, to make a real-time optimization (RTO) in a Fluid Catalytic Cracking unit that uses atmospheric residue as feed (RFCC). It is expected to achieve a good financial return.

6. Conclusions

Actually, PETROX has shown to be an efficient computational tool to simulate several complex plants, as well as to aid process engineers in operating, controlling and designing process units. Some important features of commercial simulators have been included in its code in order to make it useful to the Engineering environment in Petrobras. Besides, it has a multi-disciplinary team that provides support and develops state-of-the-art applications required by users, based on either Petrobras' proprietary technologies which could not be feasible with commercial simulators or would be much more difficult and complex.
It is anticipated that real time tools will be very important to determine the best operating conditions in Petrobras refineries.
Finally, development of an in-house process simulator, besides being a competitive advantage, is strategic to protect and disseminate the technological innovations in the whole company.

References

Castier, M., 1985, Desenvolvimento de um programa executivo para a simulação de processos químicos, M. Sc. Dissertation Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Química, 239 p.

HTRI (2004), Xist User's Manual

Niederberger, J., Zech, I.A., da Silva, J.A., Mizutani, F.T., Aires, J.S.S, 2005, Petrox – Petrobras' Process Simulator, in proceedings of the 2nd Mercosur Congress on Chemical Engineering / 4th Mercosur Congress on Process Systems Engineering, Rio de Janeiro, Brazil.

Schultz, J. M. (1962), The Polytropic Analysis of Centrifugal Compressors. Transaction of the ASME Journal of Engineeringfor Power, 69-82, January. Errata 222, April 1962.

Wilson, G, M., Grantz, R. G. (1978), A new Correlation of NH3, CO2 and H2S Volatility Data from Aqueous SourWater Systems. API Publication 955, March.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A bi-level decomposition methodology for scheduling batch chemical production facilities

Georgios M. Kopanos[a], Luis Puigjaner[a] and Michael C. Georgiadis[b]

[a]*Universitat Politècnica de Catalunya - ETSEIB, Diagonal 647, Barcelona, 08028, Spain*
[b]*University of Western Macedonia, Department of Engineering Informatics & Telecommunications, Karamanli & Lygeris, Kozani, 50100, Greece*

Abstract

This work presents a two-layered decomposition methodology aimed at the resolution of difficult batch scheduling problems arising in chemical production facilities. The method relies on partial relaxation and aggregation of a detailed discrete time-indexed scheduling model, effectively transforming it into a computationally tractable approximate scheduling problem. The output of the approximate model forms the basis of an order-based decomposition scheme, within which scheduling sub-problems are resolved with the aid of the detailed scheduling model. Computational results are shown for several instances of the well-known "Westenberger-Kallrath" batch scheduling benchmark, which represents a typical multi-purpose multi-product batch chemical process.

Keywords: scheduling, decomposition, batch plants, MILP.

1. Introduction

The batch chemical production scheduling problem, consists in determining the optimal allocation of a number of resources (processing units, raw materials, utilities, manpower, etc.) over time to a number of processing tasks transforming raw materials into a number of desired final products. The extent of each processing task (total amount of materials consumed/produced by the task) is constrained by minimum and maximum batch sizes. Interactions between resources and processing tasks are discrete, i.e. occurring at fixed time intervals relative to the tasks' starting times. The problem is NP-hard in the strong sense, and even simplified versions of it are NP-hard too (Burkard et al., 2000). Accordingly, a number of approaches have been developed in order to maintain computational tractability in systems of industrial relevance; including rigorous and heuristic time-based decomposition of detailed scheduling models, mixed-integer linear programming, heuristics, local search, methods relating to project scheduling, and combined batching/scheduling techniques. The later have been shown to yield good solutions to realistic short-term production scheduling problems with reduced computational cost.

2. Motivation: The ''Westenberger-Kallrath'' process

The "Westenberger-Kallrath" (WK) benchmark represents a typical multi-purpose, multi-product batch chemical process. A thorough description of the process can be found in the literature (Kallrath, 2002).

The WK example consists of 17 distinct processing tasks (T1-T17) that can take place in 9 multi-purpose units, and 19 material states (S1-S19) with 5 desired products (S15-S19). The process features convergent, divergent, and cyclic production flows, constituting a sophisticated scheduling problem. Fig.1 depicts the State-Task-Network (STN) representation (Kondili et al., 1993) of the WK process flow.

The process involves a number of perishable products, for which both initial and maximum inventory levels are zero. It is worth mentioning that two main issues complicating scheduling considerations in the WK process are: (i) the existence of one processing task (T2) with flexible output proportions, and (ii) the need for sequence-independent batch setup times. For task T2, the first output state (S3) can be produced in amounts ranging from 20 to 70% of the task's batch size. Consequently, the second output state (S4) can be produced in amounts ranging from 80 to 30% of the task's batch size.

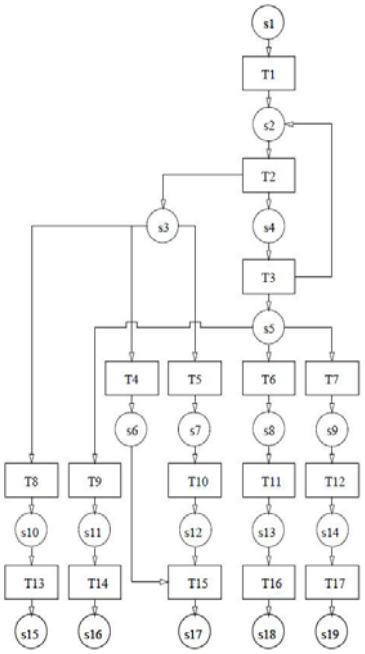


Figure 1. WK STN representation

For each processing unit, cleaning must be performed whenever production switches between two different task types. For a given set of final product demands, the scheduling objective is the minimisation of makespan subject to a number of operational constraints.

3. Detailed problem formulation

The detailed scheduling problem (DSP) formulation is based on the STN process representation. Let S denote the set of material states in the process (raw materials, intermediates, final products), U the set of available processing units, and I the set of processing tasks. For each material state $s \in S$, there is an initial inventory R_s^0, a maximum inventory level R_s^{max} and an external demand d_s. For each processing task $i \in I$, there exist minimum and maximum allowable batch sizes; B_i^{min} and B_i^{max}, respectively. For each processing unit $u \in U$, there is a set of production tasks $I_u \subset I$ that can be executed in the unit with integer processing times p_{iu}. Finally, for each material state $s \in S$, there is a set of consumption tasks I_s^c with consumption coefficients $\alpha_{is} \leq 0$, and a set of production tasks I_s^p with production coefficients $\alpha_{is} \geq 0$. These coefficients are expressed as fractions of the total batch size of task i.

Let T denote the total scheduling horizon, B_{iut} the batch size of task i starting in unit u at time $t \subset T$, and R_{st} the amount of state s at time t. Binary variable x_{iut} equals to one if task i starts in unit u at time t; otherwise, it is set to zero. To continue with, let $I^F \subset I$ denote the set of processing tasks with variable production coefficients. For each state s and processing task $i \in (I^F \cup I_s^p)$, there exist minimum and maximum production coefficients; α_{is}^{min} and α_{is}^{max}, respectively. Finally, b_{isut} denotes the amount of state s processed by task i in unit u at time t. It follows the DSP mathematical model.

$$B_i^{min}\, x_{iut} \le B_{iut} \le B_i^{max}\, x_{iut} \qquad \forall u \in U, i \in I_u, t \in T \tag{1}$$

$$\alpha_{is}^{min}\, B_{iut} \le b_{isut} \le \alpha_{is}^{max}\, B_{iut} \qquad \forall u \in U, s \in S, i \in (I_u \cup I^f \cup I_s^p), t \in T \tag{2}$$

$$\sum_{s \in I_s^p} b_{isut} = B_{iut} \qquad \forall u \in U, i \in I_u, t \in T \tag{3}$$

$$\begin{aligned} R_{st} = R_{st-1} + \sum_u \sum_{i \in (I_u \cup I_s^c)} \alpha_{is} B_{iut} + \sum_u \sum_{i \in (I_u \cup I_s^p \cap I^f)} \alpha_{is} B_{iut-p_{iu}} \\ + \sum_u \sum_{i \in (I_u \cup I_s^p \cup I^f)} b_{isut-p_{iu}} - d_s \qquad \forall s \in S, t \in T \end{aligned} \tag{4}$$

$$R_{st} \le R_s^{max} \qquad \forall s \in S, t \in T \tag{5}$$

$$\sum_{i \in I_u} \sum_{t'=t-p_{iu}+1}^{t'=t} x_{iut} \le 1 \qquad \forall u \in U, t \in T \tag{6}$$

$$\sum_{i' \in I_u, i' \ne i} \sum_{t'=t+p_{iu}}^{t'=t+p_{iu}+c_{iu}-1} x_{iut} \le \lceil \frac{p_{iu} + c_{i,u}}{\min(p_{i'u})} \rceil (1 - x_{iut}) \qquad \forall u \in U, i \in I_u, t \in T \tag{7}$$

$$t \sum_{i \in I_u} x_{iut}\, (p_{iu} + c_{i,u}) \le C_{max} \qquad \forall u \in U \tag{8}$$

Batch size constraints are expressed by eqs.(1)-(3). Flexible production recipes are taken into account by incorporating eqs.(2)-(3). Eqs.(4)-(5) correspond to material balance constraints. Eq.(6) forbids any two tasks to be performed in the same unit concurrently, and eq.(7) enforces unit unavailability during changeovers from any task *i* to a different task *i'*. Eq.(8) is only related to the makespan (C_{max}) objective.

4. Approximate scheduling model

The approximate scheduling model is derived from the detailed scheduling formulation (DSP) with the following modifications. First, alternative processing units for the same production task are aggregated based on a mean processing time. Second, the computationally expensive unit allocation constraints are partially relaxed. The functional form of all other constraints is retained. In effect, the suggested approximate model concentrates on ordering the production steps within the process relative to each other, without explicitly considering scheduling details at the unit level. Afterwards, let U_i denote the set of alternative processing units for task *i*, and $\overline{p}_i$ its mean processing time. Moreover, let z_{it} denote the number of batches for task *i* starting at time *t*, Q_{it} the cumulative batch size of task *i* starting at time *t*, and q_{ist} the cumulative amount of state *s* being processed by task *i* at time *t*. Finally, let *I'* denote any non-empty set of tasks that can be performed in at least one common unit. The approximate model constraints are summarised below.

$$B_i^{min} z_{it} \le Q_{it} \le B_i^{max} z_{it} \qquad \forall i \in I, t \in T \tag{9}$$

$$\alpha_{is}^{min} Q_{it} \le q_{ist} \le \alpha_{is}^{max} Q_{it} \qquad \forall s \in S, i \in (I \cup I^f \cup I_s^p), t \in T \tag{10}$$

$$\sum_{s \in I_z^p} q_{ist} = Q_{it} \qquad \forall i \in I, t \in T \tag{11}$$

$$R_{st} = R_{st-1} + \sum_{i \in I_s^c} \alpha_{is} Q_{it} + \sum_{i \in (I_s^p \cap I^f)} \alpha_{is} Q_{it-\bar{p}_i} + \sum_{i \in (I_s^p \cup I^f)} q_{ist-\bar{p}_i} - d_s \qquad \forall s \in S, t \in T \tag{12}$$

$$z_{it} \leq |U_i| \qquad \forall i \in I, t \in T \tag{13}$$

Eqs.(9)-(12) are the aggregated form of eqs.(1)-(4) in the DSP formulation, and eq.(13) is the approximate form of eq.(6) in DSP. Changeovers are ignored in the approximate model, thus makespan cannot be meaningfully related to the approximate model either.

5. Decomposition

The total number of batches required to satisfy demand is known by solving ASP. This information can be utilised to split the detailed scheduling problem into a number of tractable sub-problems. Hence, if Z_k is the (constant) number of batches to be scheduled at each iteration k, then the total number of sub-problems to be solved will be at most $\sum_i \sum_t z_{it} / Z_k$. In practice, an adaptive adjustment of Z_k can be employed based on the makespan achieved at the previous sub-problem.

The iterative procedure examines the solution of ASP problem sequentially starting from the first time period. When the desired number of batches Z_k has been reached (let t^* denote the corresponding time period in ASP), model DSP is solved by adding the following cuts:

$$\sum_{u \in U_i} \sum_t x_{iut} \geq \sum_{t' \leq t^*} z_{it'} \qquad \forall i \in I \tag{14}$$

$$\sum_{u \in U_i} \sum_t B_{iut} \geq \sum_{t' \leq t^*} Q_{it'} \qquad \forall i \in I \tag{15}$$

Let us refer to this sub-problem as DSP-k. The solution to DSP-k will be used in the next iteration in two ways: (i) the makespan achieved in DSP-k will provide a lower bound for the makespan of DSP-*(k+1)*, and (ii) the timings of scheduled batches from DSP-k will be fixed and provided as input to DSP-*(k+1)*. In practice, the timings obtained by DSP-k can be fixed only partially, i.e. with the exception of the last batch of each task i scheduled in DSP-k, so that feasibility of next sub-problem DSP-*(k+1)* is facilitated. For each sub-problem *DSP-k*, the choice of an appropriate upper bound on makespan is critical in the computational performance of the algorithm. In this study, a greedy heuristic has been employed, which schedules all batches in Z_k with two simple dispatching rules: (i) batches starting earlier in ASP are scheduled first, and, (ii) each batch is scheduled on the available unit with the earliest completion time.

6. Experimental results

Two sets of computational experiments are considered; one ignoring and the other considering cleaning times. Consequently, the solutions of 22 problem instances of the WK example are compared with publicly available results in order to shed light on the advantages of the proposed approach.

Comparisons are attempted with the Time-Grid Heuristic (TGH) methodology (Blömer and Günther, 1998; Blömer and Günther, 2000) and the Batching/Batch Scheduling (B+BS) approach (Neumann et al., 2002). TGH employs relax-and-fix as well as dive-and-fix heuristics for the resolution of the detailed scheduling model (Wolsey, 1998). B+BS methodology employs a simple batching model in combination with resource-constrained project scheduling algorithms. ASP is used here as an acronym for our combined bi-level approach.

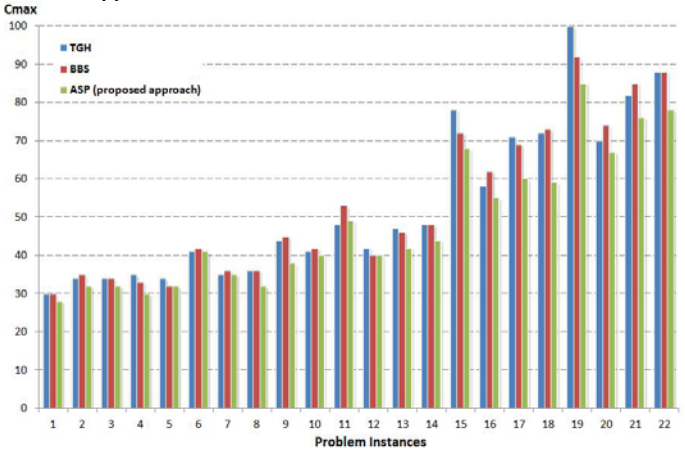


Figure 2. Experimental results for the case without cleaning times (Cmax in hours).

For the computational experiments without cleaning times (see Fig.2), equal or better solutions than TGH are obtained in 21 out of 22 cases, with an average makespan improvement of 8%. Compared to B+BS, ASP finds equal or better makespans in all 22 cases, with an average improvement of 8.16%. For the computational experiments with cleaning times (see Fig.3), ASP manages to find equal or better solutions than TGH in all cases, with an average improvement of 15.5%. Compared to B+BS, equal or better solutions are obtained in 20 cases with an average improvement of 9.3%.

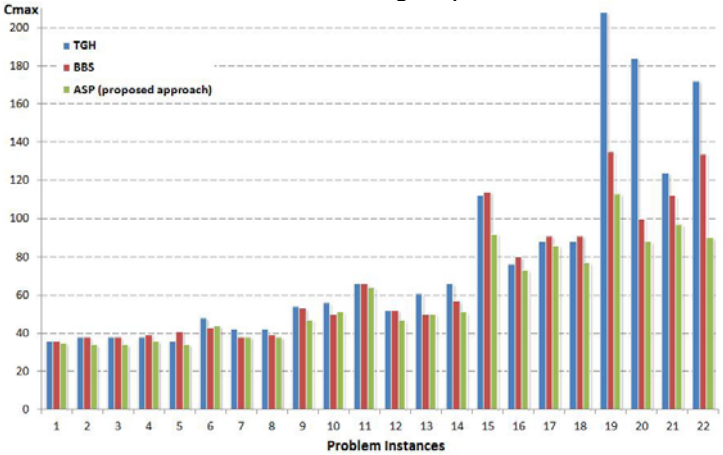


Figure 3. Experimental results for the case with cleaning times (Cmax in hours).

The results indicate that ASP manages to find good quality solutions in reasonable computational time (see Table 1). Problem ASP and all DSP sub-problems were

resolved with the aid of CPLEX 6.6 solver via a GAMS (Brooke et al. 1998) interface on a single 400 MHz Unix processor. The number of DSP sub-problems was determined dynamically for each problem instance, so that at least 20 batches were scheduled in every DSP iteration. For completeness, it is noted that TGH results were obtained on a single 266 MHz PC and B+BS results were obtained on an 800 MHz PC.

Table 1. CPU times (in seconds) for the cases with (a) and without (b) cleaning times.

	PROBLEM INSTANCES										
METHODS	1	2	3	4	5	6	7	8	9	10	11
TGH	1,100	2,247	2,487	1,550	1,778	3,600	2,587	3,123	3,600	3,600	3,600
BBS	3	13	17	60	60	60	60	60	60	60	60
ASP[a]	24	41	29	27	31	44	38	65	73	64	191
ASP[b]	44	66	111	103	116	133	108	50	59	97	64
	PROBLEM INSTANCES										
METHODS	12	13	14	15	16	17	18	19	20	21	22
TGH	3,600	3,600	3,600	3,600	3,600	3,600	3,600	5,152	3,600	3,600	3,600
BBS	60	60	60	60	60	60	60	60	60	60	60
ASP[a]	182	167	197	234	247	271	232	265	228	263	211
ASP[b]	97	89	179	107	110	126	219	291	279	284	261

7. Final considerations

A decomposition approach has been developed for the resolution of large-size batch process scheduling problems. The method has been tested on several published instances of a benchmark batch process scheduling problem with good results; particularly, for the larger problem instances. Comparisons between the decomposition approach and the full scheduling model in terms of quality solution and CPU effort will show the effectiveness of the proposed approach. To conclude, it is pointed out that as it is not certain that the minimum number of batches corresponds to a schedule of minimum length, the theoretical determination of a performance guarantee for the decomposition approach presents a challenging direction for future research.

Acknowledgments

Financial support received from the Spanish Ministry of Education and Innovation (FPU grant) is fully appreciated.

References

F. Blömer, H.O. Günther, 1998, Scheduling of a multi-product batch process in the chemical industry, Computers in Industry, 36, 245-259

F. Blömer, H.O. Günther, 2000, LP-based heuristics for scheduling chemical batch processes, International Journal of Production Research, 35, 1029-1052

A. Brooke, D. Kendrick, A. Meeraus, R. Raman, 1998, GAMS: A User's Guide, GAMS Development Corporation

R.E. Burkard, M. Hujter, B. Klinz, R. Rudolf, M.T. Wennink, 2000, A process scheduling problem arising from chemical production planning, Optimization Methods & Software, 10, 175-196

GAMS/Cplex 6.6, 2000, User Notes, GAMS Development Corporation

J. Kallrath, 2002, Planning and scheduling in the process industry, OR Spectrum, 24, 219-250

E. Kondili, C.C. Pantelides, R.W.H. Sargent, 1993, A general algorithm for short-term scheduling of batch operations - I. MILP formulation, Computers & Chemical Engineering, 17, 211-227

K. Neumann, C. Schwindt, N.Trautmann, 2002, Advanced production scheduling for batch plants in process industries, OR Spectrum, 24, 251-279

L.A. Wolsey, 1998, Integer programming, Wiley Interscience

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Generating Portable, High-level Process Models and Stand-Alone, Special-Purpose Simulators

Heinz A Preisig, Tore Haug-Warberg

Department of Chemical Engineering, NTNU, Trondheim, Norway

Abstract

The concepts behind a computer-aided modelling environment are being presented.

Keywords: modelling simulation, CAD, software

1. The Scene

Models are a core commodity in process systems engineering. If it is design or any operation activity, models are being used in one or the other form. The models being used are tailored for the application, because it is the performance of the application of the model that determines the measure of quality. Different people will have a different view on what constitutes a model. Thus the term *mode*l is used in many different contexts such as a physical, scaled replica, a set of equations or a computed trajectory. The PSE community most commonly uses the term *model* for a set of equations representing the behaviour of the modelled physical-chemical-biological system and not any derived products, such as computer code, which is also the view we take here.
Traditionally flowsheeting packages represent the state-of-the-art in terms of the common practice. Tools to construct models are limited to providing 'pick, insert and combine" library models. The linking of the sub-models may be done graphically using fixed connector definitions. The commercial side of the business has taken an approach that appears monolithic from the outside. The interior of these systems is increasingly protected in terms of accessibility and providing a view on the "inside". Whilst the vision of tools generating models based on fundamental principle has been around as long as the flowsheeting effort, the realization has been inhibited successfully based on either commercial arguments or it is seen as to large of a task, too abstract and too complex. Academic efforts approaching the problem by defining languages for capturing the information have failed to make a curb. Reviews on the subject are available in abundance (Tu:2006, Cameron:2008). Also the software providers do have little interest in supporting the exchange of models and it is only the pressure excerpted by industry that made the CAPE-OPEN effort to become a partial success.

2. The View

What one would want is a graphical representation of the process, possibly providing different modes on how to look at it. The components of these representations should give access to the model that represents its behaviour including all the assumptions that have been made to get there. Editing the structure should be supported so as to handle complexity and multiplicity of possible underlying behaviours. Thus it should allow defining different models for the same plant. The definition of the models should be built on a theoretical framework that is as generic as possible and enforces the basic concepts. Thus for first-principle rules should be implemented as hard facts. The repre-

sentation should be based on what one considers canonical in the respective sciences, in particular thermodynamics, geometry and kinetics. The fundamental concepts should be handled analytically, thus the basic concepts like energy, entropy functions and their derivatives should be available to any order analytically. Engines should exist that generate code for structure-tailored solvers, for any of the user's favourite solver, for the generation of the stand-alone, special-purpose solvers; an engine that ensures completeness and structural integrity of the models on the background of the theory; an engine that handles commonly made assumptions and eliminates the consequent mathematical problems in the model; an engine that generates stand-alone models, which can be wrapped and match an appropriate solver environment. Engines/software components that are connected by pipelines enabling the replacement of any of the software components and the need for servers should not be a must but one should have the option of inheriting all the required information down to the resulting application.
The thinking thus goes towards a kind of factory in contrast to an integrated computing environment.

3. The Modelling Process

3.1. The Model

The process of modelling a plant can be captured in Fig 1: The first domain, on the top left, is to map the process into a model, which requires two major operations: the first being an abstraction, breaking the overall process down into a set of interacting control volumes and a second being the description of each control volume and each interaction. Since one may have more than one alternative in terms of theories for the individual parts, one may be able to generate several different alternative representations using a set of combinations of available theories.

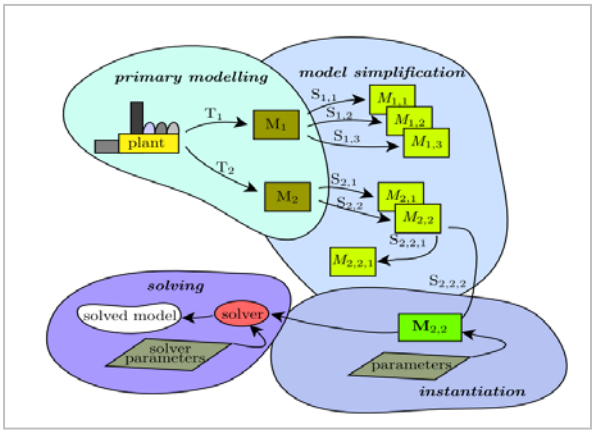


Figure 1: *Modelling overview: three major domains 1. primary modelling: maps the world into a mathematical object using theory T 2. model simplification: simplifies to match use of model using simplification S 3. model instantiation: define parameters characterizing the specific plant and solve the defined problem using the problem-specific method settings*

3.1.1. Example 1:

A distillation can be seen as a tower of dynamic flashes, where each flash consists of two phases enclose in a fixed volume. Three flashes are special, the feed splitting the column into a rectifying and stripping section, the boiler, where energy is added in form of heat and the condenser, where the heat is again removed. The mathematical description of a generic flash is shown in Figure 2

3.2. Simplifications

The second domain, on the right, is a set of simplifications and approximations. This may be done recursively, thus a simplification may be followed by another one, etc. This results in the possible generation of several models all of which represent the same plant but with an emphasis on different characteristics.

3.3. Instantiation

In the third domain the process model is instantiated with parameters making it complete and ready for simulation or other mathematical problem solvers. The transition to the solver incorporates also adding information that is specific to the solver being used.

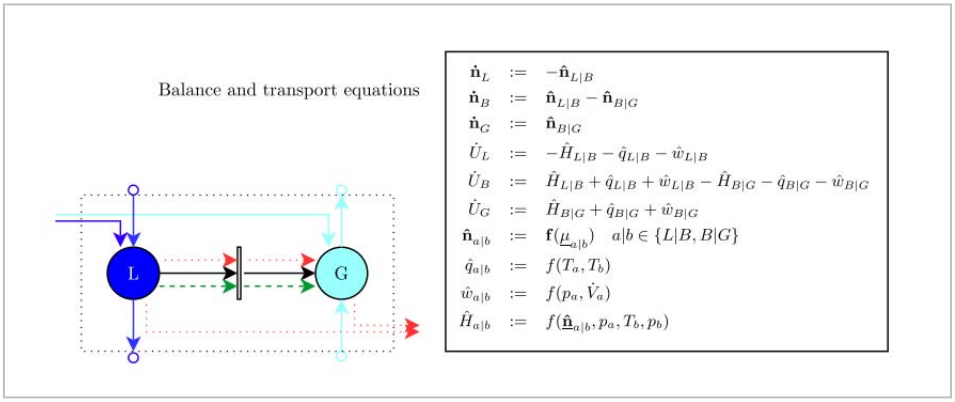


Figure 2: *A possible abstraction of a tray consisting of a lump for each phase communicating heat, work and component mass through the three connections. Heat may be exchanged with the environment and material may enter each of the lumps.*

3.4. Solving

The fourth domain solves the mathematical problem specified. In the figure a simulation is implied, which is indeed underlying many problems, including optimisation.

4. The Structures

The *abstraction* results in a network of capacities and connections, which can be very large depending on the granularity of the model and the size of the plant being modelled. This complexity can only be handled by overlaying a hierarchy. Mathematically seen a hierarchical graph is generated in which the arcs representing the connections are defined between leave nodes, with the leave nodes being the primitive capacities. Defining a graph as a set of nodes and arcs with the nodes being graphs ***G*** on the next lower level connected by a set of arcs, and defining the bottom out rule that on level 0 is a simple node ***N***, one has a hierarchical graph.

The connections on any level represent connections between pairs of primitive nodes each of which belongs to two parts of the tree on the respective level.

What one requires thus is a way of generating and manipulating a graph consisting of nodes and arcs and a tree, thus a *hierarchical graph tool*. The graph manipulation requires a number of operations: add, delete and move node, add arc, connect two nodes, delete and move an end of an arc to a new node. In order to support the hierarchy a number of operations are required such as zoom into a graph, group nodes/graphs to form a new graph (obviously including the connections), and the opposite, namely exploding a graph. These operations can be done on an abstract object that generates a set of identifiers. There is no need to make the objects anything special. The actual interpretation of the graphs' components can simply be identified using the identifiers by the graph tool. Thus this component can be designed completely independently.

If one chooses for a graphical handling of the graph generating and manipulation, the graph tool need only to generate the events for accessing the identifier-handling module.

One could also choose to generate a language for the definition of the hierarchical graph. This then would have some ways of defining the hierarchy and the underlying graph with syntax implementing adding nodes and arcs, and refining commands. In terms of algebra, one requires a tree and an incidence list for the description of the two key components, namely the tree and the underlying graph.
Depending on the application, the nodes and the arcs may be consistently typed. For example in Figure 2 we clearly have different types of connections namely mass, heat and work. This can easily be implemented as attributes to the graphs, nodes and arcs.
The individual components in the underlying graph, namely the nodes and the arcs must be given an interpretation. The *theory* for their representation can be generated separately (Hackenberg:2005). So in the chemical engineering domain one will have nodes representing physical-chemical-biological systems being of the class of 0, 1, 2, 3-D distributed systems, but it may also be an information processing system, which is used for the representation of controllers, for example. The *theory* for each component can be generated separately. The application then is only choosing the combination of *theories* appropriate for hers/his application. Using the state-concept, the *theories* for physical-chemical-biological systems implements the respective basic laws, in particular the conservation principles being the dynamic part of the description.

4.1.1. Example 2

These include the graph in typed form, explicitly the typed incidence matrices of the graph, which provide the hard link to the graph representation. Next the introduced flows and reaction/transposition parts must be extended applying the transport phenomena concepts. Transport is in turn driven by the intensive properties being the conjugates to the various fields: chemical potential, temperature and pressure, to mention the main ones. Latter are derived from the thermodynamic part of the model including a material model in the form of Helmholtz surfaces as a function of the canonical variables. Legendre transformations provide the necessary connection to the basic state variables being the conserved quantities. Providing the ability to differentiate the thermodynamic models provides access to any of the required variables.

$$
\begin{aligned}
U &:= A + T\,S \\
A &:= A(T, V, \underline{\mathbf{n}}) \\
\underline{\mu} &:= \frac{\partial A}{\partial \underline{\mathbf{n}}} \\
p &:= \frac{\partial A}{\partial V} \\
H &:= U + p\,V := U + \frac{\partial A}{\partial V}\,V \\
&:= U + \frac{\partial A}{\partial V}\,V
\end{aligned}
$$

It is absolutely feasible to *split the definition of the "theory" into different parts* and use different means for their definition, as long as the internal representation is providing the necessary link. For example it seems natural to split off the thermodynamics and possibly even the transport and kinetic part and define them as separate model components. In nearly all cases one requires a language capable of capturing equations representing specific phenomena. The equations themselves are composed of usually rather straightforward algebraic expression for which it is easy to define a simple language and generate a respective parser. The main problem really is in providing a powerful indexing procedure, which resolves the internal broadcasting at the time the equations are being instantiated. Equations consist of two expressions one for each side. An expression can be captured in an abstract syntax tree making it easy to later generate target code as desired. This applies to all sub-domains transport, thermo, etc. Each of these implementations must provide consistent equations. Those variables that are not defined by one module must be provided by one of the other modules forming an overall bipartite network. Providing alternative models defines a *super*-structure. As part of the specific model definition all choices must be made in a given super-structure reducing it

to an internally consistent set of equations[1]. All the consistency checks can be done on this level. Providing "clean" connections of the models automatically guarantees consistent process models. Thus we use the term *model synthesis* for this last stage though the sequence of the synthesis is user-controlled by either an input file or an interactive extension of the graph editor.
The internal representation of a model consists of sets of: typed nodes and arcs, typed incidence matrices, theory components representing nodes and arcs as consistent sets of equations, index maps (nodes, arcs, species, reactions, etc.) and plant-characterizing parameters. Since the equations are given as abstract syntax trees, it is straightforward to define various sets of templates for the different target languages and the available operations. These templates are then instantiated with the respective components from the abstract syntax and written to a file.

5. The Components

The Graph Tool generates and manipulates graphs. It imports the information on how its interface behaves, which enables the implementation of rules guaranteeing application-consistent graphs, such as connections must be done between primitive systems, leave nodes, in the hierarchical graph. The connection to the next layer is provided through an interface for a plug-in ***the Typing Tool*** plugs into the graph tool providing a well defined interface to perform the typing of the graph component. The typing tool imports its knowledge labelled "theory" from ***the Theory User Tool*** which provides an interface to the defined *theory* being the super-bi-partite graph of the equations representing the behaviour of the individual nodes and arcs. Here the choices reducing the super-graph into a simple equation set is performed and the index maps for the broadcasting operation in the final code are prepared. This tool may also implement model simplifications (Preisig:2009). **The Instantiation Tool** is used to instantiate the model's parameters, conditions and inputs. Intrinsically this specifies the type of mathematical problem to be solved. For example knowing the initial conditions, the model parameters and the inputs, defines a simulation problem. ***The Code Generation Tool*** computes first the final index maps (Preisig:2009) and then generates the compiled code for the selected model components using a template engine (Parr:2004). Finally this code is spliced into the solver-specific interface generating the final code being written to a target file. Later operation can also be solved with a template machine. Providing different sets of templates for the primitive operations and the external functions to be embedded, provides the means to generate code for the respective different target solvers. Target language and target solvers may be split for obvious practical reasons. ***The Theory Definition Tool*** may be split into several parts each specific to any of the sub-fields mass, energy, transport, thermo, etc. (Loevfall:2008). Each must though provide a consistent model; that is a set of equations that have a well defined set of input variables from, which a set of output variables are defined all of which is consistent with the applied theory. The set of equations must be well determined, though may have a superstructure imposed. Any of those tools must perform the possible structural check to the extent possible. For example physical units in connection with any mathematical operation must be checked etc. The tool should also provide analytical derivatives of the equations it generates. So as to enable transformations as briefly touched upon in Example 2.

[1] Consistency is defined as having zero degrees of freedom once the parameters and initial conditions of a simulation problem have been defined.

6. Conclusions

The concept of a software/solver structure has been introduced and, which mirrors the outlined step-wise modelling process. The core is a communicating network of components forming a factory for generating all kind of tools, including stand-alone simulators, code for legacy solvers, simulation languages, optimisation environments etc. The resulting components provide a dense checking on structures being generated thereby essentially eliminating the structural problems one experiences using the traditional tools (Szyperski:1998). The generation of internal code enables automatic analytical differentiation of any of the expressions appearing in the theory. The making of order-of-magnitude assumptions can be implemented including the reconciliation of the introduced mathematical problems, as shown in (Preisig:2009, Preisig:2008)

The generation of such a software environment: is a truly interdisciplinary activity, with information processing being the common "glue". Essential for the decomposition of the overall problem is the definition of well defined, small, highly structured, easy digestible pieces, which utilize *very few* concepts. For example the thermo module uses essentially only 3 principle ways of combining objects.

The paper addresses six different types of subjects all lying in different domains but also all of which must be resolved in order to generate an integrated modelling environment. 1) Graphical user interface for easy handling of hierarchical graphs. 2) Template engine, a technique to generate target code recursively. 3) Small languages that are tailored to capture individual components such as mathematical expressions and equations for the representation of the theory for the components represented in the respective graphs; latter being process models (capacities and extensive quantity transfer, reactions) or expressions, representing mathematical operations, or thermodynamic models, or geometry. 4) Solvers for well defined mathematical problems, including integrators, root solvers, symbolic differentiation. 5) Theory knowledge, including physical chemistry in particular thermodynamics, but also mathematical system theory, control etc. 6) Communications, that enable and support the interaction of the different programs using communication frameworks such as .COM or .NET environments.

The concept of a graph appears in different contexts: The graphical representation of a hierarchical decomposition of the processes' control volume; the representation of an equation, an expression as an abstract syntax tree; the representation of equation sets as bi-partite graphs; the GUI-controlling automaton as a state-graph.

References

I T Cameron and G D Ingram. A survey of industrial process modelling across the product and process lifecycle. Comp & Chem Eng, 32:420–438, 2008.

J¨org Hackenberg. Computer support for theory-based modelling of process systems. PhD thesis, RWTH Aachen, Germany, 2005.

Børn Tore Løevfall. Computer Realisation of thermodynamic modles using algebraic objects. PhD thesis, NTNU, Trondheim, Norway, 2008.

Terence John Parr. Enforcing strict model-view separation in template engines. In WWW '04: Proceedings of the 13th international conference on WWW, p 224–233, New York, 2004.

H A Preisig. Three principle model reductions based on time-scale considerations. ESCAPE 18, (ISBN978-0-4444-53228-2), 2008.

Heinz A Preisig. Constructing, modifying and maintaining consistent process models. FOCAPD, 2009.

Clemens Szyperski. Component software - beyond object-oriented programming. Number ISBN 0-201-17888-5. Addison Wesley, 1998.

Hong Tu and Irven Rinhard. ForeSee – a hierarchical dynamic modeling and simulation system of complexprocesses. Comp & Chem Eng, 30:1324–1345, 2006

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Application of computer software for system identification and PID control loop tuning to a petrochemical plant

Péricles R. Barros,[a] Mario C. M. M. de Campos,[b] George Acioli Júnior,[a] Rodrigo C. d'A. de Carvalho,[b] João B. M. dos Santos, [a]

[a]*Electronic Instrumentation and Control Laboratory – UFCG, Campina Grande, Brazil*
[b]*CENPES - PETROBRAS, Rio de Janeiro, Brazil*

Abstract

PID controllers are widely used in the process control industry due to their simplicity and ease of re-tuning on-line. A software package has been developed for system identification and PID control loop tuning by a partnership between UFCG and CENPES-PETROBRAS. This system implements literature and industry standard PID tuning algorithms and developed in UFCG. Application results of tuning software tools in a petrochemical plant, at PETROBRAS' are presented in this paper

Keywords: System Identification, PID Tuning, Relay Experiments, Industrial Application.

1. Introduction

The use of PID (proportional-integral-derivative) control loops is common in a large number of industrial processes (Aström and Hägglund, 2006). It is estimated that over 90% of industrial controllers are implemented based around PID algorithms, particularly at regulatory level. Its wide application has motivated the development of various PID tuning techniques and tuning software tools (Ang et al., 2005). A research effort has been made to improve system identification techniques and PID tuning using operational plant data.

The PETROBRAS (Brazilian Petroleum Company), through your Research Center - CENPES, invested in developing PID tuning software package tools during partnership with Federal University of Campina Grande (UFCG). In this manner, the academic and industry competence was combined. These software tools are named BR-TUNING.

The BR-TUNING software tool implements literature and industry standard PID tuning algorithms as well the algorithms developed at UFCG. Open and closed-loop identification techniques are implemented. In this paper, descriptions of implemented algorithms are presented. The main objective of this paper is present application results of the BR-TUNING in a natural gas processing plant (PETROBRAS' UN-RNCE in Guamaré–RN-Brazil).

This paper is organized as follows: Initially, the algorithms for identification and PID tuning implemented in BR-TUNING are presented (Sections 2, 3 and 4). In the section 5, the application results are presented and, finally, conclusions are presented in section 6.

2. Open-Loop Identification and PID Tuning

In this section, step response and relay-based open-loop identification techniques are presented. For each identification technique, a set of standard PID tuning are available in BR-TUNING.

2.1. Identification of FOPDT model from a step excitation and model-based PID Tuning

In the BR-TUNING it is implemented an identification technique of FOPDT model from open-loop step response. This technique was presented in (Wang et al., 2000). Consider the open-loop shown in Figure 1. It applies a step excitation signal (MV input) and store the PV signal. The FOPDT model is given by $\hat{G}(s) = \frac{b}{s+a} e^{-L_d s}$.

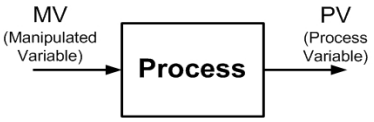


Figure 1: Open-Loop Process

The process is assumed to be in steady-state at $t = 0$ so that, $MV(t) = 0$ for $t < 0$ and zero initial conditions are assumed. For $t \geq L_d$ the process is assumed to satisfy $PV(t) = -a\int_0^t PV(\tau)d\tau + bht - bhL_d$, where h is the amplitude of step excitation.

From which a regression model can be obtained on model parameters $\{a, b, bL_d\}$. The parameters can be computed in one step using the least-squares method, and a model-based PID tuning method can be used. Several methods can be found in literature (O'DWYER, 2006). The following techniques: Ziegler-Nichols, Cohen-Cohn, Chien-Hrones-Reswick-Load-0% (CHR-Load-0%), CHR-Load-20%, CHR-Setpoint-0%, CHR-Setpoint-20%, IMC - First Order and SIMC are implemented in BR-TUNING.

2.2. Relay-based Identification and PID Tuning

The standard relay feedback experiment (see Figure 2) was introduced in (Åström and Hägglund, 1984). The theory for analyzing the plant under relay feedback was based on the well knows describing function analysis, which also contributed to make the method widely accepted. With this experiment, a single point of process frequency response, the critical point (phase angle –180°), can be estimated. Applying the relay experiment, the output, for most systems, has a limit cycle which occurs near the critical frequency. The values of critical gain (K_c) and critical frequency (ω_c) are estimated using DFT (Discrete Fourier Transform) of input-output process data (MV and PV signals).

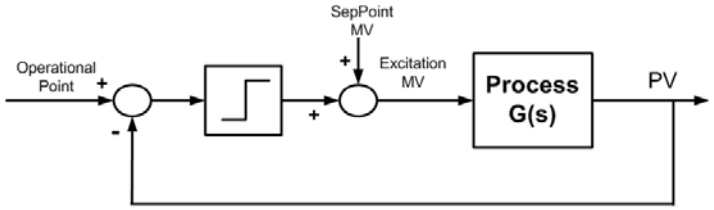


Figure 2: Relay Feedback Experiment

With the critical point, it is possible to design a PID controller. The BR-TUNING implements following techniques: Ziegler-Nichols, Clair, Tyreus-Luyben, With Some Overshoot and No Overshoot, for details about these techniques a good reference is (O'DWYER, 2006).

3. Relay-Based Closed-Loop Evaluation

In this section, modified relay experiments are presented to evaluate the closed-loop gain and phase margins. Consider the closed-loop shown in Figure 3. The unknown process transfer function is given by $G(s)$ while the controller is given by $C(s)$.

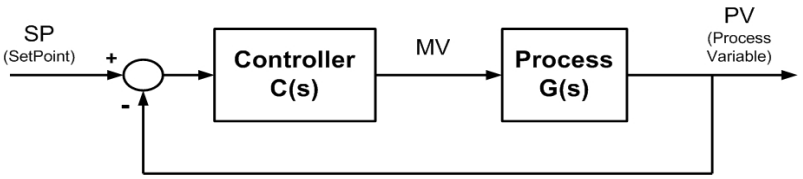


Figure 3: Closed-Loop Process

The closed-loop transfer function from setpoint signal SP to PV signal is given by

$$\frac{\mathrm{PV(s)}}{\mathrm{SP(s)}} = \frac{G(s)C(s)}{1+G(s)C(s)} = T(s)$$

The gain margin (GM) can be estimated with the relay structure, as shown in Figure 4 (See (Schei, 1992)), applied to the closed-loop T(s). Applying a relay with amplitude d to the closed-loop yields a limit cycle near the critical frequency (ω_u) of the loop transfer function. The loop gain at that frequency can be computed as

$$L(j\omega) = G(j\omega)C(j\omega) \cong \frac{m}{1-m},$$

where $m = \frac{-\pi a}{4d}$ and a is the closed-loop output amplitude.

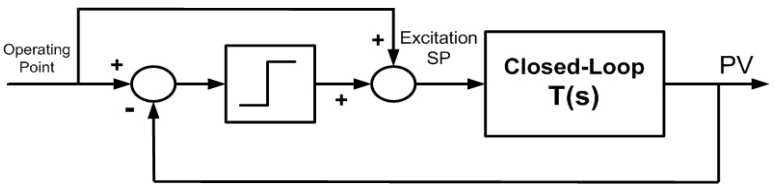


Figure 4: The Gain Margin Relay Test Structure

In BR-TUNING, a more refined algorithm to estimate closed-loop critical frequency information is implemented. This algorithm using DFT (Discrete Fourier Transform) of the SP and PV signals to define ω_u and $T(j\omega_u)$. With this, we can define $L(j\omega_u)$ as

$L(j\omega_u) = \frac{T(j\omega_u)}{1-T(j\omega_u)}$, and the gain margin is defined as $GM = \frac{1}{|L(j\omega_u)|}$.

The closed-loop phase margin can be obtained using the relay structure shown in Figure 5. This structure was introduced in (Schei, 1992) and generalized in (Arruda and Barros, 2003a).

Estimate the crossover frequency (ω_g), $T(j\omega_g)$ and $L(j\omega_g)$ in the same way using DFT. The phase margin is defined as $PM = \pi + \angle L(j\omega_g)$.

For more details about the algorithm to estimate closed-loop margins using DFT and relay tests structures can be found in (Arruda and Barros, 2003a).

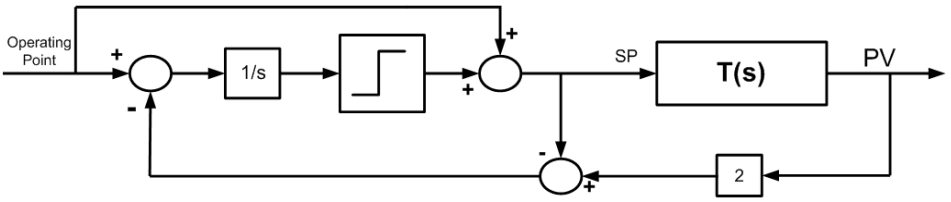


Figure 5: The Phase Margin Relay Test Structure

4. PI Controller Iterative Redesign

Consider the closed-loop shown in Figure 2 and gain and phase margins specifications given by A_m and φ_m respectively. For the desired specifications the controller must be designed to satisfy the following set of equations:

$$\angle G(j\omega_u)C(j\omega_u) = -\pi\,, \quad |G(j\omega_u)C(j\omega_u)| = \frac{1}{A_m}\,,$$

$$|G(j\omega_g)C(j\omega_g)| = 1 \text{ and } \angle G(j\omega_g)C(j\omega_g) = -\pi + \varphi_m$$

If $G(j\omega)$ is known, the PI controller design requires the solution of the set of equations for K_p, T_i, ω_u and ω_g a total of four unknowns and four equations and it is a non-linear problem. In (Arruda and Barros, 2003), an iterative algorithm is used which requires the estimation of the frequencies ω_u and ω_g and the gain and phase margins at each iteration. After each relay test the controller is updated to correct for the gain or phase margin. The scheme can be interrupted whenever a both specifications are met within tolerated limits. The design procedure, presented in details and with safety considerations, in (Arruda and Barros, 2003) is as follows: At each iteration assume an existing controller $C^0(s)$, with parameters $K_p{}^0$, $T_i{}^0$.

4.1. Gain Margin Redesign

Estimate GM and frequency ω_u, and redesign the controller for Gain Margin using

$$K_p = \frac{K_p{}^0 GM}{A_m}\,, \quad \text{so that the new controller is } C(s) = K_p\left(\frac{s+1/T_i}{s}\right).$$

4.2. Phase Margin Redesign

Estimate PM and the gain crossover frequency ω_g:

a) Compute the integral time $$T_i = \frac{\tan\left(-\pi - \varphi_m - PM + \tan^{-1}\left(\omega_g T_i{}^0\right)\right)}{\omega_g}.$$

b) Adjust the proportional gain. $$K_p = K_p{}^0 \frac{\sqrt{\left(1/T_i{}^0\right)^2 + \omega_g{}^2}}{\sqrt{\left(1/T_i\right)^2 + \omega_g{}^2}}.$$

5. The Experimental Results

In this section, application results of the BR-TUNING in PETROBRAS' UN-RNCE are presented.

5.1. Temperature Control Loop TIC_29051 – Open-Loop Example

The first loop is a temperature control named TIC_29051. An open-loop step excitation was applied. The FOPDT model estimated using BR-Tuning is $\hat{G}(s) = \frac{0.925}{93.62s+1} e^{-18s}$.

In Figure 6, the step responses of open-loop process (blue curve) and estimated model are shown. The mean-squared error computed for estimated model was 0.008. The PID controller was designed using IMC technique with filter time constant 93.6 sec.

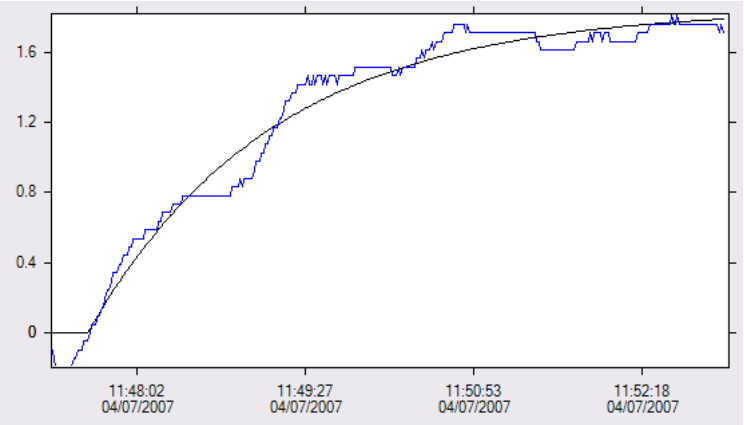


Figure 6: BR-TUNING Window – Open-Loop Identification

The designed controller is $C(s) = 0.99\left(1 + \frac{1}{1026.1s} + 8.2s\right)$. The control loop performances of TIC_29051 with the older and the new controller are shown in Figure 7.

5.2. Flow Control Loop FIC_29020 – Closed-Loop Example

The second loop is a flow control named FIC_29020. In this loop, relay tests for closed-loop gain and phase margins evaluation are performed. The estimated gain and phase margins are $GM = 1.104$ and $PM = 49.36°$. This is very aggressive from the gain and phase margins side. The gain margin specification $A_m = 2$ yields a controller gain $K_p{}^1 = 0.25$ and $T_i{}^1 = 0.05\,\text{sec}$, while a redesign for a phase margin $\varphi_m = 70°$ results in a controller with com $K_p{}^1 = 0.633$ and $T_i^1 = 0.0336\,\text{sec}$. The objective is to achieve a more stable control loop, the chosen controller was the phase margin with the gain from the gain margin:

$$C(s) = 0.25\left(1 + \frac{1}{0.0336s}\right).$$

The control loop performance for SP signal step changes of FIC_29020 with older and new controller is shown in Figure 8a and 8b respectively. Note that the response with new controller is less aggressive when compared with the response with older.

6. Conclusion

In this paper application results of the BR-TUNING in a petrochemical plant was presented. Moreover, the algorithms for identification and PID tuning implemented in BR-TUNING were reviewed. The obtained results showed the properties of the algorithms and these practical effectiveness.

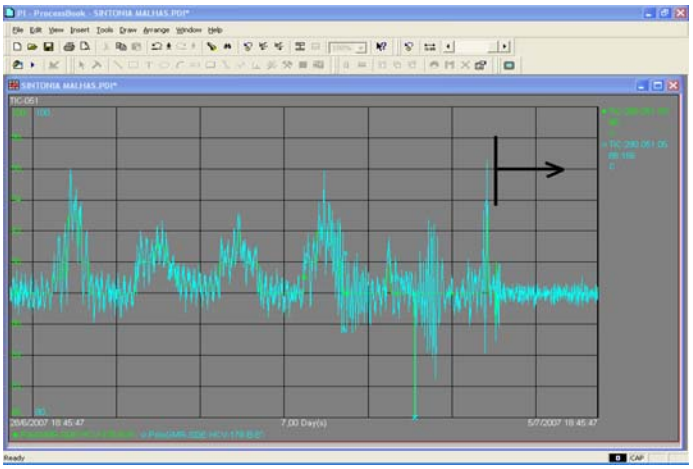


Figure 7: Supervisory Window for TIC_29051

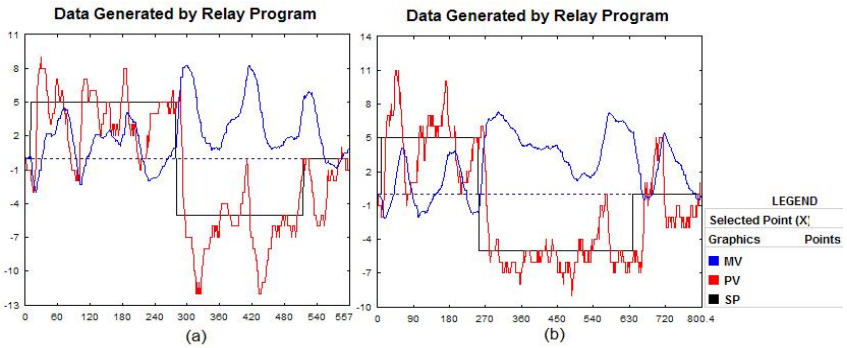


Figure 8: Step response: (a) Older controller, (b) New (redesigned) controller

References

K. H. Ang; G. Chong; Y. Li, 2005, PID Control System Analysis, Design, and Technology, IEEE Transactions on Control Systems Technology, v.13, n.04, 559-576.

G. H. M. Arruda; P. R. Barros, 2003, Relay-based gain and phase margins PI controller design, IEEE Transactions on Instrumentation and Measurement Technology, v.52, n.05, 1548-1553.

G. H. M. Arruda; P. R. Barros, 2003a, Relay-based transfer function frequency points estimation, Automatica, v.39, n.02, 309-315.

K. J. Astrom; T. Hagglund, 2006, Advanced PID Control, Instrument Society of America, Research Triangle Park, North Carolina.

K. J. Astrom; T. Hagglund, 1995, PID Controllers: Theory, Design and Tuning, 2nd ed., Instrument Society of America, Research Triangle Park, North Carolina.

A. O'dwyer, 2006, Handbook of PI and PID controller tuning rules, Imperial College Press.

T. S. Schei, 1992, A method for closed-loop automatic tuning of PID controller, Automatica, v.28, n.03, 587-591.

Q.-G. Wang; B. Hwang; X. Guo, 2000, Auto-tuning of tito decoupling controllers from step tests, ISA Transactions, v.39, 407-418.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimization of MMA Batch Polymerization Reaction by the Implementation of IDP Technique

Paulo R. A. Pereira,[a] Ana M. F. Fileti,[b]

[a]*School of Chemical Engineering, Centro Univ. das Fac. Assoc. de Ensino – FAE Largo Eng. Paulo de A. Sandeville, 15, CEP 13870-377, S. J. da Boa Vista - SP, Brazil.*
[b]*State University of Campinas, School of Chemical Engineering, CP 6066, CEP 13083-970, Campinas – SP, Brazil*

Abstract

The present work is concerned with a dynamic optimization problem involving the methyl methacrylate (MMA) free-radical batch solution polymerization reaction to produce poly-methyl methacrylate (PMMA). Ethyl acetate is used as solvent and benzoyl peroxide (BPO) as the reaction initiator. A dynamic model that represents the batch reaction process is presented. The model equations are solved in order to maximize the performance index defined as MMA conversion rate (X_m). The optimization process will indicate the trajectories for the reactor temperature (T), the cooling water flow rate (F_{cw}) and the heating power (Q) supplied to the reactor through an on/off switch system. X_m is optimized in accordance with the experimental reactor operation and safety conditions. The reaction time of 350 minutes is divided in discretized intervals. The reactor temperature and the control variables (F_{cw} and Q) are assumed piecewise constant in each time interval. To perform the optimization a program based on the Iterative Dynamic Programming (IDP) technique is successfully developed. The optimization results demonstrated that IDP technique is suitable for the optimal control of MMA-PMMA system, and probably for other batch polymerization processes with high complexity and nonlinearity.

Keywords: Batch Reactor, Dynamic Programming, Optimization, IDP, MMA.

1. Introduction

In this work, the MMA batch polymerization reaction was chosen as case-study. MMA polymerization process in batch reactors has been largely studied in the last decades. Important contributions on MMA polymerization can be found in the specialized literature (Ellis et al. (1988); Mutha et al. (1997); Soroush e Kravaris (1992); Rho et al. (1998); Jeong e Rhee (2000); Ahn et al. (1998), Chang e Liao (1999); Chae et al. (2000); Rafizadeh (2001)). By nature batch reactors are transient and are modeled by a combination of Differential and Algebraic Equations (DAEs) of distinct levels of complexity. To solve these equations important techniques have been used as Pontryagin's Maximum Principle (Zeybek et al. (2004)), Genetic Algorithms (Upreti et al. (2005)), Control Vector Parameterization (CVP) (Ekpo and Mujtaba, 2007), and IDP (Luus, 1991) chosen for this work. In the IDP approach, a piecewise constant control policy is found by employing dynamic programming in an iterative fashion. It uses a contraction of the region over which allowable values of the control are taken randomly (Palanki & Vemure, 2005). The choice of random values was proposed by Bojkov & Luus (1992) to overcome the difficulty of dealing with control vectors of high dimensionality presenting faster convergence to optimum. This strategy has been

applied in a variety of batch applications (Luus, 1993a; Luus, 1993b). Sayer et al. (2001) applied IDP to compute optimal monomer and chain transfer agent feed profiles to produce polymer with specified copolymer composition and molecular weight distribution. More recently, Sayer et al. (2006) implemented IDP strategy to optimize the off line optimal trajectories for the less reactive feed flow rate and for the reactor set-point temperature. In the present study, the IDP technique is applied to optimize X_m and obtain the optimal trajectories of T, F_{cw} and Q, for the system MMA-PMMA.

2. Mathematical Modeling

The same experimental reactor scheme initially designed by Pedrosa et al. (2001) and used by Antunes et al. (2005) is considered for the IDP off line optimization. The reaction is supposed to be carried out inside a jacket glass batch reactor (Figure 1). The feed of F_{cw} for the jacket is regulated by a control valve, and monitored in terms of inlet and outlet temperatures. The system is equipped with two electrical heaters of 0.075 and 0.045 kJ/s, respectively. The first one remains on during the reaction time while the second is hypothetically controlled by an on/off basis, which occurs respectively, when the temperature deviates –4 or +4 ºC from set-point of 63ºC (336.2 K). These conditions for T were defined by Antunes et al. (2005), according to the real reactor characteristics, and operation/safety conditions, which are considered as the basis of this study. To represent the MMA solution polymerization reaction in the batch reactor, it was chosen the principle model studied by Ray et al. (1995) as well as Seth and Gupta (1995), which considers the mass balance and moments of the live and dead polymer concentrations distributions. Chakravarthy (1997) improved this model by introducing the gel, cage and glass effects. According to Gosh et al. (1998), predictions of this model demonstrated a perfect agreement when compared to their experimental results and with those reported earlier in the literature. Once the conditions adopted in the present study for MMA polymerization reaction are close to those used by Gosh et al. (1998), the model was considered reliable and was not experimentally validated. The following DAEs for the model are presented below:

BPO initiator consumption:

$$\frac{dI}{dt} = -k_d I + R_{Ii}(t) \tag{1}$$

Mass balance to the primary radical:

$$\frac{dR}{dt} = 2fk_d I - k_i \frac{RM}{V_l} \tag{2}$$

Monomer consumption:

$$\frac{dM}{dt} = -(k_p + k_f)\frac{\lambda_o M}{V_i} - k_i \frac{RM}{V_l} - k_s S \frac{\lambda_0}{V_l} + R_{Im}(t) - R_{vm}(t) \tag{3}$$

Solvent consumption:

$$\frac{dS}{dt} = R_{Is}(t) - R_{vs}(t) \tag{4}$$

Moments of the live polymer concentrations distributions, λ_i:

$$\frac{d\lambda_0}{dt} = k_i \frac{RM}{V_l} - k_t \frac{\lambda_0^2}{V_l} \tag{5}$$

$$\frac{d\lambda_1}{dt} = k_i \frac{RM}{V_l} + k_p M \frac{\lambda_0}{V_l} - k_t \frac{\lambda_0 \lambda_1}{V_l} + \left(k_s S + k_f M\right)\frac{\left(\lambda_0 - \lambda_1\right)}{V_l} \tag{6}$$

$$\frac{d\lambda_2}{dt} = k_i \frac{RM}{V_l} + k_p M \frac{(\lambda_0 + 2\lambda_1)}{V_l} - k_t \frac{\lambda_0 \lambda_2}{V_l} + (k_s S + k_f M)\frac{(\lambda_0 - \lambda_2)}{V_l} \quad (7)$$

Moments of the dead polymer concentration distributions, μ_i:

$$\frac{d\mu_0}{dt} = \left(\frac{1}{2} ktc + k_{td}\right)\frac{\lambda_0^2}{V_l} + (k_f M + k_s S)\frac{\lambda_0}{V_l} \quad (8)$$

$$\frac{d\mu_1}{dt} = k_t \frac{\lambda_0 \lambda_1}{V_l} + (k_f M + k_s S)\frac{\lambda_1}{V_l} \quad (9)$$

$$\frac{d\mu_2}{dt} = k_{tc} \frac{\lambda_1^2}{V_l} + k_t \frac{\lambda_0 \lambda_2}{V_l} + (k_f M + k_s S)\frac{\lambda_2}{V_l} \quad (10)$$

Variables to account addition and vaporization of monomer after time t = 0, kmol/m^3:

$$\frac{d\zeta_m}{dt} = R_{lm}(t) - R_{vm}(t) \quad (11)$$

$$\frac{d\zeta_{ml}}{dt} = R_{lm}(t) \quad (12)$$

Volume of reacting mixture at time t, m^3:

$$V_l = \frac{S(MW_s)}{\rho_s} + \frac{M(MW_m)}{\rho_m} + \frac{(\zeta_m - M)(MW_m)}{\rho_p} \quad (13)$$

Energy balance in the reactor:

$$\rho\, c\, \frac{dT}{dt} = (-\Delta H_P) k_P M \lambda_0 + U\, A\, (T_j - T) \quad (14)$$

Energy balance in the jacket:

$$\rho_w c_w \frac{dT_j}{dt} = UA(T - T_j) + U_\infty A_\infty (T_\infty - T_j) + [Q - F_{cw}\rho_w c_w (T_j - T_{cw})] \quad (15)$$

Luus (2000) describes that the best way of dealing with difficult equality constraints is the use of a quadratic penalty function incorporating a shifting term. Then, to complete the model equations, X_m is used as an augmented performance index as follow:

$$I = X_m + \theta(X_m - s)^2 = \left(1 - \frac{M}{\zeta_{m1}}\right) + \theta\left[\left(1 - \frac{M}{\zeta_{m1}}\right) - s\right]^2 \quad (16)$$

where θ is a positive penalty function factor and the shifiting term is inittialy put to zero and then, after every pass q, updated by:

$$s^{q+1} = s^q - X_m \quad (17)$$

The state vector $\boldsymbol{x}$ can be generically expressed as:

$$x = [I, M, R, S, \lambda_0, \lambda_1, \lambda_2, \mu_0, \mu_1, \mu_2, \zeta_m, \zeta_{m1}, T, T_j, I_{pf}]^T \quad (18)$$

Additional equations and values for cage, gel and glass effects, besides other parameters used in this work are found in the sources: Achilias et al. (1992), Chakravarty et al. (1997); Ellis et al. (1988); Gosh et al. (1998); Soroush e Kravaris (1992).

3. Optimal Control Technique

To optimize X_m, a program in Matlab for the IDP technique implementation was developed. As all the steps involving the IDP algorithm are very well documented in the literature, they are not described in details in this work. To implement IDP, it was necessary to create a program subdivided in main routine and additional subroutines. To solve the model equations, Matlab's ode23tb function was inserted as a subroutine in

the IDP main program. It is suitable for solving stiff differential equations as in the case of MMA polymerization. The developed program permits to divide the time interval [0, *tf*] into *P* [0, 20] stages of equal time length *L*. To choose the values for the control variable $\boldsymbol{u}$ (F_{cw}) and the number of $\boldsymbol{x}$-grid points *N*, the Matlab's function *rand* (an *N*-by-*N* matrix with random entries) was used to generate random values to accelerate the convergence process (Luus, 2000). To assure the random values inside the region permitted for control (in agreement with reactor operation and safety conditions), the clipping technique was used:

For the control variable: $F_{cw} = 26.7x10^{-5}\, m^3/s \quad if \quad F_{cw} \geq 26.7x10^{-5}\, m^3/s$ *(20)*

$F_{cw} = 8x10^{-6} \quad if \quad F_{cw} \leq 0$ *(21)*

For the state variable: $T = 340.2\, K \quad if \quad T \geq 340.2\, K$ *(22)*

$T = 332.2\, K \quad if \quad T \leq 332.2\, K$ *(23)*

It is important to emphasize that the clipping rules are used only to supply the lower and upper limits for the IDP comparisons, but are not used as constrains during the integration process. The heating system was composed by two on/off electrical heaters following the clipping rules:

Both heaters on: $Q = 0.12\, kJ/s \quad if \quad T \leq 332.2K$ *(24)*

Only one heater on: $Q = 0.075\, kJ/s \quad if \quad T \geq 340.2K$ *(25)*

Figure 2 (Schröder and Mendes, 1999) illustrates an example for the optimal trajectory obtained for $\boldsymbol{x}$ at the end of the first IDP iteration. The configuration used was *P* = 10 stages (vertical lines) and **N** = 5 $\boldsymbol{x}$-grid points (black dots). The solid line from stage 1 to *P* is the trajectory for each stage obtained during the integrations for the first iteration. All the lines at the stages *P* and *P* – 1 represent the partial results obtained by the integrations using different *N* $\boldsymbol{x}$-grid points and *R* random values for **u** control variables.

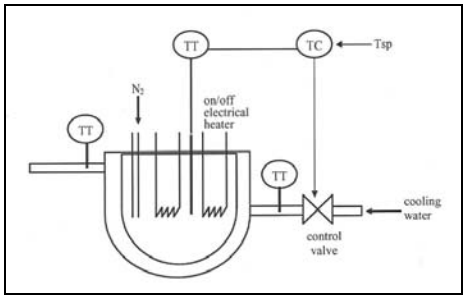


Figure 1 – Batch Reactor Scheme for MMA polymerization (Antunes, 2005)

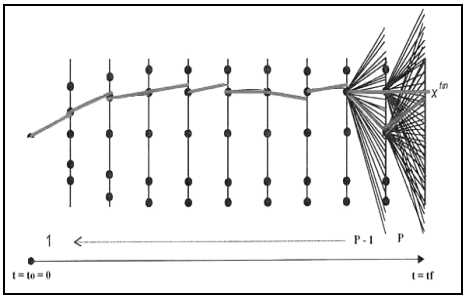


Figure 2 – Optimal Trajectory for $\boldsymbol{x}$-Grid Points in the First Iteration

4. Results

During the optimization process several IDP configurations for *N* $\boldsymbol{x}$-grid points and *R* random values for F_{cw} were tested. The best results for X_m were obtained by an IDP configuration with *P* = 20, *N* = 20 and *R* = 20 (Figure 3). The trajectories for *T*, F_{cw} and *Q*, are shown in Figures 4 to 6. The optimization result of 0.5322 for X_m is 33.05% higher than 0.4 obtained experimentally by Antunes (2005) using a fixed *T* set point of 336.2 K tracked by a PID controller. Probably this improvement on the value of X_m is due to *T* trajectory obtained by the IDP off line optimization, which suggests the use of different set points for each time interval (Figure 4), instead of a fixed set point.

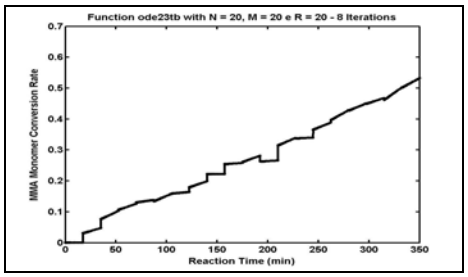


Figure 3 – MMA Conversion Rate

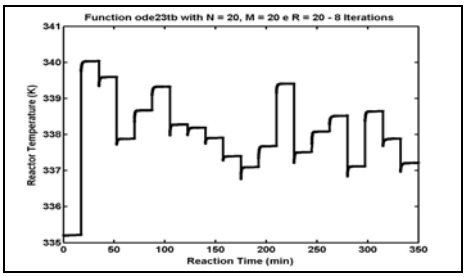


Figure 4 – Reactor Temperature

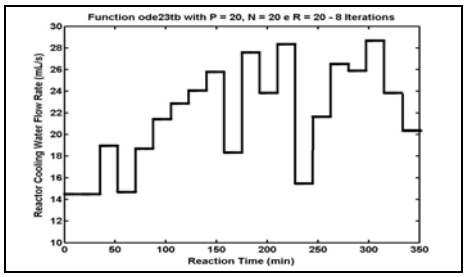


Figure 5 – Cooling Water Flow Rate

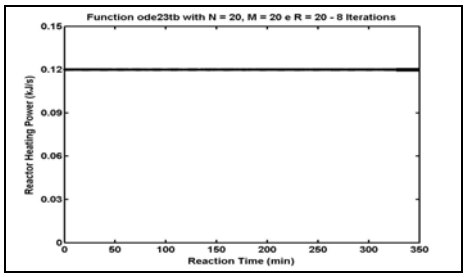


Figure 6 – Reactor Heating Power

The behavior of T, F_{cw} and Q (Figures 4 to 6) are in accordance with the reactor operation and safety conditions. At the second stage in Figure 4, T reached the superior permissible limit and F_{cw} was increased to balance the system, while the two heaters kept switched on during the reaction time (Figures 5 and 6). At 200 minutes, X_m calculated by the model has a small decrease (Figure 3) of about 0.02, which is naturally unfeasible. Additionally, the data between some of the discretization intervals apparently shows a pseudo discontinuity in X_m trajectory. However, both observations can be explained by the IDP methodology. For example, when the system equations are integrated from $tf - 2L$ to $tf - L$, it is unlikely that the state $\boldsymbol{x}(tf - L)$ will be exactly one of the $\boldsymbol{x}$-grid points at stage P. This problem of not hitting an $\boldsymbol{x}$-grid point exactly is overcame by taking the optimal control policy corresponding to the $\boldsymbol{x}$-grid point that is closest to the state $\boldsymbol{x}(tf - L)$. As has been shown by DeTremblay and Luus (1989), this gives a good approximation if a sufficiently large number of $\boldsymbol{x}$-grid points and of allowable values for control are used. Figure 4 shows that T increases and decreases very fast. Although it is well known that U changes considerably with X_m, it was not taken into account by the model. It is reasonable due to the high dilution used in this work (S = 8.833 kmol/m^3 and M = 2.413 kmol/m^3). For this dilution condition U does not change significantly. Naturally T affects the characteristics of PMMA produced. For example, the number average molecular weight (Mn=Mm[$(\mu_1+\lambda_1)/(\mu_0+\lambda_0)$]) is different when T is constant and when T varies. The experimental value obtained by Antunes et al. (2005) after the runs by the gel permeation chromatography technique was 4.1x10^5 kg/kmol for Mn, while the IDP optimization results indicated 4.9x10^5 kg/kmol for Mn. As a matter of fact, Mn will be the major concern as performance index in future works and this justifies not using a maximum constant T during the MMA polymerization to maximize X_m.

5. Conclusions

The present work is concerned with the optimization of MMA batch solution polymerization reaction. It was successfully developed a program in Matlab to perform an off line optimization through IDP technique. The control variables used were F_{cw} and

Q, and the model equations were reliable to predict the polymerization reaction behavior. The choice of IDP technique and X_m as performance index supplied relevant subsidies on the trajectories for T, F_{cw} and Q, which can be used in future for experimental works in order to improve the reactor performance and PMMA production. These subsidies showed that the definition of different set points for T at each time interval affects positively the process, and contributes to achieve better results for X_m in comparison with the use of a fixed set point for the whole reaction time. In the experimental reactor of Figure 1, the T profile obtained from IDP procedure (Figure 4) would be implemented as set point tracking using a conventional or advanced temperature controller. Regarding the relevance and power of IDP, for future works the reactor performance should be evaluated in terms of optimum global efficiency and PMMA characteristics. An augmented perform index could be created by unifying the properties *Mn*, *Mw* (number average molecular weight) and X_m, (Upreti, 2005).

Nomenclature

A, $A\infty$	Reactor-jacket and surrounding-jacket heat transfer surface areas, m^2
c, c_w	heat capacity of reacting mixture and water, kJ/kg.K
f_o, f	initiator initial efficiency and at any time t
k_d, k_i, k_f, k_p,k_s, k_{tc}, k_{td}	rate constants for the equations 1 to 10 at any time t (s^{-1} or $m^3.mol^{-1}.s^{-1}$)
I, I_{pf}	concentration of initiator, $kmol/m^3$; or performance index, dimensionless
M	concentration of monomer, $kmol/m^3$
MW_I, MW_m, MW_s	molecular weight of pure initiator, monomer and solvent, kg/mol
R_{li}, R_{lm}, R_{ls}	primary radical of initiator, monomer and solvent in liquid phase, $kmol/m^3$
R_{vm}, R_{vs}	primary radical of monomer and solvent in vapor phase, $kmol/m^3$
S, s	concentration of solvent, $kmol/m^3$, and shifting term for the penalty function
T, T_{cw}, T_j, T_∞	temperature of reactor, inlet cooling water, jacket and room respectively, K
$\boldsymbol{u}$	control variable vector
U, U_∞	overall heat-transfer coefficients of reactor-jacket and jacket-surrounding, $kJ/m^2.s.K$
$\boldsymbol{x}$, X_m	state variable and monomer conversion rate
ΔH_p	heat of propagation reactions, kJ/kmol
ρ, ρ_{cw}, ρ_s, ρ_m, ρ_p	density of reacting mixture, water, MMA liquid, polymer or solvent respectively, kg/m^3

References

Achilias, D. S., Kiparissides. (1998). C. Journal of Applied Polymer Science, v. 35, p.1303.

Antunes, A. J. B., Pereira, J. A. F. R. & Fileti, A. M. F. (2005). Fuzzy control batch reactor: development and experimental testing. Computers and Chemical Engineering, 30, 268-276.

Chakravarthy, S. S. S., Saraf, D. N. & Gupta, S. K. (1997). Use of genetic algorithms in the optimization of free radical polymerization exhibiting the Trommsdorf effect. J. of Applied Polymer Science, 63, 529-548.

Ekpo, E. E. & Mujtaba, I. M. (2007). Optimal control trajectories for a batch polymerization reactor. International Journal of Chemical Reactor Engineering, 5, Article A4.

Ellis, M. F., Taylor, T. W., Gonzales, V. & Jensen K. F. (1988). Estimation of the molecular weight distribuition in batch polymerization. AICHE Journal, 34 (3), 1341-1353.

Gosh, P., Gupta, S. K., Saraf, D. N. (1998). An experimental study on bulk and solution polymerization of methyl methacrylate with responses to step changes in temperature. Chem. Eng. Journal, v. 70, p.25-35.

Luus, R. (2000). Iterative dynamic programming. Monographs and Surveys in Pure and Applied Mathematics. Florida: Chapman & Hall/CRC. It contains all the citations of this author.

Pedrosa, L. S., Nunes, R. L., Pereira, J. A. F. R., Fileti, A. M. F. (2001). A neural based inference system to the on-line estimation of polymer quality. 6th World Congress of Chemical Engineering, Melbourne.

Ray A.B., Saraf D.N ., Gupta S .K . (1995). Free radical polymerization associated with the Trommsdorff effect under semibatch reactor conditions. I : Modeling, Polym. Eng. And Sci., 35, 16, 1290-1299.

Schröder, A., Mendes, M. J. (1999). Time optimal control of distillation columns: a mixed IDP-SQP approach. Computer and Chemical Engineering Supplement, S491-S494.

Seth, V. & Gupta, S. K. (1995). An experimental study on bulk and solution polymerization of methyl methacrylate with responses to steps changes in temperature. Journal of Polymer Engineering, 15, 283.

Soroush, M., Kravaris, C. (1992). Nonlinear control of a batch polymerization reactor: an experimental study. AIChE Journal, v. 38 (9), p.1429-1448.

Upreti, S. R., Sundaram, B. S., Lohi, A. (2005). Optimal control of MMA polymerization in non-isothermal batch reactor using bifunctional initiator. European Polymer Journal, 41, 2893-2908.

Zeybek, Z., Yuse, S., Hapoglu, H., Alpbaz, M. (2004). Adaptive heuristic control of a batch polymerisation reactor. Chem. Eng. Proc., Vol. 43, 911-920.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimum heating configuration of pultrusion process

Lizandro S. Santos[a], Rogério L. Pagano[b], Evaristo C. Biscaia[b], Verônica M. A. Calado[a].

[a]*Chemical School, Federal University of Rio de Janeiro, 21941-909, Rio de Janeiro, Brasil.*
[b]*Chemical Engineering Program/COPPE, Federal University of Rio de Janeiro, 68502, 21941-914, Rio de Janeiro, Brasil.*

Abstract

Pultrusion is a continuous polymeric manufacture process for producing composite parts with constant cross-sectional profiles. In this process, the fiber/resin mixture is continuously pulled through a heated die at constant speed. As the fiber/resin system goes through the die cavity, a cure reaction occurs and this reaction releases a considerable amount of energy. But initially, the die needs to be heated in order to start the cure reaction. For this objective, electrical heaters are generally used on die surface to provide enough energy to start the polymerization reaction. The heating power must be carefully designed to achieve a minimum degree of cure. In this present work, an "I" beam geometry using a configuration with internal heaters was simulated by Finite Volume Method, based on a computational fluid dynamics package. Thus, a sum of all heating power was considered as the objective function in order to find the minimal energy consumption. This process heating overall cost was minimized applying a stochastic algorithm, the so-called particle swarm optimization (PSO) algorithm. The hybrid nature of the developed computational program evolves CFD tools to simulate the process allied with a PSO algorithm to optimize it. The obtained results showed that the heat flux spent to cure the resin-fiber system may be considerably reduced producing a significant decrease in process cost. The superior quality of the results was characterized by their comparison with reported literature values. At the same time, the developed methodology was always capable to find an optimum design of pultrusion process.

Keywords: cure reaction, computational fluid dynamics, optimization, particle swarm.

1. Introduction

Pultrusion is a continuous polymeric manufacture process for producing composite parts with constant cross-sectional profiles. In this process, the fiber and resin mixture is continuously pulled through heated die at a constant speed. As the fiber/resin system goes through the die, a cure reaction occurs and this reaction causes the change of composite physical properties besides a considerable energy production, because of the reaction exothermic characteristic (Pagano *et al*. 2006). A schematic diagram of pultrusion is shown in Fig. 1.

Some studies have been conducted in order to gain an understanding of the pultrusion optimization. Li *et al*. (1999) optimized the pultrusion process based on the uniformity of the degree of cure at the die exit. The optimization procedure has been applied in

order to obtain a part characterized by a good quality, i.e. a uniform distribution of the degree of cure with a satisfactory mean value, at the exit of the forming die. This procedure was based on an iterative procedure by using the Steepest Descent coupled to finite element method. Carlone *et al*. (2007) used a hybrid method based on genetic algorithms and simplex method for minimizing the same function proposed by Jianhua *et al*. (2001). However, these researchers failed to provide methods that regard economic objectives. Coelho and Calado (2001) proposed an economical objective function with the aim to find the optimal temperature profile using the Simulating Annealing Method. Srinivasagupta *et al*. (2003) used a multi-objective optimization which involves control, environmental and quality aspects for injected pultrusion optimization. The approach was successful in determining optimal values of the processing parameters such as heating zone temperatures, as well as the equipment specifications (die dimensions and heater) that satisfy the multiple objectives. However, it was seen that the economic objective did not coincide with quality aspects objectives and a tradeoff becomes necessary.

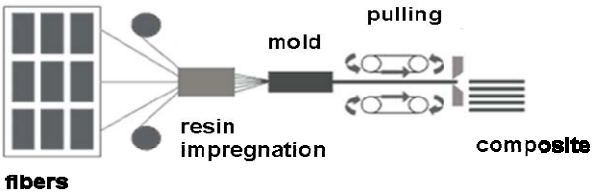


Figure 1. Schematic diagram of pultrusion process.

In the present investigation, a numerical procedure based on CFD technique and a stochastic optimization algorithm is developed. Our strategy is to join the desirable characteristics of CFD tools and Particle Swarm Optimization (PSO) algorithm (Kennedy and Eberhart) to result in an efficient optimization procedure. The aim is to minimize the process cost, represented by an objective function. The application of this technique on pultrusion process has not been explored yet.

2. Mathematical model

2.1. Equations

The mathematical model of the steady state pultrusion process is composed by the energy balance and the kinetic cure reaction. These equations can be written as:

$$\rho_c \; C_{pC}\left(u_i \nabla T\right) = k_c \nabla^2 T + c_{in}(1-\phi)\left(-\Delta H\right)R_a \tag{1}$$

$$\frac{d\alpha}{dz} = R_a = \left(Ae^{\left(\frac{-Ea}{RT}\right)} \right)(\alpha)^m (1-\alpha)^n \tag{2}$$

where T is temperature, u velocity vector, C_c specific heat of the composite, k_c thermal conductivity of the composite, ($-\Delta H$) total heat generated, ρ_c the density of the composite, R_a resin reaction rate, ϕ fiber volume fraction, c_{in} resin initial concentration, α degree of cure, A frequency factor, E activation energy, and m, n is the reaction order. In this work, the degree of cure was assumed to be zero *(α = 0)* before the fiber-resin system entering into the die, and the inlet temperature was

considered as 298 K. In order to compare our results with the results presented by Liu X.L (2001) the same physical and kinetic parameters were considered here, shown in Table 1.

Table 1. Physical and Kinetic Parameters

Fiber fraction ***(Φ)***	0.55
Density of the resin ***(ρ_r)***	1100 (kg/m^3)
Density of the fiber ***(ρ_f)***	2560 (kg/m^3)
Thermal conductivity of the resin ***(k_r)***	0.169 (J/m.s)
Thermal conductivity of the fiber ***(k_f)***	1.04 (J/m.s)
Thermal capacity of the resin ***(Cp_r)***	1640 (J/kg.K)
Thermal capacity of the fiber ***(Cp_f)***	640 (J/kg.K)
Initial temperature ***(T_0)***	300 K
Initial concentration of resin ***(cin)***	1100 (kg/m^3)
Frequency Factor ***(A)***	1.869x10^6 (s^{-1})
Activation Energy ***(E_{a1})***	71.688x10^3 (J/mol)
Order of reaction ***(n+m)***	2
Universal constant of gases ***(R)***	1.98 (cal/mol. K)
Heat generated ***(ΔH)***	398 (J/g)

2.2. Optimization procedure

In order to propose the pultrusion process optimization it was used an objective function similar to that proposed by Coelho and Calado (2001), which represents the process profit per unit time:

$$F = Q_{spent}.P_e \tag{3}$$

where Q_{spent} is the total heat rate spent to cure the system and P_e is the price of energy consumed.

The constrained non-linear optimization problem is written as:

$$\min_{\theta,\alpha\in\Re^n} F(q,\alpha) \tag{4}$$

subjected to:

$$g(\theta) = \alpha_{\min} - \alpha$$

where $\alpha_{\min}$ is the minimum desired value of degree of cure.

The problem is formulated as a new unconstrained optimization problem by adding a quadratic penalty function:

$$F(q,\alpha) = F + \psi.[\max(0, g(\theta))]^2 \tag{5}$$

where ψ is the penalty parameter.

The optimization procedure method applied to minimize the objective function was PSO (Kennedy and Eberhart). This method allows a global minimization in a predefined search region, without the need of initialization parameter guesses and without differentiation of the objective function. In this method, a set of points, nominated particles, is configured to find the optimal value of objective function.

CFD packages enable to model engineering problems in a realistic geometry with good mesh resolution. The software used here is the Ansys CFX^{TM}, a CFD package

based on finite volume method. In this technique, the region of interest is divided into small sub-regions, called control volumes. The equations are discretized and solved iteratively for each control volume. However, it is not provided of optimization tools. Our strategy is to compound characteristics of CFD tools with a PSO algorithm to design optimal heating strategies of the process. In the proposed strategy, PSO estimates heat flux values and transfers these values to a management algorithm that starts the CFD package. At the end of simulation, the algorithm reads the final value of degree of cure. According to simulation results, PSO estimates new values of heat flux and starts the same procedure until the convergence, as illustrated in Fig. 3.

In a previous paper, Liu *et al.* (2001) used a three-dimensional model for simulating the pultrusion of vinyl-ester-fiberglass I-beam composite. The procedure is successful in predicting the temperature and curing profiles in the pultruded material under different process conditions. The temperature predictions were verified by comparing to the experimental results published in the literature.

Assuming this fact, our methodology is to decrease the superficial area of heaters intended to minimize the energy consumption used in the process. We believe that the typical heater configuration, suggested by Liu *et al.* (2001), is not efficient because it spends high amount of thermal energy. We proposed a heat configuration consisted of five cylindrical heaters inside an insulated die. This configuration increases algorithm possibilities to find optimal heat arrangements.

Pultrusion was conducted using an insulated steel die of 1020 mm in length. Heater dimension is 5 cm of diameter and 14 cm of length. Adopted mesh consists of 170617 nodes and 888242 elements. Using symmetry, only a quarter of pultrusion tool should be considered in the numerical model.

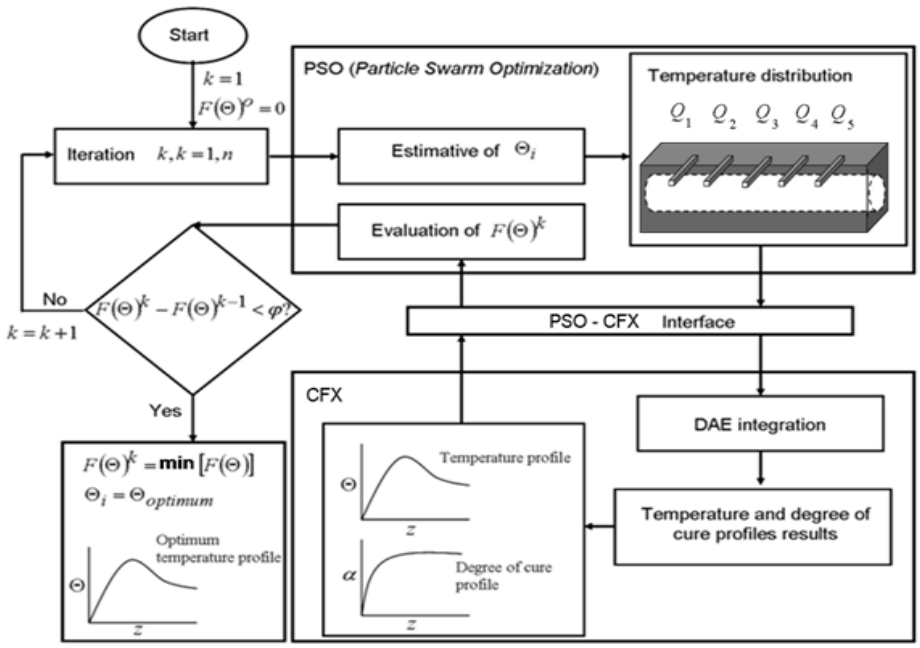


Fig.3. Illustration of optimization procedure.

3. Results

Optimization results of composite and die are shown in the following figures.

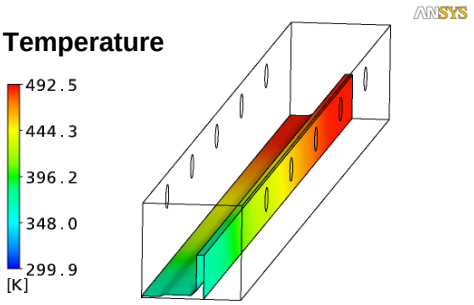


Fig. 4. Composite temperature profile.

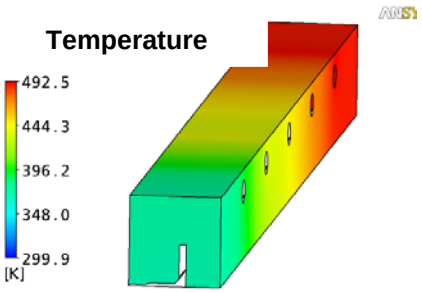


Fig.5. Die temperature profile.

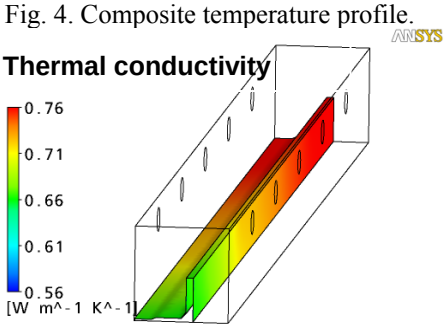


Fig.6. Thermal conductivity.

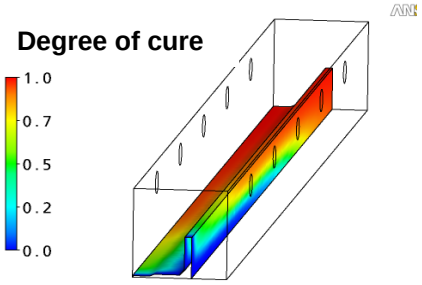


Fig.7. Degree of cure profile.

Temperature distribution of composite and die part are shown in Fig. 4 and 5. When the exothermic reaction starts, released heat raises the composite temperature until 492 K. Fig 6 shows that thermal conductivity increases with temperature and degree of cure. Fig.7 shows that the degree of cure achieves its maximum value at die exit; this figure clearly demonstrates that the die was adequately heated providing complete cure of the resin.

Optimization results are shown in Table 2. At each iterative step, PSO estimates heat flux values of each heater. Efficiency of the algorithm is well characterized in Table 2 through the value reduction of heat rate in successive iterations. Based on the obtained results, energy used to cure the resin/fiber system may be considerably reduced. The results show that the best choice is to use only the third heater with a heat flux of 67.69 $kW \cdot m^{-2}$ or 446 K. This arrangement results in a total heat rate of 6.7 kW while Liu *et al.* (2001) used 9.2 kW, meaning that there is the possibility to cure the material with lower energy consumption and our configuration is approximately 26% more economic.

Table 2. Optimization results

Iteration	q_1 ($kW.m^{-2}$)	q_2 ($kW.m^{-2}$)	q_3 ($kW.m^{-2}$)	q_4 ($kW.m^{-2}$)	q_5 ($kW.m^{-2}$)	Total heat rate *(kW)*
1	23.35	62.15	42,05	26,42	32,41	18.24
2	18.87	4.70	53.72	48.98	32.10	15.84
3	0.00	0.00	38.30	37.69	18.19	16.66
4	0.00	0.00	47.98	33.80	0.00	8.17
5	4.24	5.32	45.07	22.63	1.26	7.85
6	0.83	0.00	38.54	38.91	0.00	7.82
7	3.18	0.00	42.29	31.89	0.00	7.73
8	0.00	0.00	67.69	0.00	0.00	**6.77**

4. Conclusions

In the present paper, it was verified that smaller energy consumption can be obtained with the suggested heat strategy in a pultrusion process. This economy was verified by the reduction of the surface area of the heaters causing a significant decrease of the energy consumption. This present heat configuration seems to be a good alternative to improve global thermal efficiency of the process.

PSO algorithm coupled to CFD code was successful in finding an optimal temperature profile resulting in a minimal process cost. Our results showed that the algorithm converges to the minimal value of heat rate. Furthermore, the solution did not violate the constrained region defined in the mathematical formulation. This characteristic guarantees the efficiency of PSO in finding the optimal solution. The results obtained herein suggest that new studies about heater arrangements of the pultrusion process must be made.

Acknowledgements

Authors would like to thank CAPES, CTPETRO/FINEP (# 0104081500) and CNPq for financial support.

References

Coelho, R.M.L., Calado, V.M.A, 2001. An optimization procedure for the pultrusion process based on a finite element formulation. Polymer Composites, 23, 329-341.

Jianhua, L., Joshi. C.S., Lam, Y.C, 2001. Curing optimization for pultruded composite sections. Composite Science and Technology, 62, 457-467.

Liu, X.L. Numerical modeling on pultrusion composite I beam, 2001. Composites Part A: Applied Science and Manufacturing, 32, 663-681.

Liu, X.L., Hillier, W, 1999. Heat transfer and cure analysis for the pultrusion of a fiberglass-vinyl ester I beam. Composite Structures, 47, 581-588.

Kennedy, J., Eberhart, R, 1995. Particle Swarm Optimization. In: Proc. IEEE International conference on neural networks Perth, Australia, 1942-1948.

Carlone, P., Palazzo, G.S., Paquito, R, 2007. Cure optimization by hybrid computational methods. Computers & Mathematics with Applications, 53(9): 1464-1471.

Pagano, R.L. Santos, LS., Calado, VMA., Biscaia, E.C, 2006 Pultrusion process: modeling of heat transfer and cure kinetic. In: World Polymer Congress Rio de Janeiro, Brazil.

Srinivasagupta, D., Kardos, JL., Joseph B, 2003. Rigorous dynamic model-based economic design of the injected pultrusion process with controllability considerations. Journal of Composite Materials, 37(20): 1851-1880.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Published by Elsevier B.V.

Optimization Technique Applied for the Solid-Liquid Equilibrium of Binary Fatty Mixtures

Stella Alonso Rocha,[a] Reginaldo Guirardello,[a]

[a]*LSOPQ/FEQ – Department of Chemical Process - University of Campinas – Unicamp Av. Albert Einstein, n. 500, CEP 13081-970, Campinas, Brazil*

Abstract

Optimization technique for the solid-liquid equilibrium calculation of binary fat mixtures with natural origin was used with the software GAMS. The calculation of phase equilibrium is based on the minimization of Gibbs free energy of the system. For the thermodynamic representation of mixtures Margules and the Slaughter & Doherty Model were used the former referring to the liquid phase and the latter for the solid one. The mixtures studied are composed by saturated, insaturated fat acids and triglycerids. The models developed in this work are written in a convex nonlinear programming, so that the optimal solution found is the global optimum. Therefore, the results found are the equilibrium points.

Keywords: Solid-Liquid Equilibrium, Gibbs free energy, eutectic and peritectic points, convexity analysis, Global minimization.

1. Introduction

The solid-liquid equilibrium data have an important hole in separation process, like extraction, distillation, absorption, and others, they are also used in foods, pharmaceutical, chemical, and recently in bio-diesel industry. Then (tenta evitar o so), the reliable determination of phase equilibrium is necessary for the correct prediction of phase formation and their composition [1].

Thus, for theoretical determinations by computational programs it is necessary to improve the thermodynamic models that could represent the behavior of the solid and liquid phases, particularly the solid phase, due to the existence of the peritectic point [2,3,4] since most of the available models are able to represent only the eutectic point [4,5,6].

The solid-liquid equilibrium can be calculated in many different ways. In the present work, the minimization of Gibbs free energy was chosen for its calculation. This technique is considered sufficient whereas it guarantees that the equilibrium point was found if the global minimum is too [7,8]. The thermodynamic representation of the solid and liquid phases was made using specific models for each phase. The liquid phase was modeled using Margules 2-sufix equation [2,4]. The solid phase was modeled using a modification of the activity model described by Slaughter & Doherty [9,4], using a mathematical limit. The final equations obtained resulted in an expression for Gibbs free energy.

Oils and fats, which are composed of mixtures of fatty acids, are the main compounds in this work [10]. Three of these mixtures were studied: myristic - palmitic acids, oleic - capric acid and palmitic acid – triestearin. Their Diagram phases could be determinate by the technique developed here.

In this context, the main objective of this work was obtained points which can represent the solid liquid equilibrium of fatty mixtures using mathematical models associated a computational programs by the minimization of Gibbs free energy, using GAMS software, where the convexity analysis guarantee the global minimum, consecutively the equilibrium state.

2. Mathematical model

2.1. *Fundamentals of Minimization of Gibbs free energy technique*

The phase equilibrium could be obtained by Minimization of Gibbs free energy given by [1]:

$$G = \sum_{i=1}^{NC} \sum_{j=1}^{NF} n_{ij} m_{ij} \tag{1}$$

Subject by:

$$n_{ij} \geq 0\,, \quad i = 1,..., NC; \quad j = 1,..., NF \tag{2}$$

That represents the non negativity of mol numbers, and:

$$\sum_{j=1}^{NF} n_{ij} = n_i\,, \qquad i = 1,..., NC \tag{3}$$

That represents the mass balance without chemical reaction.

2.2. Thermodynamics models – Solid and liquid phases

The solid-liquid equilibrium is described by the ratio of fugacities, given by [7]:

$$\ln \frac{f_i^{l,o}}{f_i^{s,o}} = \frac{\Delta h_{f\,i}}{R \cdot T_{t\,i}} \cdot \left(\frac{T_{t\,i}}{T} - 1 \right) - \frac{\Delta c_{p\,i}}{R} \cdot \left(\frac{T_{t\,i}}{T} - 1 \right) + \frac{\Delta c_{p\,i}}{R} \cdot \ln \frac{T_{t\,i}}{T} \tag{4}$$

In equation 4, two simplifications are necessary: the fusion temperature is used instead of the triple temperature and the variation of heat capacity can be neglected The equation is approximated by:

$$\ln \frac{f_i^{l,o}}{f_i^{s,o}} = \frac{\Delta h_f}{R \cdot T_f} \cdot \left(\frac{T_f}{T} - 1 \right) \tag{5}$$

Defining the parameter B_i as:

$$B_i = \ln \frac{f_i^{l,o}}{f_i^{s,o}} = \frac{\Delta h_f}{R \cdot T_f} \cdot \left(\frac{T_f}{T} - 1 \right) \tag{6}$$

For the liquid phase, the equation used to describe the comportment was Margules – 2 suffix, given by:

$$\underline{G}^{ex} = R \cdot T \cdot \sum_{i=1}^{NC} x_i \cdot \ln g_i = \frac{1}{2} \cdot \sum_{i=1}^{NC} \sum_{j=1}^{NC} A_{ij} \cdot x_i \cdot x_j \tag{7}$$

On the other hand, the solid phase was described for an approximation of Slaughter and Doherty [9], given by:

$$g_i^s = \frac{1}{x_i^s + e} \tag{8}$$

The parameter ε in equation 7, was used by Slaughter and Doherty to avoid some numerical problems in the calculation of the equilibrium constant [9]. However, in this

work it is used a different approach for the equilibrium calculation, it can use the limit of $e \to 0$ in the Gibbs free energy.

In this work, the model for the solid phase was considered to be one that obeys the Gibbs-Duhem equation and that has the following properties:

$$g_i^s \cdot x_i^s = 1 \qquad \text{for} \qquad 0 < x_i^s \le 1 \tag{9}$$

$$g_i^s \cdot x_i^s = 0 \qquad \text{for} \qquad x_i^s = 0 \tag{10}$$

$$\lim_{x_i^s \to 0} \left(x_i^s \cdot \ln\left(g_i^s \cdot x_i^s \right) \right) = 0 \tag{11}$$

2.3. *Minimization Problem*

Associating the thermodynamics model for the liquid and solid phases, the Gibbs free energy can be written as:

$$G = \sum_{i=1}^{NC} n_i^0 \cdot m_i^{s,o} - n_p^0 \cdot \Delta G_R^o + n_p^s \cdot \Delta G_R^o + R \cdot T \cdot \sum_{i=1}^{NC} n_i^l \cdot \left[B_i + \ln n_i^l - \ln \sum_{j=1}^{NC} n_j^l \right] + \frac{\sum_{i<j}^{NC} A_{ij} \cdot n_i^l \cdot n_j^l}{\sum_{j=1}^{NC} n_j^l} \tag{12}$$

and satisfying the restrictions:

$$n_i^s + n_i^l = n_i^0 + n_i \cdot \left(n_p^s - n_p^0 \right) \qquad i = 1, \mathrm{K}, NC \qquad i \neq p \tag{13}$$

$$n_i^s \ge 0 \qquad i = 1, \mathrm{K}, NC \tag{14}$$

$$n_i^l \ge 0 \qquad i = 1, \mathrm{K}, NC \qquad i \neq p \tag{15}$$

$$n_p^l = 0 \tag{16}$$

The equilibrium composition is then calculated by minimizing G, given by equation (12), with respect to the variables n_i^l and n_i^s, satisfying the restrictions (14) – (16), at constant T and P. Other quantities that are also kept constant during the minimization are n_i^0, n_i, B_i, A_{ij}, ΔG_R^o, and $m_i^{s,o}$.

3. Algorithm for Optimization

Considering the mathematical problem proposed, the solution was an algorithm applied in GAMS, using the Gibbs free energy minimization with the global minimization, based on the equations developed in the previous section.

The necessaries parameters to obtain the phase diagrams were taken from Rolemberg [2]. The parameters are physical properties of pure compounds, such as melting temperatures (T_f), enthalpy variation at the melting point (Δh_f), the interaction phase parameter of Margules equation (A_{12}), and the Gibbs free energy of reaction (ΔG_R^o).

In mixtures where the formation of the intermediate compound (characterized by the peritectic point) does not occur, the parameter ΔG_R^o can assume any positive value or zero.

Other initial parameters were necessary, the atmospheric pressure *P* = 1.00 atm and the universal gas constant *R* = 1,9872 *cal/mol.K*

After the parameters and models implementation in GAMS, using CONOPT solver, that used GRG solving method, can be possible to obtain diagrams phase of the mixtures studied.

4. Case studies and Numerical Results

In this work, three case studies are considered: myristic acid – palmitic acid, oleic acid - capric acid, palmitic acid - triestearin. These examples and the necessaries parameters were taken from Rolemberg [2], and are presented in Table 1.

Table 1: Parameters for the solid-liquid equilibrium [2]

Compound	T_f (K)	Δh_f (kcal/mol)	A_{12} (cal/mol)	ΔG_R^o (cal/mol)
Myristic acid	327.07	10.800	-297.2904	-308.0957
Palmitic acid	335.02	13.100		
Oleic acid	286.59	9.460	-648.5496	≥ 0
Capric acid	303.98	6.690		
Palmitic acid	335.02	13.100	44.8203	≥ 0
Tristearin	345.27	48.580		

The determination of solid-liquid phase equilibrium can be done by minimization of gibbs free energy method, for the three fatty binary systems. The mathematical model presented resulted in a convex NLP, applied in GAMS, using CONOPT solver.

4.1. Case Study 1 – myristic acid – palmitic acid

The solid-liquid phase diagram for this mixture is the set of points (x_1, T), where compound 1 is myristic acid and compound 2 is palmitic acid.

Figure 1 shows the calculated values for the equilibrium temperatures at different molar fractions of myristic acid, where the calculated curve is compared with the experimental points from Rolemberg [2]. It can be seen a good agreement between the calculated and the experimental data. Both eutectic and peritectic points are shown in the phase diagram, where the presence of the intermediate solid compound occurs for x_1 between 0.5 and 0.7.

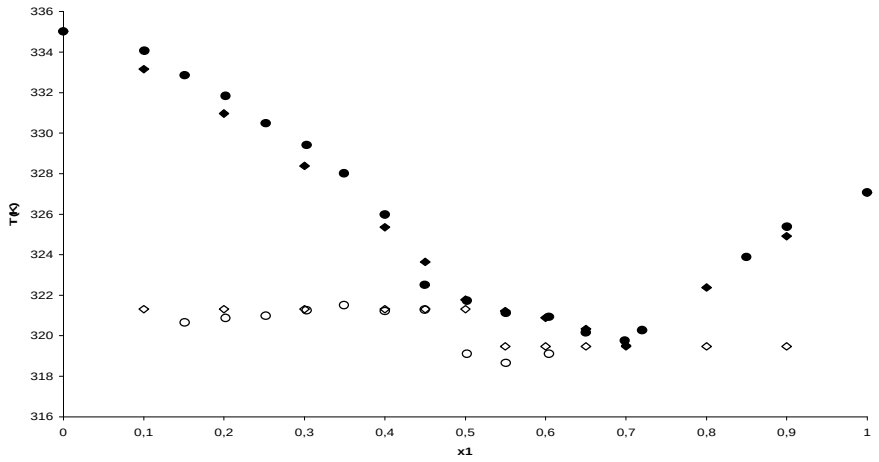


Figure 1: Phase Diagram for myristic (1) – palmitic (2) acids: ♦ fusion temperature - calculated; ◊ transition temperature – calculated; ? fusion temperature – experimental; ? transition temperature - experimental.

4.2. Case Study 2 – oleic acid – capric acid

The solid-liquid phase diagram for this mixture is the set of points (x_1, T), where compound 1 is oleic acid and compound 2 is capric acid.

Figure 2 shows the calculated values for the equilibrium temperatures at different molar fractions of oleic acid, where the calculated curve is compared with the experimental points from Rolemberg [2]. It can be seen that the calculated data have a good agreement with the experimental ones. Only the eutectic point is shown in the phase diagram, where there is not presence of the intermediate compound.

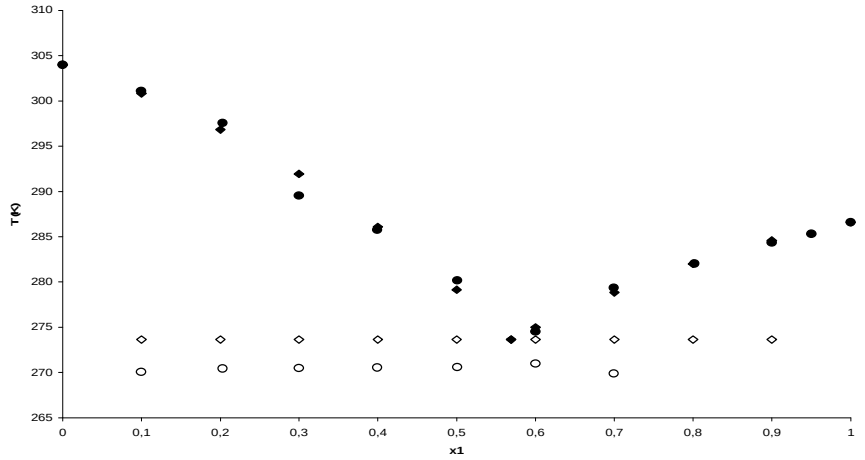


Figure 2: Phase Diagram for oleic (1) – capric (2) acids: ♦ fusion temperature - calculated; ◊ transition temperature – calculated; ? fusion temperature – experimental; ? transition temperature - experimental.

4.3. Case Study 3 – palmitic acid – tristearin

This case study is a mixture composed of a saturated fatty acid and a triglyceride, with palmitic acid ($C_{16}H_{32}O_2$) and tristearin ($C_{57}H_{110}O_6$), respectively.

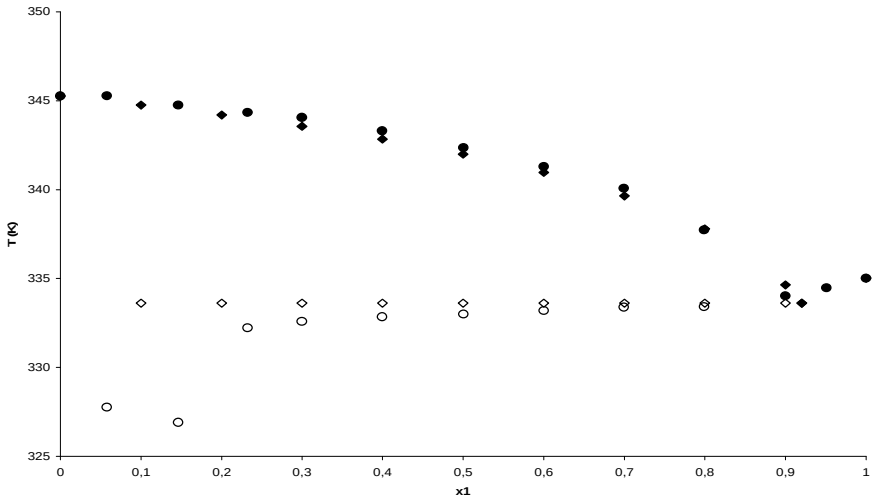


Figure 3: Phase Diagram for palmitic acid (1) – Tristearin (2): ♦ fusion temperature - calculated; ◊ transition temperature – calculated; ? fusion temperature – experimental; ? transition temperature - experimental.

Figure 3 shows the calculated values for the equilibrium temperatures at different molar fractions of palmitic acid, where the calculated curve is compared with the experimental points from Rolemberg [2]. It can be seen that there are a good agreement between the calculated and the experimental data. Only the eutectic point is shown in the phase diagram, where there is not presence of the intermediate compound.

5. Conclusion

This study deals with an optimization of the solid-liquid equilibrium point of binary fatty mixtures. The phase diagrams were determinate for satured, insatured and triglycerids binary fatty mixtures. The results obtained by the methodology developed in this work shows good agreement with the experimental data of Rolemberg [2]. The software GAMS shows to be an important tool for optimization, because of its implementation facility of the mathematical model. This new developed technique is capable to detect the intermediate compound, characterized by peritectic point in any temperature that this point could exist, including the perfect identification of the eutectic point. This work shows good results in determining the solid-equilibrium point of fatty mixtures.

References

[1] G. P. Rangaiah, Fluid Phase Equilibria, 187-188 (2001) 83-109.
[2] M. P. Rolemberg, Equilíbrio Sólido-Líquido de Ácidos Graxos e Triglicerídios: Determinação Experimental e Modelagem, Doctorate Thesis on Chemical Engineering – State University of Campinas (2002).
[3] L. A. D. Boros, Modelagem Matemática e Termodinâmica do Equilíbrio Sólido-Líquido de Sistemas Graxos, Master Thesis on Chemical Engineering – State University of Campinas (2005).
[4] S. A. Rocha, Cálculo do Equilíbrio de Fases Sólido-Líquido em Misturas binárias por meio de Técnicas de Minimização e Análise de Convexidade, Master Thesis on Chemical Engineering – State University of Campinas (2008).
[5] S. Bruin, Fluid Phase Equilibria, 158-160 (1999) 657-671.
[6] C. C. Huang, Y. P. Chen, Chemical Engineering Science, 55 (2000) 3175-3185.
[7] J. M. Prausnitz, R. N. Lichtentaler, E. G. Azevedo, Molecular Thermodynamics of Fluid-phase Equilibria, 2nd ed., Prentice-Hall PTR, 1986.
[8] J. M. Smith, H. C. Van Ness, M. M. Abbott, Introduction to Chemical Engineering Thermodynamics, 7th ed., McGraw-Hill, 2004.
[9] D. W. Slaughter, M. F. Doherty, Chemical Engineering Science, 50, nº 11, (1995) 1679-1694.
[10] K. W. Won, Fluid Phase Equilibria, 82 (1993) 261-273.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal production scheduling for the sausage industry

Paulo Eduardo Polon[a], Cid Marcos Andrade Gonçalves[a], Luiz Mario de Matos Jorge[a], Paulo Roberto Paraiso[a].

[a]*Department of Chemical Engineering, State University of Maringá (UEM) On Colombo avenue, at 5790, Edifice D90, Zip Code 87020-900, Maringá – PR, Brazil Phone: (0-XX-44) 3261-4752 – Fax: (0-XX-44) 3261-4792.*

Abstract

The increased consumption of processed foods, as well as a greater variety of products offered by the food industry has required for greater flexibility and the most efficient coordination in the allocation of existing resources. The sausage industry in particular in recent years had an increase in consumption, and this has led to a search for a more efficient use of available resources to achieve increased productivity, cost reduction, reliability in attendance and reduction of lead times. These aspects of production and management can be achieved more easily through tools of programming production. The sausage production is a batch process. The process can be divided into two main stages, the first refers to the preparation, the mixture of raw materials and stuffing; the second stage refers to the cooking of products that were stuffed in the first stage. A variety of types of products are obtained in the sausage industry. Most of these products do not have an exclusive line and therefore, there is a dispute between equipment and intermediate states. The products are classified as perishable, and it is recommended that production should not exceed the demand under the risk of keeping products in stock and overcome the shelf life validity of the product. The objective of this paper was to optimize the production for the sausage industry using values of demands for each product. The proposed approach states the problem as a model of mixed integer linear programming (MILP) to represent the decisions involved and was solved using the GAMS (General Algebraic Modelling System) language. The model was applied to the problem of production scheduling with real-world operational data from an industry and was able to generate better results than those used by industry. Thus, the model proved to be appropriate to support the decisions of production scheduling in the sausage industry.

Keywords: scheduling, sausage manufacture, mixed integer linear programming

1. Introduction

The increased consumption of processed foods and greater variety of products offered by the sausage industry has required for greater flexibility and efficient coordination in the allocation of existing resources.

The scheduling has attracted the attention of both academia and industry. The reason for this is the pressure on industries to improve productivity and reduce costs and also the substantial progress related to the models and techniques of solution and the rapid growth of computational power. The problem of interest is to determine the most

efficient way to produce a set of products in a time horizon, given a set of limited resources and production recipes. The activities of scheduling generally take place in multiproduct or multipurpose plants, in which a variety of different products can be manufactured via the same or different recipes sharing of limited resources such as equipment, material, and time.

Mathematical programming, especially Mixed Integer Linear Programming (MILP) because of its flexibility and extensive modeling, has become one of the most exploited for scheduling problems. Applications of MILP based scheduling methods range from the simplest single-stage single-unit multiproduct processes to the most general multipurpose processes. These planning processes scheduling problems are inherently combinatorial in nature because of the many discrete decisions, such as equipment assignment and tasks allocation over time (Floudas and Lin, 2005).

Production recipes in chemical processes can be very complex. For scheduling, the representation in networks has been developed to represent the proceeds of production unambiguously. Kondili, Pantelides and Sargent (1993) proposed a general structure State Task Network (STN). The STN representation of a process is a direct graph with two distinct types of nodes: state nodes represented by a circle, representing raw materials, intermediate products and final products, and tasks nodes represented by a rectangle, representing a physical or chemical operation, as reaction, mixing, heating.

According to Mendez et al. (2006) to formulate a mathematical model for any problem scheduling, the first major issue that arises is how to represent the time. This article used to the discrete time model, where a time horizon of interest is divided into a number of time intervals of uniform durations. The start/end of a task and other important event are associated with the boundaries of these time intervals.

Kondili, Pantelides and Sargent (1993) and Shah, Pantelides and Sargent (1993) proposed a discrete-time formulation for short term scheduling of batch operations, based on the STN representation that was used with some adjustments in this article.

The use of models of MILP programming to represent the decisions involved and resolve them using the GAMS software is used frequently for the optimization of production of real-world industries (Araujo, Arenales and Clark, 2004, Toso and Morabito, 2005 and Dogan and Sarimveis, 2008).

2. Production Process

This work is based on case study of the sausage industry. This industry produces 8 different types of sausage, such as: fresh sausage, cooked sausage, hotdog sausage, bologna and ham. The process production consists of two steps. In the first step of the process production meat products are selected and then cut into pieces small sufficiently to pass easily by the mouthpiece of grinder machine, for each type of sausage use an appropriate degree of comminution. To work with the meat close to freezing point it is recommended, this facilitates the cut operations, reduces the abrasion of the equipment and mainly minimizes the contamination. After grind, the meat is placed in appropriate containers and stored in refrigerated room. Additives and spices are carefully weighed and well blended. Then this blend is added to the grind meat, these materials are mixed in a mixer until they form a compact mass and well homogenized, to avoid formation of air bubbles during the stuffing (Rocco, 1996). Before the stuffing some products are cured, the process, which is the development of characteristics of color, flavor and texture of each product, depending on the product this process can vary from 8 to 24 hours. Natural and artificial casings are used as forms and containers for sausages.

In the process of stuffing the emulsion prepared is placed on top of the stuffer machine, it is compressed by a piston electric drive, the force to leave the equipment in a stuffing tube, of various sizes, according to type of product being stuffing. After stuffing the products are placed in special cooking cabinets or rods.
In a second step, the products are cooked in smokehouses or tanks of cooking. After cooking the products are cooled, packaged and transported to storage or shipment.
The real process analysed, is composed of 7 stuffing proper for stuffing 8 different products, the production capacity of stuffers varies from 700 kg/h to 3000kg/h according to the product being produced. The stuffers 1 and 2 are suitable for stuffing the following products: ham, hamlike, bologna type 1 and bologna type 2. The stuffers 3, 4 and 5 are suitable for stuffing the following products: fresh sausage, cooked sausage cooked sausage type 1 and type 2 and stuffers 6 and 7 are suitable for stuffing the hotdog sausage. The process of cooking in smokehouse has 6 smokehouse with capacity ranging from 3,200 kg to 1,600 kg and the cooking time varies from 2 hours for hotdog sausage until to 8 hours for bologna type 1. The process of cooking in tanks with hot water suitable for cooking the ham and hamlike products has 3 cooking tanks, 2 heat shock tanks and 5 cooling tanks, each tank has a capacity of 1,500 kg of material and processing time is 4 hours for cooking, 1 hour for the shock heat and 6 hours for cooling.

3. Methodology

Primarily, an examination of the industrial production of sausage was making. To facilitate the implementation of the model, divided the process into two sections according to the scheme of work seen in the industry under study. In the first stage, which the products are stuffed, the production time is 18 hours. In the second stage, which the products earlier stuffed are cooked in a smokehouses or in tank of cooking, the production time is 24 hours. Although the production process of sausage is in various stages, any these stages not require to be scheduled because the equipment to work with idle time, such as grinders, mixers and tanks of cure, so the main bottlenecks of production are the stuffed and cooked sections.
To develop the model, the representation of the production process in the STN as shown in Figure 1 was made. The production of sausage has three production lines distinct: (a) Fresh sausage, (b) sausage cooked in a smokehouses and (c) sausage cooked in tanks of cooking.

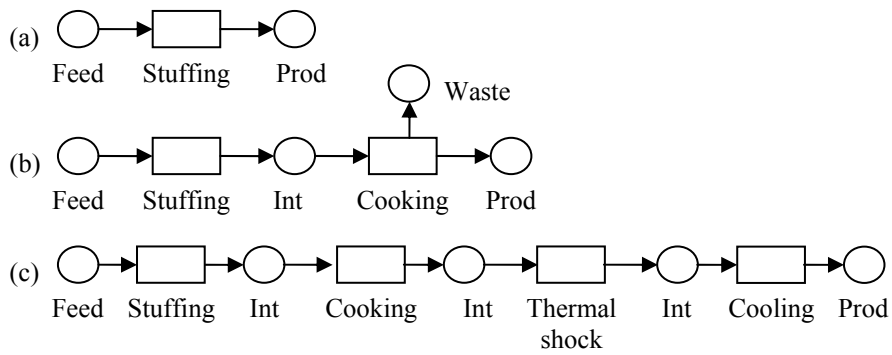


Figure 1 STN representation for production of sausage.

In addition to STN, the set, the equipment available for processing of each task, the intermediate storage policy, the production requirements and a time horizon of interest are defined. The objective is to determine a program that includes the sequence of tasks to be performed in each piece of equipment and the amount of material to be processed by each task.
The production process has two stages, so the solution of the problem was conducted in two modules, one for the processing time of 18 hours (Stuffing step) and another for the processing time of 24 hours (cooking step). The results will be obtained in stuffing step serve as input to the cooking step. It considered in this model there is full availability of raw materials and labour.
The criteria of performance were used to maximize the plant overall production. The production demand is not seasonal, with historical data obtained from industry, the average production of each product is estimated. Due to the perishable nature of products in this work used minimum and maximum production demands.
The model used was adapted and modified the MILP problem presented by Kondili, Pantelides and Sargent (1993) and Shah, Pantelides and Sargent (1993) and resolved by the modeling language GAMS.

4. Results

The optimal schedule is shown in Figure 2 and Figure 3 for the production time of 18 hours and production time of 24 hours respectively. The number above each bar is the tasks, while the number below the bar denotes the amount of material being processed.
The optimal combined global production for the Product was 145,828 kg; this result was 19.53% above the production industrial.

Equipment																				Tasks
stuffer1	4	4	4	4	4	4	4	4	1	1	3	4	2	4	1	1	1	2		1 stuffing1
	3000	3000	3000	3000	3000	3000	3000	3000	1000	1000	2500	3000	2200	3000	1000	1000	1000	2200		
stuffer2	4	4	4	4	3	4	4	4	1	4	1	4	4	4	2	2	2	2		2 stuffing2
	3000	3000	3000	3000	2500	3000	3000	3000	1000	3000	1000	3000	3000	3000	2200	2200	2200	2200		
stuffer3	7	7	5	7	7	5	7	5	5	5	5	5	7	5	5	5	5	5		3 stuffing3
	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800	1800		
stuffer4	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6		4 stuffing4
	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000		
stuffer5	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6	6		5 stuffing5
	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000		
stuffer6	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8		6 stuffing6
	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500	2500		
stuffer7	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8	8		7 stuffing7
	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200	1200		
																				8 stuffing8
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Time

Figure 2. Gantt chart for first step.

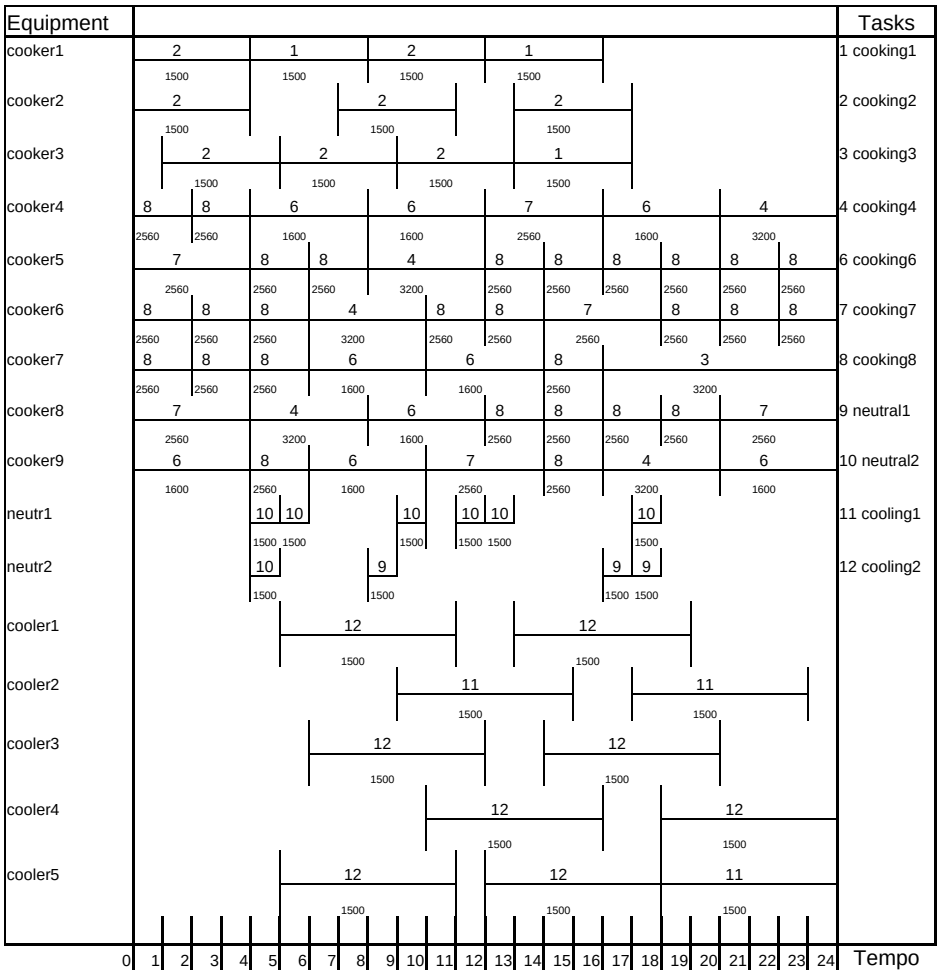


Figure 3. Gantt chart for second step.

Table 1 Production

Product	Industrial Production (kg)	Scheduling Production (kg)
Ham	4500	4500
Hamlike	11200	12000
Bologna 1	2200	2880
Bologna 2	14800	15040
Fresh sausage	15100	21600
Cooked sausage 1	12000	12240
Cooked sausage 2	5000	13056
Hotdog sausage	57200	64512
Total	122000	145828

Table 1 shows a comparison between the results obtained in the optimization and the results of the industry. Three products (fresh sausage, cooked sausage 2 and hotdog sausage) obtained very best production that the industrial production.

5. Conclusions

This paper presented the optimization of production to a sausage industry, using a general formulation adapted to the problem that considers both the sequencing and scheduling of production. A comparison between the production manufactures by industry and the production obtained by the program, the efficiency of the method was verified and the using the method in the industry achieve better utilization of equipment and the plant productivity was improved.

References

S. A. Araujo, M. N. Arenales and A. R. Clark, (2004), Dimensionamento de lotes e programação do forno numa fundição de pequeno porte, Gestão & Produção, 11, n.2 p165-176.

P. Doganis, and H. Sarimveis, (2008), Optimal production scheduling for the dairy industry. Annals of Operations Research, 159, 315-331.

C. A. Floudas and X. Lin, (2005), Mixed Integer Linear Programming in Process Scheduling: Modeling, Algorithms, and Applications, Annals of Operations Research, 139, 131-162.

E. Kondili, C. C. Pantelides, and R. W. H Sargent, (1993), A General Algorithm for Short-Term Scheduling of Batch Operations - I. MILP Formulation, Computers and Chemical. Engineering, 17, n.2, 211-227.

C. A. Méndez, et. al., (2006), State-of-the-art review of optimization method for short-term scheduling of batch processes, Computers and Chemical Engineering, 30, 913-946.

S. C. Rocco, (1996), Embutidos, Frios e Defumados, Embrapa – SPI, Coleção Saber, 94p.

N. Shah, C. C. Pantelides, and R. W. H Sargent, (1993a), A General Algorithm for Short-Term Scheduling of Batch Operations - II. Computational issues, Computers and Chemical. Engineering, 17, n.2, 229-244.

E. A. V. Toso and R. Morabito, (2005), Otimização no Dimensionamento e Sequenciamento de Lotes de Produção: Estudo de Caso numa Fábrica de Rações, Gestão & Produção, 12, n.2 p 203-217.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Development of a Morphological Population Balance Model in Principal Component Space

Cai Y Ma and Xue Z Wang*

Institute of Particle Science and Engineering, University of Leeds, Leeds LS2 9JT, UK
** Correspondence author: Xue Z Wang. Email: X.Z.Wang@leeds.ac.uk*

Abstract

Multi-dimensional and morphological population balance models for crystallization processes can simulate the dynamic evolution of particle shape as well as particle size distribution. These models however can become computationally expensive when the crystal has a large number of independent faces, and are not applicable to non-crystalline, irregularly shaped particles such as those encountered in granulation and milling. This paper addresses these challenges by introducing principal component analysis (PCA) to morphological population balance modeling. PCA transforms the shape description of a particle from a high dimensional domain to a lower dimensional, principal component space. Morphological population balance models can then be built in this latent variable space, greatly reducing the computational complexity. It also makes it possible to model non-crystalline irregularly shaped particles. The original particle shape at any time can be reconstructed from the principal components. The methodology is demonstrated via modeling the cooling crystallization of potash alum.
Keywords: Principal Component analysis; Morphological population balance; Multidimensional population balance; Crystal shape; Model reduction

1. Introduction

Population balance (PB) models have been widely used in engineering for modelling the dynamic evolution of particle size distribution (PSD). However, in these models, the size of a particle is often defined as the diameter of a sphere having the same volume as the particle, missing important information about particle shape. Motivated by this observation, research has been carried out to introduce particle shape information into PB simulation, and multi-dimensional PB models were proposed for crystallization processes. The models can predict the evolution of PSD in more than one size dimension [1-4] but the majority of efforts have been limited to using two characteristic size dimensions, e.g. the length and width distributions of plate-like crystals. Multidimensional PB models in literature are not always able to fully reconstruct the particle shape unless for very specific situations such as plate like crystals. Ma et al. [5, 6] proposed a morphological PB model for crystallization processes. It defines the shape and size of a crystal as a function of the normal distances of all its faces to the geometrical centre, and predicts the dynamic evolution of size distribution of every face for the population of crystals. Morphological PB model emphasizes the capability of full reconstruction of the particle shape in high fidelity at any moment of the simulation. Since a morphological PB model is able to retain the particle shape, it provides an effective multi-scale modelling tool, linking single particle simulation e.g. single crystal morphology prediction, with a process model about the behaviour of the particle population in a processing vessel.

In principle, the morphological PB model has no limit on the number of faces for the modelled crystals, i.e., the number of size dimensions. But with the increase in the number of size dimensions, solution of the PB equations becomes computationally too expensive, resulting in unacceptably long CPU time. In applying the morphological PB model to the case studies [5, 6], symmetrical faces were treated as having identical growth behaviour, which effectively reduced the size dimensions, and so the CPU time.

The application of multi-dimensional PB models to processes such as granulation and milling that produce irregularly shaped particles has not been attempted. This does not imply that particle shape is not important for these processes but do indicate the difficulties to accurately define the shape of these irregular particles. Therefore, the way to apply multi-dimensional PB models to these processes is not obvious.

The purpose of this study is to develop a new methodology for carrying out shape based PB modeling that uses only a small number of size dimensions, but misses little shape information. Based on principal component analysis (PCA), the method uses principal components (*PC*s) to describe the particle shape so that PB modeling can be carried out in a transformed, latent variable domain or *PC* space. The *PC*s are calculated directly from the original multiple size dimensions of particles but the number of *PC*s can be much smaller than that of the original size dimensions. At any time, the real particle shape can be reconstructed from the latent *PC*s using the PCA model.

2. Shape Description in Principal Component Space

2.1. Principal component analysis

Given a data set consisting of m observations for n variables, $x_1, x_2, \ldots x_n$, PCA calculates a new variable y_1 that accounts for the variation in the n original variables. The first *PC*, y_1, is the linear combination of the n variables, i.e., $y_1 = w_{11} x_1 + w_{12} x_2 + \ldots + w_{1n} x_n$. The sample variance is greatest for all of the coefficients, $w_{11}, w_{12}, \ldots w_{1n}$ (written as a vector $\boldsymbol{w_1}$) with a constraint, $\boldsymbol{w'_1 w_1} = 1$. The second *PC*, y_2, is also given by the linear combination of the n original variables as, $y_2 = w_{21} x_1 + w_{22} x_2 + \ldots + w_{2n} x_n$ or $\boldsymbol{y_2} = \boldsymbol{w'_2 x}$, which has the greatest variance subject to two conditions, $\boldsymbol{w'_2 w_2} = 1$ and $\boldsymbol{w'_2 w_1} = 0$ (so that y_1 and y_2 are uncorrelated). Similarly the jth *PC* can be calculated as $y_j = \boldsymbol{w'_j x}$, which has greatest variance subject to $\boldsymbol{w'_j w_j} = 1$, $\boldsymbol{w'_j w_i} = 0$ $(i < j)$.

The coefficients defining the first *PC*, i.e. the elements of $\boldsymbol{w_1}$ can be determined by maximizing the variance of y_1 subject to a constraint, $\boldsymbol{w'_1 w_1} = 1$. The solution of $\boldsymbol{w_1}$ to maximize the variance y_1 is the eigenvector of variance-covariance matrix, **S**, corresponding to the largest eigenvalue. The eigenvalues of **S** are roots of the equation,

$$||S - \lambda I| = 0 \tag{1}$$

By arranging the eigenvalues, $\lambda_1, \lambda_2, \ldots \lambda_n$ in descending order, the top d eigenvectors, $d << n$, are the *PC*s that capture most of the variance of the original data while the remaining *PC*s can be mainly considered as noise. Hence the data is reduced from n dimensions to a much smaller dimension of d with loss of mainly the noise information.

The operation can be described in matrix algebra as follows:

$$[y]_{m\times n} = [x]_{m\times n}[w]_{m\times n} \qquad \text{and} \qquad [y]_{m\times d} \approx [x]_{m\times n}[w]_{m\times d} \tag{2}$$

Then the original variables can be re-constructed by the following operation:

$$[x]_{m\times n} \approx [y]_{m\times d}[w_{m\times d}]^T \tag{3}$$

2.2. Shape and size description in principal component space

For a crystalline particle, such as a potash alum crystal having 26 faces, its shape and size can be represented by the normal distance of each face to the geometrical centre [5, 6]. Using the same method to define the shape and size of a non-crystalline, irregularly shaped particle in two-dimension (2D), the distance of every pixel at the edge to the geometric centre is required. To describe the shape and size of the same object in three-dimension (3D), it requires the distances to the geometric centre from all the points on the particle surface. In summary, the shape and size of a single particulate object in 2D or 3D, either a structured crystal or an irregularly shaped particle, can be described by a finite number of sizes, namely, x_1, x_2, …, x_n. To consider a population of m particles, each with n size descriptors, a matrix is formed as shown in (E-1).

	x_1	x_2	----	x_n
particle - *1*				
⋮				
particle - *m*				

(E-1)

Applying PCA to this matrix and selecting d *PC*s, $d < n$, leads to a new matrix with the same number of columns but d rows of *PC*s which represent the new latent size dimensions with each being calculated by the linear combination of all the original sizes x_1, x_2, …, x_n. To conduct multi-dimensional PB modelling in the original space, one has to consider all the n dimensions. The idea of multi-dimensional PB modelling in a *PC* space is to model size distribution evolution in each of the *PC* directions, *PC1*, *PC2*, …, *PCd*, via solving d ($<< n$) dimensional PB equations without compromising much on the shape information. At any time of simulation, shape and size of particles represented by x_1, x_2, …, x_n, can be reconstructed from the selected *PC*s using eq (3).

3. Morphological PB Modeling of Potash Alum Crystallization in *PC* Space

A case study of a seeded potash alum crystallization process is presented here to demonstrate that morphological PB modeling can be carried out in *PC* space. The crystal morphology of potash alum crystals is dominated by the large face {111} and two smaller faces {100} and {110}. In this study, secondary nucleation, crystal breakage and aggregation are excluded [6]. Therefore, all needed to carry out morphological PB modelling in the *PC* space are the initial size distributions of the selected *PC*s, and the growth rate in each *PC* direction as a function of operating conditions.

time	particles	x_1	x_2	----	x_{26}
t_1	particle - *1*				
	⋮				
	particle - *m*				
⋮	⋮				
t_T	particle - *1*				
	⋮				
	particle - *m*				

(E-2)

3.1. Calculation of initial PC size distributions and growth rates in PC directions

The simulated results [6] based on three independent faces was used to generate data for the other 23 faces. The obtained data is arranged into a matrix (E-2). The columns, x_1, x_2, …, x_{26}, represent the sizes of the 26 faces. The rows are divided into sections, t_1, t_2, …, t_T, representing the time points of simulation. At a given time point, there are m rows, representing m particles, particle-*1*, …, particle-*m*. So the data matrix shown in (E-2) is of the size $(m \times T) \times 26$.

For each time point of t_1, t_2, …, t_T, however, there would be too many rows due to too many particles (in this study, $m = 2.2 \times 10^7$). As a result, only a limited number of representative particles of different shape and size were selected (3375 in this study). The simulated time span was 3300s for 57 time steps. So the matrix (E-2) has elements

of (3375×57)×26. PCA was applied to the matrix (E-2). Three *PC*s were selected based on the slope of variance vs. *PC* number. A new matrix with three columns and the same number of rows were formed. The new matrix was then used to estimate the growth rates in each *PC* direction as follows:

$$G_1=1.05\times10^{-5}\sigma^{1.24}\times Size_{PC1}{}^{0.16};G_2=6.28\times10^{-7}\sigma-5.84\times10^{-8};G^3=-1.50\times10^{-8}\sigma+1.22\times10^{-9} \quad (4)$$

3.2. PB model in the PC space

The PCA-based three dimensional PB equation can be written as follows:

$$\frac{1}{V_T(t)}\frac{\partial}{\partial t}[\psi(PC1,PC2,PC3,t)V_T(t)]+\frac{\partial}{\partial PC1}[G_1(PC1,t)\psi(PC1,PC2,PC3,t)] \quad (5)$$
$$+\frac{\partial}{\partial PC2}[G_2(PC2,t)\psi(PC1,PC2,PC3,t)]+\frac{\partial}{\partial PC3}[G_3(PC3,t)\psi(PC1,PC2,PC3,t)]=0$$

where ψ is the number population density function, G_1, G_2 and G_3 are the growth rates in three *PC* directions, and V_T is the total volume of the suspension in the reactor.

To solve Eq. (5), the initial distributions of crystal population in 26 dimensions were converted to new distributions in the three *PC* dimensions, which were curve-fitted using a Gaussian function. The PCA-based PB equation (Eq. 5) can be solved using a discretisation method, moment of classes, which is only briefly described here with its detailed description to be found in literature [1, 5].

The 3D size domain was equally discretised into n_1, n_2, n_3 classes, which form a set of $n_1\times n_2\times n_3$ ordinary differential equations:

$$\frac{1}{V_T(t)}\frac{d}{dt}\left[N_{i,j,k}(t)V_T(t)\right]+[FX^{O}_{i,j,k}(t)-FX^{I}_{i,j,k}(t)]+[FY^{O}_{i,j,k}(t)-FY^{I}_{i,j,k}(t)]+[FZ^{O}_{i,j,k}(t)-FZ^{I}_{i,j,k}(t)]=0 \quad (6)$$

where $N_{i,j,k}(t)$ is the number of crystals in the class $Cl_{i,j,k}$, and *FX*, *FY* and *FZ* are the crystal flow fluxes in three PC directions with the superscripts, *O* and *I*, denoting the crystal flowing outletting from and inletting into the $Cl_{i,j,k}$ class [5, 6].

The discretised PCA-based PB equations (Eq. (6)), together with the initial size distribution and boundary conditions, were solved simultaneously using the Runge-Kutta-Fehlbergh solver [7] in the mesh ranges of 22~114μm, -24~51μm and -26~79μm in *PC1*, *PC2* and *PC3* directions, respectively, with the classe sizes being 100×100×100. For a crystal with known values for *PC1*, *PC2* and *PC3*, the reconstruction of its values of x_1, x_2, ...x_{26} can be easily performed using equation (3).

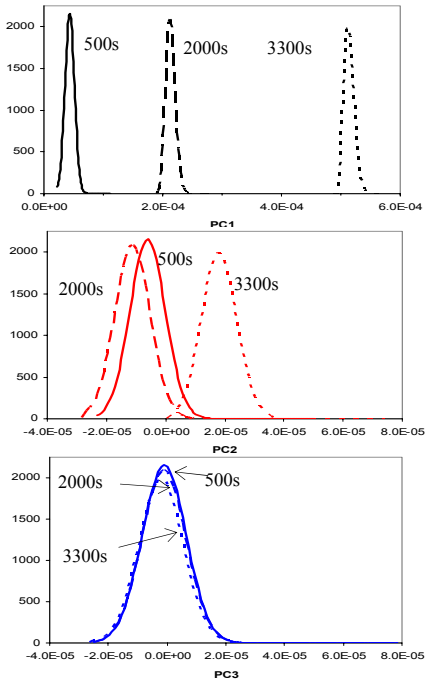


Figure 1. Typical size distributions for individual PCs.

4. Results and Discussions

Figure 1 shows typical size distributions of three *PC*s at 500th, 2000th and 3300th second. The most significant *PC*, *PC1*, grew much faster with time, while the second significant *PC*, *PC2*, was firstly decreasing with time, then increasing after 2150th second. The third *PC*, *PC3*, did not show obvious change during crystallization. Using the PCA reconstruction technique, the size distributions of

the 26 faces were obtained. Three typical distributions in x_1, x_9 and x_{15} face directions, representing {111}, {100} and {110} faces, were illustrated in Figure 2. Also plotted in Figure 2 are size distributions in x_1, x_9 and x_{15} obtained from the original morphological PB model [6]. Figure 2 shows that PCA reconstruction produced size distributions very close to that of the original simulation.

Since the sizes for the 26 faces can be reconstructed from *PC*s, the shape of crystals can be drawn at any time during crystallization process. Figure 3 shows the shape evolution of a crystal. At 3300s, both the crystal shape reconstructed from *PC*s and the original shape from the original simulation were drawn.

Figure 4 shows comparison between simulated evolution trajectories of the {110} boundary sector using PCA approach and and the origianl morphological PB model [6].

The case study proved the principle that morphological PB modeling can be effectively performed in the transformed, latent variable space. The real particle shape at any time of simulation can be reconstructed with high fidelity from only a few *PC*s. It also brings about a question: how can the methodological PB be used in practice? It clearly cannot be expected that the original data of crystal size distributions of all 26 faces over the time span is always available for other processes.

The PCA based morphological PB modeling technique can be used in various ways, and is open for future exploration. Below we look at one example. We still consider a seeded cooling crystallization process with the assumptions that secondary nucleation, aggregation and breakage can be ignored. The seed particles can be characterized to get data of the shape and size. The data can be arranged into a matrix in the form of (E-1). Applying PCA to the matrix and then selecting *d* number of *PC*s will produce a new matrix in the form of (E-2). The size distribution for each of the *d PC*s can be estimated. The above procedure produces: 1) the number of *PCs*, *d*, for which population balance simulation will be carried out; 2) the initial size distributions for the selected *d* principal components; and 3) a transformation matrix containing the loadings that can be used at any time of the simulation to reconstruct the original *n* variables (original shape descriptors) from the values of the *d PCs*.

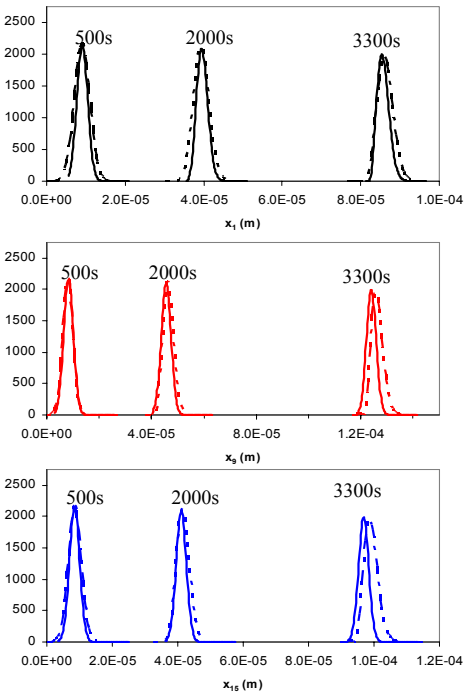


Figure 2. Size distributions of individual faces: solid curves (reconstructed from PCs); dotted curves (original simulation).

The only missing item prior to PCA based morphological PB simulation is the growth rate of each *PC* as a function of supersaturation (σ) and its size:

$$G_{PCi} = f(\sigma, \text{size of } PC_i, \text{parameters}) \quad (7)$$

Suppose the structure of equation (7) is known, the PCA based PB equations can be formulated with its parameter values to be determined via model identification using data from a real crystallization experiment.

A real cooling crystallization experiment can be carried out using the characterized seeds with operating conditions, e.g. temperature and supersaturation being continuously recorded. The product at the end should be characterized in the same way as for seeds. The product characterization data can then be used as the targets in model identification [3] for determining the parameters in eq (7). PCA based morphological PB models can thus be built to simulate the crystallization process of the same chemical in other operating conditions and varied seed crystal shape and size.

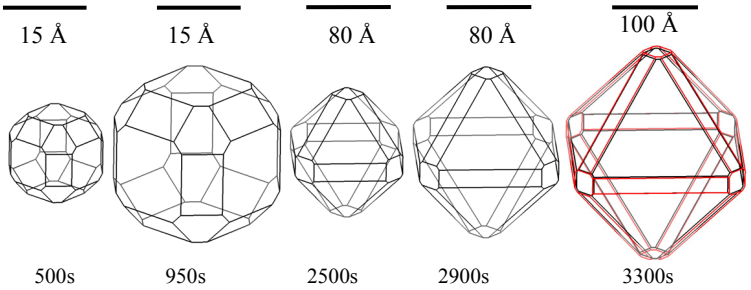
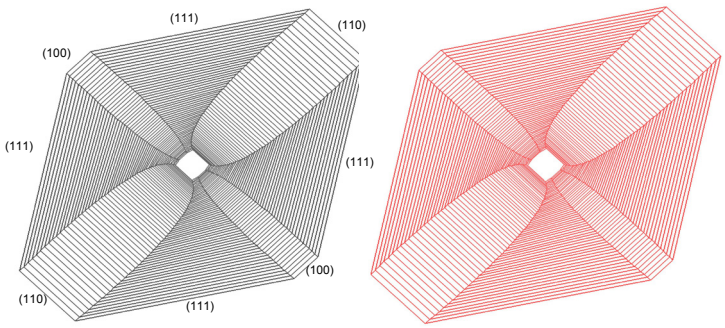


Figure 3. Shape evolution for a crystal, reconstructed from PCs (at 3300s, black is reconstructed from PCs and red is the original shape).



(a) (b)

Figure 4 Evolution of {110} boundary sector, (a) reconstructed from principal components and (b) original simulation.

Acknowledgements

The work is funded by EPSRC (Grant references: EP/C009541 and EP/E045707). We would also like to thank the industrial collaborators including AstraZeneca, Malvern Instruments, National Nuclear Laboratory, Pfizer, Syngeta and 3M Health Care.

References

1. Puel, F., G. Fevotte, and J.P. Klein, Chem. Eng. Sci., 2003. **58**:3715 & 3729.
2. Ma, D.L., D.K. Tafti, and R.D. Braatz, Ind. & Eng. Chem. Res., 2002. **41**:6217.
3. Oullion, M., F. Puel, G. Fevotte, and S. Righini, Chem. Eng. Sci., 2007. **62**:833.
4. Ma, C.Y., X.Z. Wang, and K.J. Roberts, Adv. Powder Tech., 2007. **18**:707.
5. Ma, C.Y., X.Z. Wang, and K.J. Roberts, AIChE J., 2008. **54**:209.
6. Ma, C.Y. and X.Z. Wang, AIChE J., 2008. **54**:2321.
7. Shampine, L.F., Watts, H. A. *Subroutine RKF45*. in *Computer Methods for Math Comp.* 1977. Englewood Cliffs, N. J.: Prentice-Hall.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Synthesis of large-scale heat exchanger networks by a monogenetic algorithm

Xing Luo,[a,b] Georg Fieg,[a*] Kang Cai,[c] and Xin Guan[c]

[a]*Institute of Process and Plant Engineering, Hamburg University of Technology, D-21071 Hamburg, Germany (*Corresponding author, E-mail address: g.fieg@tuhh.de)*
[b]*Institute of Thermodynamics, Helmut Schmidt University, D-22043 Hamburg, Germany*
[c]*Institute of Thermal Engineering, University of Shanghai for Science and Technology, 200093 Shanghai, China*

Abstract

Synthesis of large-scale heat exchanger networks has attracted increasing attention but it is still a very difficult task because of the geometrically increasing of the design parameters to be optimized. In such cases, even long computing time were used, the result would usually not converge to the global optimum. In order to overcome this disadvantage, a monogenetic algorithm based on the optimization of sub-networks (functional groups) was proposed to improve the initial design obtained by the hybrid genetic algorithm. In this new procedure, the first search was carried out by the use of the hybrid genetic algorithm under a small population size and limited number of evolution generations so that the first design could be finished in relatively short time. Then, the sub-networks undergo recombinations and are further optimized with the hybrid genetic algorithm, in which a large population size and more evolution generations are used for the optimization of the sub-networks to ensure that the optimal solution of each sub-network can be found. An example from literature was calculated with this new method which yields a better network construction.

Keywords: large-scale heat exchanger network, hybrid genetic algorithm, monogenetic algorithm, optimization.

1. Introduction

The optimal design of large-scale process systems is a difficult task not only due to its non linear characteristics but also due to a number of local optima in its solution space. A typical example is the synthesis of a heat exchanger network (HEN) used for heat recovery. In refineries the amount of the streams to be heated up and cooled down could be up to 100. However, most of existing targeting techniques, particularly those from Pinch Technology (Linnhoff and Hindmarsh, 1983), are suitable only for small HENs. In recent years the genetic

algorithm was also applied to the HEN synthesis (Lewin, Wang and Shalev, 1998; Lewin, 1998; Chen et al., 2007), but it is not efficient for the synthesis of large-scale HENs. In such cases, even a large population size and long computing time were used, the result would usually not converge to the global optimum. Björk and Pettersson (2003) and Pettersson (2005) developed a genetic algorithm with sub-networks as genes which seems effective for synthesis of large-scale HENs.

In the present work the monogenetic algorithm is proposed which is based on the idea that any improvements in sub-networks (functional groups) will yield an improvement in the whole network. Since a sub-network usually contains less process streams, it is easy to find its optimal network structure. The recombination operation will stochastically applied to the sub-networks to form new sub-networks. An example from literature was used to illustrate the procedure and a better design was found.

2. Mathematical model for HEN synthesis

The HEN synthesis problem can be stated as follows: Given are N_h hot process streams, N_c cold process streams, N_{hu} hot utilities HU and N_{cu} cold utilities CU. Specified are heat capacity flow rates and the supply and target temperatures of each stream. Given also are the temperature levels and costs of the hot and cold utilities, the costs and overall heat transfer coefficients of heat exchangers, heaters and coolers. Determine the configuration of the HEN and the values of heat transfer areas and thermal flow rates of each heat exchanger in the HEN which bring the total annual cost of the HEN to the minimum.

The problem is solved basically by means of the hybrid genetic algorithm developed in our early work (Chen et al., 2007; Luo et al. 2009). It is assumed that the HEN has a stage-wise superstructure proposed by Yee et al. (1990), of which one restriction is that at each branch in a stage there is only one heat exchanger. However, the restriction of isothermal mixing used in their model is rescinded in the present analytical solution of stream temperature. The variables to be optimized are $(UA)_{ijk}$ and heat capacity flow rates $\dot{W}_{h,ijk}$ and $\dot{W}_{c,ijk}$, in which U is overall heat transfer coefficient, A is heat transfer area, and the subscripts i, j, and k denote the structure stage, hot stream, and cold stream, respectively. For a given set of $(UA)_{ijk}$, $\dot{W}_{h,ijk}$ and $\dot{W}_{c,ijk}$, the stream temperatures in the HEN can be explicitly obtained. To avoid the use of binary variables, we stipulate that a non-positive value of $(UA)_{ijk}$ means no heat exchanger at the position (i, j, k). With this stipulation, the cost of a heat exchanger in the HEN can be calculated by

$$C_{E,ijk} = \begin{cases} a + bA_{ijk}^{c}, & (UA)_{ijk} > 0 \\ 0, & (UA)_{ijk} \le 0 \end{cases} \tag{1}$$

in which a, b and c are known parameters for the calculation of heat exchanger cost.

In order to heat or cool the streams to their target temperatures, the strategy of excessive use of utilities (Lewin et al., 1998) is adopted. If the exit temperature of stream n, t_n'', is higher than the upper bound of its target value $t_{\mathrm{OUT},n}^{+}$, a cold utility should be used. Similarly, if t_n'' is lower than the lower bound of its target value $t_{\mathrm{OUT},n}^{-}$, it should be heated by a hot utility, no matter whether it is a hot stream or a cold one. Thus, the sum of utility cost for stream n (hot stream: n = 1, 2, …, N_h; cold stream: $n = N_\mathrm{h} + 1, N_\mathrm{h} + 2, \ldots, N_\mathrm{h} + N_\mathrm{c}$) and the cost of corresponding heater or cooler can be expressed as

$$C_{\mathrm{U},n} = \begin{cases} \min\{C_{\mathrm{HU},p}\dot{W}_n(t_{\mathrm{OUT},n}^{-} - t_n'') + a_p + b_p A_{\mathrm{HU},n}^{c_p};\ p = 1,2,\cdots,N_{\mathrm{hu}}\}, & t_n'' < t_{\mathrm{OUT},n}^{-} \\ 0, & t_{\mathrm{OUT},n}^{-} < t_n'' < t_{\mathrm{OUT},n}^{+} \\ \min\{C_{\mathrm{CU},q}\dot{W}_n(t_n'' - t_{\mathrm{OUT},n}^{+}) + a_q + b_q A_{\mathrm{CU},n}^{c_q};\ q = 1,2,\cdots,N_{\mathrm{cu}}\}, & t_n'' > t_{\mathrm{OUT},n}^{+} \end{cases} \tag{2}$$

$$A_{\mathrm{CU},n} = \frac{\dot{W}_n(t_n'' - t_{\mathrm{OUT},n}^{+})}{U_{\mathrm{CU},n}\Delta t_{m\mathrm{CU},n}} \tag{3}$$

$$A_{\mathrm{HU},n} = \frac{\dot{W}_n(t_{\mathrm{OUT},n}^{-} - t_n'')}{U_{\mathrm{HU},n}\Delta t_{m\mathrm{HU},n}} \tag{4}$$

in which $n = 1, 2, \cdots, N_\mathrm{h} + N_\mathrm{c}$ and Δt_m is the logarithmic mean temperature difference.

The objective function is the sum of heat exchanger costs, utility costs and costs of heaters and coolers calculated by Eqs. (1) and (2). Thus, the HEN synthesis problem becomes to find the values of $(UA)_{ijk}$, $\dot{W}_{\mathrm{h},ijk}$ and $\dot{W}_{\mathrm{c},ijk}$ so that the total annual cost reaches the minimum,

$$\min\ C(\mathbf{x}) = \sum_{n=1}^{N_\mathrm{h}+N_\mathrm{c}} C_{\mathrm{U},n} + \sum_{i=1}^{N_\mathrm{s}}\sum_{j=1}^{N_\mathrm{h}}\sum_{k=1}^{N_\mathrm{c}} C_{\mathrm{E},ijk}$$
$$x \in \{(UA)_{ijk}, \dot{W}_{\mathrm{h},ijk}, \dot{W}_{\mathrm{c},ijk} \mid i = 1,2,\cdots,N_\mathrm{s};\ j = 1,2,\cdots,N_\mathrm{h};\ k = 1,2,\cdots,N_\mathrm{c}\} \tag{5}$$

This problem is solved by means of the hybrid genetic algorithm. The values of $(UA)_{ijk}$, $\dot{W}_{\mathrm{h},ijk}$ and $\dot{W}_{\mathrm{c},ijk}$ are restricted in the regions of $[(UA)_{\min}, (UA)_{\max}]$, $[0, \dot{W}_{\mathrm{h},j}]$ and $[0, \dot{W}_{\mathrm{c},k}]$, respectively, with an empiric value of

$$(UA)_{\min} = -3(UA)_{\max} \tag{6}$$

and the physical restriction: If $(UA)_{ijk} \le 0$, then $\dot{W}_{\mathrm{h},ijk} = \dot{W}_{\mathrm{c},ijk} = 0$. The fitness of a HEN (individual) used in the hybrid genetic algorithm is evaluated with the reciprocal of its total annual cost and is adjusted with a commonly used linear transform,

$$f = \frac{C^{-1} + C_{\min}^{-1} - 2(C^{-1})_{\mathrm{avg}}}{C_{\min}^{-1} - (C^{-1})_{\mathrm{avg}}} \tag{7}$$

in which C is the total annual cost of the individual (See Eq. (5)), $C_{\min}$ the minimum total annual cost in the population, and $(C^{-1})_{\text{avg}}$ the average value of the reciprocal of total annual cost of the population.

3. Monogenetic algorithm

For a large-scale HEN, there might be more than ten thousand design parameters, and even the hybrid genetic algorithm often yields the solution to an unexpected local optimum. This is due to the probabilistic characteristics of the algorithm. The more the possible heat exchangers in a network exist, the less the possibility to obtain an optimal design. In practice a large-scale HEN usually consists of some functional groups. In traditional genetic algorithms the probability that such functional groups appear simultaneously in a HEN is so small that in fact it would never appears.

The new method for the large-scale HEN synthesis consists two optimization levels. At the upper level the evolution is carried out among the functional groups of a HEN, i.e. the functional groups are taken as individuals. Meanwhile, the functional groups themselves are optimized at the lower level with the hybrid genetic algorithm. Because of their small size, the optimization of the functional groups is usually easy to be performed. The sub-networks of the HEN can be taken as such functional groups. The sub-network is defined as a part of the HEN in which the streams have no thermal relation with the streams outside the sub-network. If a HEN can be divided into two or more sub-networks, then, these sub-networks are independent from each other. Therefore, any improvement in a sub-network will bring an improvement in the whole network at the same time.

The procedures are as follows: (1) Apply the hybrid genetic algorithm to the whole network with small individual number and small evolution generations to obtain an initial set of sub-networks. (2) Recombine two or more sub-networks stochastically to form a new sub-network. (3) Apply the hybrid genetic algorithm to this sub-network. If the new sub-network is better than the original parent sub-networks, this new sub-network replaces its parent sub-networks and will be decomposed into new sub-networks if possible. Otherwise, the original parent sub-networks will remain in the HEN. (4) Repeat steps (2)-(4) until no improvements can be found. More details can be found in our recent work (Fieg, Luo and Jezowski, 2009).

4. Case study

An example taken from Björk and Pettersson (2003) is used to illustrate the procedure. The problem data are given in Table 1. The annual cost of heat exchanger is $8000 + 800A^{0.8}$\$/yr where A is the area given in m^2. The hot utility cost is 70\$/kWyr and the cold utility cost is 10\$/kWyr. Their best design has an total annual cost of 2.073M\$/yr. A new design was obtained by Pettersson (2005) with the total annual cost of 1.998M\$/yr.

Table 1 Problem data

Hot streams	t' °C	t_{OUT} °C	$\dot{W}$ kW/K	h kW/m^{2}K	Cold streams	t' °C	t_{OUT} °C	$\dot{W}$ kW/K	h kW/m^{2}K
H1	180	75	30	2	C1	40	230	20	1.5
H2	280	120	15	2.5	C2	120	260	35	1
H3	180	75	30	2	C3	40	190	35	1.5
H4	140	45	30	2	C4	50	190	30	2
H5	220	120	25	1.5	C5	50	250	60	2
H6	180	55	10	2	C6	40	150	20	2
H7	170	45	30	2	C7	40	150	20	2
H8	180	50	30	2	C8	120	210	35	2.5
H9	280	90	15	2	C9	40	130	35	2.5
H10	180	60	30	2	C10	60	120	30	2.5
H11	120	45	30	2	C11	50	150	10	3
H12	220	120	25	2	C12	40	130	20	1
H13	180	55	10	2	C13	120	160	35	1
H14	140	45	20	2	C14	40	90	35	1.75
H15	140	60	70	2	C15	50	90	30	1.5
H16	220	50	15	2.5	C16	50	150	30	2
H17	220	60	10	2.5	C17	30	150	50	2
H18	150	70	20	2	HU	325	325	-	1
H19	140	80	70	2	CU	25	40	-	2
H20	220	50	35	2					
H21	180	60	10	2					
H22	150	45	20	2.5					

In the present work, an initial solution was obtained by hybrid genetic algorithm with 100 individuals and 1000 generations, which has the total annual cost of 2.062M$/yr. The initial HEN can be divided into 11 sub-networks. These sub-networks are further optimized with the hybrid genetic algorithm, which yields a better structure with 2.012M$/yr. Then, the monogenetic algorithm was applied to the sub-networks. After recombination, the new sub-networks will be improved and evaluated. This operation is repeated until no improvement can be further obtained. The total annual cost of the whole HEN reduces to 1.965M$/yr. The structure of the HEN is shown in Fig. 1, which has 44 units.

5. Conclusions

A new method for solving large-scale HEN synthesis problems is proposed as the monogenetic algorithm. In this method, the genetic algorithm is applied to the sub-networks of a HEN rather than the population of HENs. The monogenetic algorithm uses the sub-networks of a HEN as individuals. Since the sub-networks are small, they can be optimized with the hybrid genetic algorithm more efficiently. A case study was conducted and a better design than that in literature is obtained.

Acknowledgments

The present research was supported by Innovation Program of Shanghai Municipal Education Commission (No. 07ZZ89).

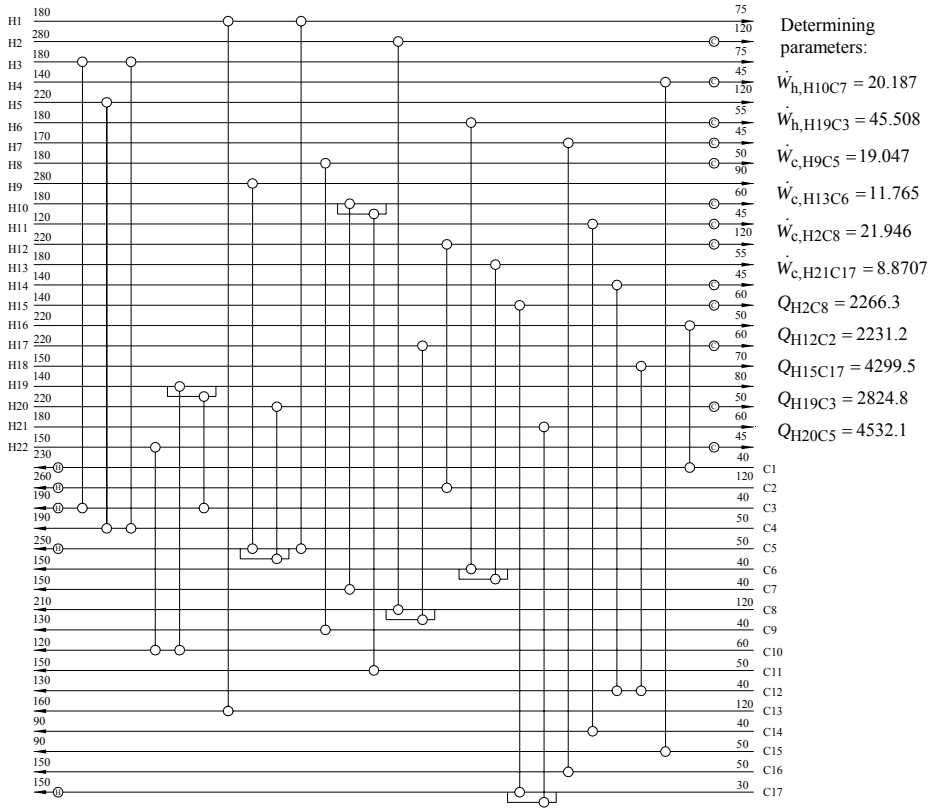


Fig. 1 Optimal heat exchanger network configuration, total annual cost 1,943,536$/yr

References

K.-M. Björk, F. Pettersson (2003), Optimization of large-scale heat exchanger network synthesis problems, Proceedings of the IASTED International Conference on Modelling and Simulation, M.H. Hamza (Ed.), Palm Springs, California, 313-318.

D.-Z. Chen, S.-S. Yang, X. Luo, et al. (2007), An explicit solution for thermal calculation and synthesis of superstructure heat exchanger networks, Chinese J. Chem. Eng. 15(2), 296-301.

G. Fieg, X. Luo, J. Jezowski (2009), A monogenetic algorithm for optimal design of large-scale heat exchanger networks, accepted by Chem. Eng. and Proc.

D.R. Lewin (1998), A generalized method for HEN synthesis using stochastic optimization - II. The synthesis of cost-optimal networks, Comput. Chem. Eng. 22(10), 1387-1405.

D.R. Lewin, H. Wang, O. Shalev (1998), A generalized method for HEN synthesis using stochastic optimization - I. General framework and MER optimal synthesis, Comput. Chem. Eng. 22(10), 1503-1513.

B. Linnhoff, E. Hindmarsh (1983), The pinch design method for heat exchanger networks, Chem. Eng. Sci. 38, 745-763.

X. Luo, Q.-Y. Wen, G. Fieg (2009), A hybrid genetic algorithm for synthesis of heat exchanger networks, Comput. Chem. Eng., doi:10.1016/j.compchemeng.2008.12.003.

F. Pettersson (2005), Synthesis of large-scale heat exchanger networks using a sequential match reduction approach, Comput. Chem. Eng., 29, 993-1007.

T.F. Yee, I.E. Grossmann, Z. Kravanja (1990), Simultaneous optimization models for heat integration - I. Area and energy targeting and modeling of multi-stream exchangers, Comput. Chem. Eng. 14(10), 1151-1164.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Simulation of distillation process in the bioethanol production using nonequilibrium stage model

Tassia L. Junqueira,[a] Rubens Maciel Filho,[a] Maria R. W. Maciel[a]

[a] *School of Chemical Engineering, State University of Campinas – UNICAMP, P.O. Box 6066, Campinas, SP 13083-970, Brazil*

Abstract

Most models available for simulation of multicomponent separation processes are based on the idealized concept of equilibrium or theoretical stages. However, accuracy of the predictions can be highly enhanced if a nonequilibrium stage model is taken into account. In this work, hydrous bioethanol production process was simulated using Aspen Plus® considering three calculation methods: equilibrium, equilibrium with constant plate efficiency and nonequilibrium stage models. Comparison of nonequilibrium and equilibrium model simulations, assuming the same number of stages, showed that energy consumption calculated is much larger when nonequilibrium model is considered and that column specifications have to be adapted in order to achieve bioethanol specification. Equilibrium model with efficiency of 70% presented a satisfactory agreement with nonequilibrium model.

Keywords: simulation, bioethanol, nonequilibrium model, distillation process

1. Introduction

Bioethanol has been increasingly used as substitute or additive to gasoline, since it is a renewable fuel and its combustion discharges less greenhouse gases when compared to fossil-derived fuels. In Brazil, second largest producer in the world, it is typically produced through fermentation of sugars derived from sugarcane.

In order to be used as a fuel, wine obtained from fermentation must be concentrated to about 93 wt% of ethanol (hydrous bioethanol), which requires distillation process. Besides, for bioethanol being used as gasoline additive, further processes are necessary, since water and ethanol form an azeotrope with 95.6 wt% ethanol at 1 atm. Thus, conventional distillation can not achieve anhydrous bioethanol specification (approximately 99.5 wt%). Azeotropic distillation with cyclohexane as entrainer, extractive distillation with monoethyleneglycol as solvent and adsorption onto molecular sieves are possible and usual alternatives for bioethanol dehydration.

Usually, simulations of distillation process consider the equilibrium stage model, although, in practice, columns rarely operate under thermodynamic equilibrium conditions. In 1985, Krishnamurty and Taylor described the nonequilibrium stage model of multicomponent separation processes and observed that equilibrium and nonequilibrium model provided results quite different from each other. In the nonequilibrium model, conservation equations are written for each phase independently and solved together with transport equations that describe mass and energy transfers in multicomponent mixtures; also it is assumed that equilibrium occurs only in the vapor-liquid interface. Besides, in this way, the empirical correcting factors, such as efficiencies used in the equilibrium model, are no longer necessary (Pescarini, 1996).

Several works have been developed regarding comparison between nonequilibrium stage model and experiments. For instance, Springer et al. (2002) realized experiments in a bubble cap distillation column with the system water – ethanol – methylacetate and also developed the nonequilibrium model, their comparison showed excellent agreement between the results. In adition, Eckert & Vaněk (2001) modelled three-phase distillation columns for ethanol dehydration with cyclohexane and observed that nonequilibrium model provided results in good agreement with experiments. Repke et al. (2004) carried out the validation of the developed nonequilibrium model in a pilot packed column and concluded that nonequilibrium model describes the experimental data with a good accuracy.

In view of the fact that distillation operations require a significant amount of energy and have a great importance in bioethanol production, the simulation of this unit operation has to be as representative as possible. In this work, simulations of distillation process in bioethanol production were carried out in Aspen Plus®, considering equilibrium model, equilibrium model with constant plate efficiency and nonequilibrium model.

2. Traditional distillation process

The distillation process simulated was based on the traditional Brazilian biorefineries configuration employed to produce hydrous bioethanol. Distillation takes place in a set of five columns, divided into distillation (A, A1 and D, each one at the top of each other) and rectification columns (B and B1). Columns B and B1 are located one above the other and have the same diameter, so they were simulated as a single column (46 stages). Since columns A (19 stages), A1 (8 stages) and D (6 stages) have different diameters, they were simulated detached. This configuration is depicted in Figure 1.

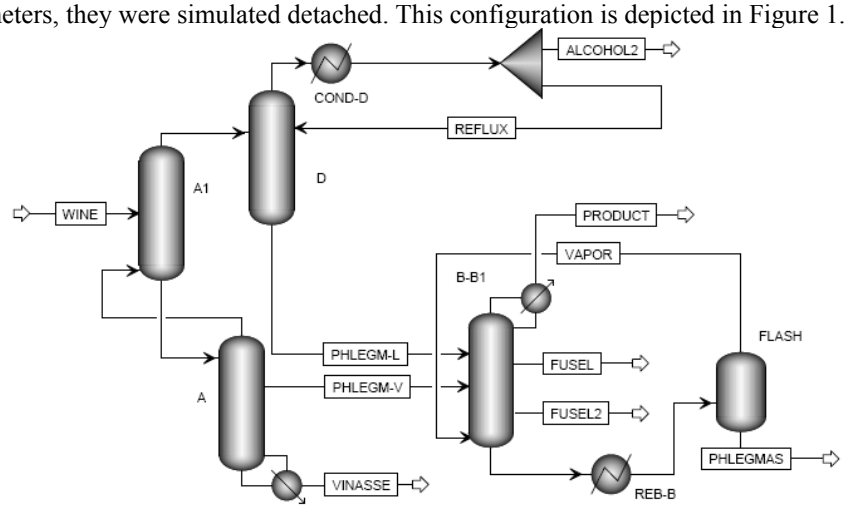


Figure 1. Configuration of the distillation process.

Wine, produced in the fermentation stage, is fed to column A1. Column D is responsible for removing volatile contaminants at the top, while a large amount of water (vinasse) is obtained in the bottom of column A. Vapour phlegm produced near to the top of column A and liquid phlegm obtained at the bottom of column D are sent to rectification columns. The task of the rectification consists on the concentration of the phlegms to 93 wt% ethanol (hydrous bioethanol). Phlegmasse, which has high contents of water, is produced in the bottom of column B-B1. Pressure drop in the columns was considered,

top/bottom pressures (kPa) adopted were 139.3/152.5, 136.3/139.3, 133.8/136.3 and 116.0/135.7 in columns A, A1, D and B, respectively.

3. Simulation procedure

Thermodynamic models adequacy to represent the system was evaluated by comparison between vapor-liquid equilibrium results given by Aspen Plus® and available experimental data (Gmehling and Onken, 1977). Aspen Plus® simulator supplies binary parameters from its databank and uses them automatically. This study analyzed NRTL and UNIQUAC activity coefficient models, PSRK equation of state and the UNIFAC method. NRTL was the model that provided better predictions, for this reason it was used to calculate the activity coefficients for the liquid phase.

Wine compositions, flow rate and temperature, given in Table 1, were based on biorefinery data. This data showed that other components, such as acetaldehyde, acetone, acetal, n-propanol, isobutanol, n-butanol and acetic acid, are also present in the wine, however they represent less than 0.1% in the composition and were disregarded in this work.

Table 1. Specification of wine fed to distillation process.

Variable	Value
Water (wt%)	92.0
Ethanol (wt%)	7.3
Glycerol (wt%)	0.4
Isoamyl alcohol (wt%)	0.2
Glucose (wt%)	0.1
Total mass flow rate (t/h)	206.9
Temperature (°C)	81.0

PHLEGMA-L and PHLEGMA-V inlet stages were the only feed positions that could be varied, since the other streams are fed to the first or last stage to substitute a missing condenser or reboiler, respectively. Therefore, optimization process was carried out in order to determine the best inlet stages concerning energy consumption. Three different approaches were used in the simulation of this distillation process regarding to column calculations: equilibrium model (EQ), which assumes as efficiency of 100%, equilibrium model with constant plate efficiency (EQ55, EQ70 and EQ85) and nonequilibrium model (NEQ). All simulations considered condenser and reboiler as an equilibrium stage.

In order to perform equilibrium model simulations, RadFrac model in Aspen Plus® simulator is used and its specification is quite simple, since number of stages and feed inlet positions are the only column design parameters required. In the equilibrium model with constant plate efficiency, RadFrac model is also utilized and a additional especification is introduced: plate efficiency. Therefore, it was specified that all stages have the same Murphree efficiency of 55, 70 and 85% in simulations EQ55, EQ70 and EQ85, respectively.

In the nonequilibrium model, which is a more rigorous calculation, software Aspen Plus® makes available RateFrac model. This model calculates the product of the binary mass transfer coefficients and interfacial areas using the correlations developed by Grester et al.; the vapor phase and liquid phase heat transfer coefficients are determined

using the Chilton-Colburn analogy. In general, these quantities depend on column diameter and operating parameters. Tray type was defined as bubble cap and column diameter was calculated by the simulator, which sizes the column based on the approach to flooding on the stage where it is most critical. Parameters used in calculations, such as vapor and liquid flows, densities, viscosities, surface tension of liquid, vapor and liquid phase binary diffusion coefficients, are all estimated by the simulator.

4. Simulation results and discussion

In order to obtain hydrous bioethanol (93 wt% ethanol), distillate rate on column A and vapor fraction on column B reboiler were manipulated. Equilibrium and nonequilibrium models had different values on these variables, since results using same specification are significantly different. Equilibrium with plate efficiency and nonequilibrium models assumed same values, since the purpose was to determine the efficiency value that most approximates nonequilibrium model predicitions.

Temperature profile in column B-B1 was analyzed aiming comparison between results given by the models. This column was chosen for this study because it has larger number of stages, thus it is the most influenced column in this process.

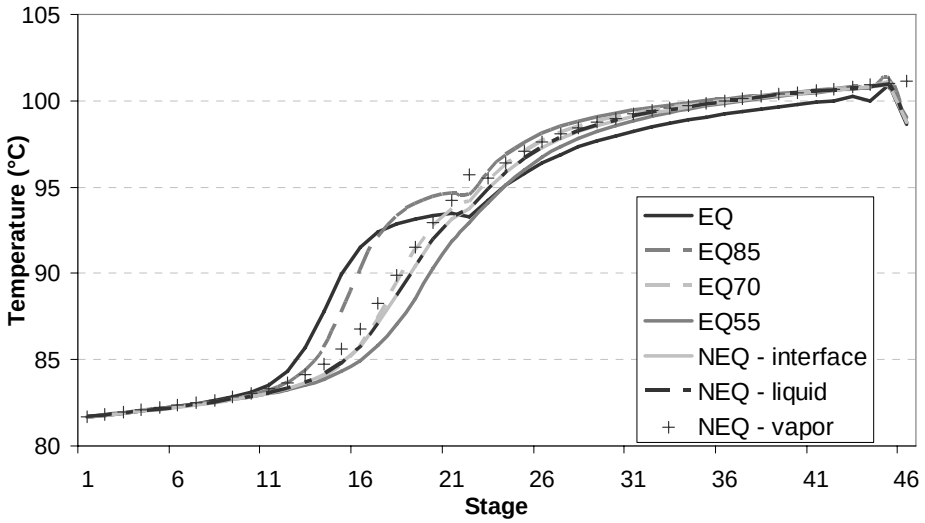


Figure 2. Temperature profiles for column B-B1.

Temperature profiles, depicted in Figure 2, showed that EQ results deviate considerably from the curves given by NEQ. Liquid and interface temperature profiles obtained in the NEQ model were coincident, indicating that there is no resistance to energy transfer between interface and liquid phase. However, vapor temperatures calculated by NEQ slightly diverged from the other temperatures, which means that vapor phase controls energy transfer in the system. Besides, analysing temperatures given by equilibrium model with constant plate efficiency (EQ55, EQ70 and EQ85), it can be observed that temperatures are higher for the superior efficiency (85%), since it means that system behavior is closer to ideal. Among simulations considering plate efficiency, taking NEQ simulation as base, EQ70 predicted better results since its curve is between vapor and liquid temperature profiles calculated by NEQ, which is an indicative that the value 70% is reasonable for efficiency in this process.

Main stream results for EQ, NEQ and EQ70 are displayed on Tables 2-4. It was verified that, in all simulations, hydrous bioethanol production was 16,100 kg/h and its concentration was around 93 wt% ethanol. Other streams also showed similar flow and composition in the simulations. It was also observed that ethanol losses on vinasse and phlegmasse are not significant in all simulations.

Table 2. Stream results for equilibrium stage model.

Stream	VINASSE	PHLEGMAS	ALCOHOL2	PRODUCT
Temperature (°C)	111.9	106.0	30.0	81.7
Mass Flow (kg/hr)	176,931	13,110	0.133	16,100
Water (wt%)	99.4	98.8	8.1	7.0
Glucose (wt%)	0.1	0.0	0.0	0.0
Ethanol (wt%)	0.0	0.1	91.9	93.0
Glycerol (wt%)	0.5	0.0	0.0	0.0
Isoamyl alcohol (wt%)	0.0	1.1	0.0	0.0

Table 3. Stream results for equilibrium stage model with efficiency.

Stream	VINASSE	PHLEGMAS	ALCOHOL2	PRODUCT
Temperature (°C)	111.9	106.3	30.0	81.7
Mass Flow (kg/hr)	175,419	14,622	0.133	16,100
Water (wt%)	99.4	99.0	10.6	6.9
Glucose (wt%)	0.1	0.0	0.0	0.0
Ethanol (wt%)	0.0	0.0	89.4	93.1
Glycerol (wt%)	0.5	0.0	0.0	0.0
Isoamyl alcohol (wt%)	0.0	1.0	0.0	0.0

Table 4. Stream results for nonequilibrium stage model.

Stream	VINASSE	PHLEGMAS	ALCOHOL2	PRODUCT
Temperature (°C)	111.9	106.0	30.0	81.7
Mass Flow (kg/hr)	175,347	14,693	0.133	16,100
Water (wt%)	99.4	98.8	14.9	7.0
Glucose (wt%)	0.1	0.0	0.0	0.0
Ethanol (wt%)	0.0	0.0	84.6	93.0
Glycerol (wt%)	0.5	0.0	0.0	0.0
Isoamyl alcohol (wt%)	0.0	1.1	0.5	0.0

Energy demand on columns reboilers for each simulation is shown in Table 5. It can be seen that equilibrium model calculates lower energy consumption when compared with

other calculation methods, because it assumes that vapor and liquid phases leave stages in equilibrium. Once mass and energy transfers or efficiency are considered, as in NEQ and EQ70 simulations, more stages and energy are required to achieve the same separation. It was also observed that energy demand was similar in EQ70 and NEQ, so 70% may be a good estimate for total efficiency in this process. In addition, values displayed in Table 5 indicate that the reboiler of column A required larger amount of energy than column B-B1, since column A operates with larger flows.

Table 5. Energy demand on hydrous bioethanol production.

Model	EQ	EQ70	NEQ
Column A reboiler (kJ/kg Product)	4503	4712	4723
Column B-B1 reboiler (kJ/kg Product)	1150	2252	2259
Total energy consumption (kJ/kg Product)	5653	6964	6982

5. Conclusions

Simulations of bioethanol production process were carried out considering equilibrium, equilibrium with plate efficiency and nonequilibrium models. Regarding to energy consumption and temperature profiles, equilibrium model gave different predictions from the other models studied. Equilibrium model with efficiency of 70% showed a satisfactory agreement with results obtained with nonequilibrium model, which is a rigorous calculation. It was also observed, from temperature profiles, that in the nonequilibrium model, liquid phase presents no resistance to energy transfer and vapor phase controls this process. In conclusion, nonequilibrium model simulation revealed to be a useful tool to predict column performance and to provide more accurate results.
Literature review showed that nonequilibrium model predicts results in good agreement with experimental results, however, further studies including experiments are recommended to validate nonequilibrium model for bioethanol production process.

6. Acknowledgments

The authors acknowledge FAPESP and CNPq for the financial support.

References

E. Eckert, T. Vaněk, 2001, Some aspects of rate-based modelling and simulation of three-phase distillation columns, Computers & Chemical Engineering, 25, 603–612

J. Gmehling, U. Onken, 1977, Vapor-Liquid Equilibrium Data Collection – DECHEMA Chemistry Data Series

R. Krishnamurthy, R. Taylor, 1985, A nonequilibrium stage model of multicomponent separation process – Part I: Model description and method of solution, AIChE Journal, 31, 3, 449-456

M. H. Pescarini, A. A. C. Barros, M. R. Wolf-Maciel, 1996, Development of a software for simulating separation processes using a nonequilibrium stage model, Computers & chemical engineering, 20, SUPA, 279-284

J. U. Repke, O. Villain, G. Wozny, 2004, A nonequilibrium model for three-phase distillation in a packed column: modelling and experiments, Computers & Chemical Engineering,28, 775–780

P.A.M. Springer, S. van der Molen, R. Krishna, 2002, The need for using rigorous rate-based models for simulations of ternary azeotropic distillation, Computers & Chemical Engineering, 26, 1265–1279

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling Electrodialysis and Photochemical Process for their integration in saline wastewater treatment

Fúlvia J. Borges[a], Hélène Roux-de Balmann[b]; Roberto Guardani[a]

[a]*LSCP/CESQ - Department of Chemical Engineering, Polytechnic School - University of São Paulo, Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil.*
[b]*Laboratoire de Génie Chimique – UMR CNRS 5503- Université Paul Sabatier, 31062 Toulouse Cedex 4, France*

Abstract

Oxidation processes can be used to treat industrial wastewater containing non-biodegradable organic compounds. However, the presence of dissolved salts may inhibit or retard the treatment process. In this study, wastewater desalination by electrodialysis (ED) associated with an advanced oxidation process (photo-Fenton) was applied to a NaCl in water solution containing phenol. The influence of process variables on the on the demineralization factor was investigated for ED in pilot scale. A phenomenological approach was used to relate the phenol, salt and water fluxes with the driving force. The oxidation process was investigated in a laboratory batch reactor and a model based on artificial neural networks was developed by fitting the experimental data describing the reaction rate as a function of the input variables. With the experimental parameters of both processes, a dynamic model was developed to ED and a continuous model, using a plug flow reactor approach, to the oxidation process. Finally, the hybrid resulting simulation could validate different scenarios of the integrated system and can be used for further optimization.

Keywords: process integration, modeling, electrodialysis, advanced oxidation process, wastewater

1. Introduction

Recalcitrant organic pollutants, like phenol and phenolic compounds, common contaminants in industrial wastewater, have been recognized as an issue of growing importance in recent years. Conventional wastewater treatments are hardly capable of removing such pollutants and oxidation processes are used for its degradation. However, the presence of ionic compounds, like sodium chloride, NaCl, normally associated with these pollutants in wastewater, may retard the process (Maciel, 2004). In this context, the possibility of integrating electrodialysis (ED) and advanced oxidation processes (AOP) for the treatment of a saline aqueous solution containing phenol has been investigated.

ED is an electrochemical separation process based on selective transport of ions from one solution to another through ion-exchange membranes under the driving forces of an electrochemical potential gradient. Typical ED equipment consists of a series of anion and cation-exchange membranes, that allow the selective passage of anion and cation,

respectively, arranged in an alternating pattern between an anode and a cathode. When ionic solutions are pumped through the cell the ion concentration increase in alternate compartments (*diluate*), while the other compartments become depleted (*concentrate*) (Strathmann, 2004). In particular, an aqueous saline solution containing phenol and water can be treated by ED resulting in a demineralized solution and a saline wastewater. In this manner, the salt can be removed prior to the wastewater treatment by oxidation processes as shown in Figure 1.

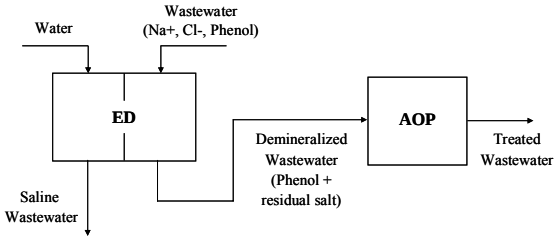


Figure 1 – Process integration schematic representation

AOP is considered a promising remediation of wastewaters containing recalcitrant compound like phenol. The processes involve the production of reactive hydroxyl radicals that can mineralize organic compounds. One of the most used AOP is the photo-Fenton process, which combines peroxide, usually hydrogen peroxide, iron ion and UV light irradiation (Esplugas, 2001).

The objective of this paper was to develop and study a hybrid model using the phenomenological parameters obtained experimentally for the demineralization and oxidation processes. Experiments were performed with synthetic saline solutions, with a concentration range of sodium chloride (NaCl) normally found in phenolic wastewater. The model will be used for further optimization.

2. Model Approach

2.1. Electrodialysis

The investigation of the desalination of water containing phenol and sodium chloride by ED was previously reported by Borges et al., 2008. Experiments were carried out in pilot equipment and a phenomenological approach was used to relate the phenol, salt and water fluxes with the driving forces (concentration and electric potential gradients) in the ED cell. Experiments were carried out at constant current, in a range between 75 and 275 A m^{-2}; the concentration of the salt and phenol in the *diluate* ranged from 0 to 55 kg m^{-3} and from 100 to 400 ppm C, respectively. Table 1 presents the flux equations and parameters obtained.

Table 1. Flux equations and overall transport coefficients obtained in the experiments with the ED system (Borges et al., 2008).

Water transport	**Salt transport**	**Phenol transport**	
$J_w = \alpha.i$	$J_s = \beta.i$	$J_p = P_p.\Delta C_p^0 + (1-\sigma).C_{pd}^0.J_w$	
α x10^{-6} (m^3 h^{-1} A^{-1})	β x 10^{-3} (kg s^{-1} A^{-1})	P_p x 10^{-4}(m h^{-1})	σ
3.74	1.98	3.02	0.24
S_1 = 0.4 m^2	S_2 = 0.2 m^2	S_1 = 0.4 m^2	

where J_W is the water flux (m h^{-1}); J_p and J_s, phenol and salt flux respectively (kg m^{-2} h^{-1}); α, electro osmotic coefficient (m^3 h^{-1} A^{-1}); β, current coefficient (kg h^{-1} A^{-1}); i, current density (A m^{-2}); S, membrane transfer area; P_p, phenol permeability coefficient (m h^{-1}); ΔC_p^0, initial phenol concentration difference between compartments, and C_{pd}^0, the initial phenol concentration in the feed (kg m^{-3}).

Salt and water transport were related to the current density applied. For the neutral solute, phenol, a convective contribution due to water flux was pointed out besides the expected diffusion contribution. The fitting of the parameters of the transport equations resulted in good agreement with the experimental results over the range of conditions investigated.

The equations presented in Table 1, which describe the flux within the ED cell, can be combined with mass balance and 10 ordinary differential equations to describe the system presented in Figure 2.

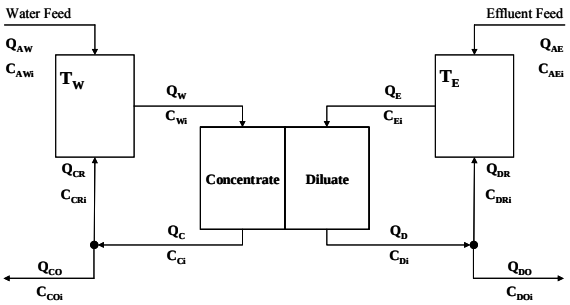


Figure 2 – Schematic representation of the ED system modeled; Q and C represent the different flow and concentration for the streams; subscript i: phenol and salt.

R_C and R_D represent the recycle fraction of Q_C and Q_D, as the following equations:

$$Q_{CR} = R_C Q_C \quad (1)$$

$$Q_{DR} = R_D Q_D \quad (2)$$

The phenol flux equation was simplified for the pilot experimental conditions; for this generic model, the former equation can be used as presented in equation (3) once the phenol concentration difference between compartments, ΔC_p, can be significant and calculated over time in this case.

$$J_p = P_p.\Delta C_p + (1-\sigma).\Delta C_p.J_w \quad (3)$$

For the second flux contribution, the original mean solute concentration on either side of the membrane was considered equal to ΔC_p.

The dynamic model can simulate the system flow, salt and phenol concentrations over time, and a steady state approach was obtained from this model.

2.2. Photochemical

Photo-Fenton oxidation processes were experimentally investigated using a laboratory batch reactor for the degradation of phenol in saline media. Experiments were carried in an annular photochemical reactor (1 L) connected to a well stirred jacketed gall

reservoir (1 L). The lamp (250W Philips medium-pressure mercury vapor lamp) was protected by a water-cooled quartz immersion well. The system was feed with a 100 ppm C of phenol saline solution and ferrous sulfate heptahydrate (Fe^{2+}). The solution was circulated through the reactor at a flow rate of about 1.5 L min^{-1} by means of a centrifugal pump. The reaction was initiated with the irradiation and onset addition of hydrogen peroxide (H_2O_2) solution at a rate of 0.83 mL min^{-1}. The pH was maintained at 3.0 and temperature at $30^oC \pm 2^oC$. Samples were withdrawn at appropriate time intervals. The organic concentration was determined by total Organic Carbon (TOC) analysis.

An experimental design based on the Doehlert model (De Aguiar et al., 1998) with three variables was used to study the effect of NaCl, H_2O_2 and Fe^{2+} concentrations on phenol degradation (variables ranges between 0 to 50 kg m^{-3}, 10 to 80 mM and 0.1 to 0.5 mM, respectively). These ranges were chosen from preliminary studies in the domain of interest.

As expected, the results confirm the negative effect of the salt concentration on the phenol removal efficiency by oxidation. Moreover, this effect was not found to be linear concerning salt concentration and degradation rate. Due to the complexity of these reaction systems, a model based on artificial neural networks has been developed to fit the experimental data (Gob et al., 1999). This model describes the evolution of the pollutant concentration i.e. phenol, by means of a reaction rate, during irradiation time under various operating conditions.

Figure 3 shows a diagram representation of the model used to obtain the reaction rate in function of the salt, Fe^{2+}, H_2O_2 concentration and TOC/TOC_0; where TOC_0 is the initial TOC value.

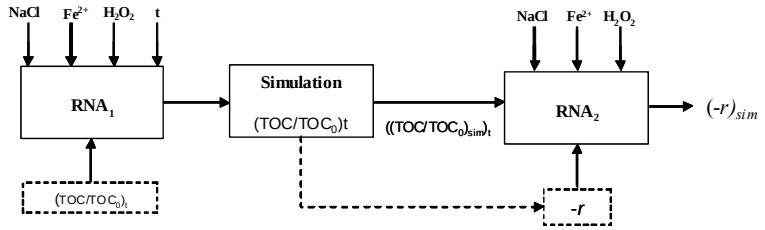


Figure 3 – Diagram representing the reaction rate mathematical model

Two artificial neural networks were used (RNA_1 and RNA_2). RNA_1 models the values of TOC/TOC_0 at a given time and condition, and RNA_2 calculates the reaction rate in function of TOC/TOC_0 value, NaCl, Fe^{2+} and H_2O_2 concentrations. The model obtained was in good agreement with experimental data.

A plug flow reactor was modeled using the reaction rate obtained by RNA_2, r, in an ordinary differential equation:

$$\frac{dTOC_{(z)}}{dV} = -\frac{r}{q} \tag{4}$$

where V is the reactor volume, z the distance down the tube from the inlet and q the flow.

The model can describe the TOC variation in a plug flow reactor in function of its volume and operation conditions, i.e. NaCl, Fe^{2+} and H_2O_2 concentrations.

3. Results and discussion

The process integration was carried out by setting the ED salt and phenol concentration output (C_{DO_S} and C_{DO_P} (Figure 2)) as the photochemical input. Several scenarios have been studied with the proposed hybrid model and some of the results obtained are summarized below.

Concerning the desalination, the model represents a generic ED system to study different configurations of the process. Firstly, the accuracy of the simulation prediction was validated comparing experimental results with a pilot configuration of the system; i.e. feed and outlet stream equal to zero (total recycle).

In terms of optimal operation conditions for the ED step, the objective is to minimize phenol loss in the saline wastewater, i.e. the concentrated stream, and minimize the cost of operation, related to energy and membrane area required. Figure 4 shows an example of how the phenol loss and the membrane area depend on the current density for 80 % demineralization of the effluent feed (Figure 4).

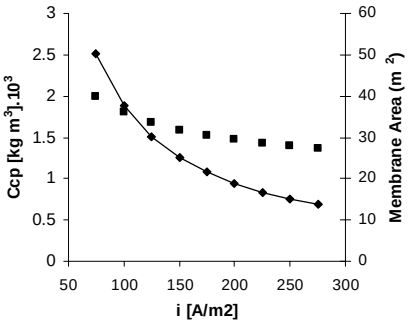


Figure 4 – Phenol loss (■) and membrane area (◆) vs. current density for 80% demineralization; C_{D_S} = 10 kg m^{-3}; C_{AE_P} =0.18 kg m^{-3}, C_{AE_S} =50kg m^{-3}, C_{AW_P} =0, C_{AWs} =0

As expected, for lower current density higher membrane area is required and higher phenol loss (C_{CO_P}) in the *concentrate* occurs in order to achieve a given amount of demineralization. The change in membrane area is more pronounced for lower current densities. The phenol loss is less sensitive to the current density than the membrane area.

On the other hand, the optimal condition for the photochemical step operation is related to the cost to demineralize the effluent to a certain level of salt concentration. The total cost is mainly the sum of reagents, energy and investment cost, in which energy is the most significant factor. Therefore, the cost can be associated directly with the reactor volume and its variation for different conditions can be obtained by the model. Figure 5 shows the reactor volume *versus* NaCl concentration for different values of H_2O_2 concentration for 70% removal of the initial TOC.

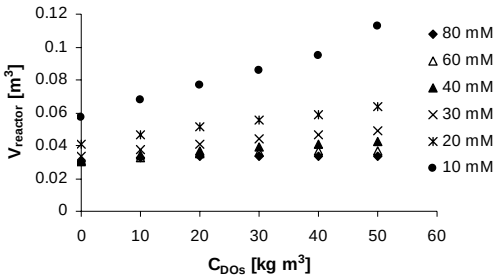


Figure 5 – Photochemical reactor volume vs. salt concentration feed for different H_2O_2 concentrations to 70% degradation; TOC_0= 0.1 kg m^{-3}; i = 275 A m^{-2}.

As can be observed for the conditions studied, the H_2O_2 concentration has an important effect on the process. The best oxidation conditions were attained when more than 30 mM H_2O_2 was added. Although values larger than 40 mM do not affect the degradation considerably, a mitigation of the NaCl negative effect can be observed for concentration above 30 kg m^{-3}. For 40 mM, the increase from 0 to 50 kg m^{-3} in salt concentration increases the reactor volume by ca. 40%, which negative significant information in economic feasibility studies.

4. Conclusions

A hybrid model in agreement with experimental data was obtained for the ED and photochemical processes. Simulations based on the model have been performed in order to estimate the behavior of the system under different conditions. The model constitutes an important tool in further studies aimed at assessing the feasibility of integrated effluent treatment processes, in which the costs of the ED and photochemical steps must be coupled in order to achieve a given level of demineralization and degradation.

Aknowledgements

The authors thank the Brazilian Federal Agency, CAPES, The National Counsel of Technological and Scientific Development, CNPq, and The State of São Paulo Research Foundation, FAPESP, for the financial support.

References

R. Maciel, G.L. Sant´Anna Jr., M. Dezotti, 2004, Phenol removal from high salinity effluents using Fenton´s reagent and photo-Fenton reactions, Chemosphere, 57, 711-719.

H. Strathmann, 2004, Ion-exchange membrane separation processes, Membrane Science and Technology Series, v.9, Elsevier.

S. Esplugas, J. Giménez, S. Contreras, E. Pascual, M. Rodríguez, Comparison of different advanced oxidation processes for phenol degradation, Water Research, 36, 1034-1042.

F. Borges, H. Roux-de Balmann, R. Guardani, 2008, Investigation of the mass transfer processes during the desalination of waster containing phenol and sodium chloride by electrodialysys, Jounal of Membrane Science, 325, 130-138.

P.F. De Aguiar, D.L. Massart, 1998, Experimental design, The Encyclopedia of Computational Chemistry, John Wiley & Sons, Chichester.

S. Gob, E. Oliveros, S.H. Bossmann, A.M. Braun, R. Guardani, C.A.O. Nascimento, 1999, Modeling the kinetics of a photochemical water treatment process by means of artificial neural networks, Chemical Engineering and Processing, 38, 373-382.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Evolutionary Multi-objective Optimization of an Activated Sludge Process

Rosana Kazuko Tomita[a], Song Won Park[a]

[a] *Department of Chemical Engineering - USP Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil.sonwpark@usp.br*

Abstract

Many real problems involve multiple objectives that are often conflicting and non comparable. Rare are the problems in which the objectives to be achieved (maximized or minimized) are single. Such problems are not adequately solved by using single objective function optimization techniques. Non comparable solutions are associated to the multi-objective optimization problems that represent the tradeoff among objectives. These solutions are named Pareto optimal solutions. In this work, a general review about different methods that deal with multi-objective optimization problems is presented. Special attention is given to evolutionary methods, detailing the NSGA II algorithm. This algorithm was implemented in Matlab to deal with crossover and mutation applied directly to real variables and constrained optimization problems. This method of evolutionary multi-objective optimization was applied in an activated sludge wastewater treatment system, in which effluent quality and operational cost are optimized. For this purpose, set-points of nitrate and dissolved oxygen concentrations were used as decision variables. The Pareto curves obtained at the end of 50 generations were analyzed. Finally, for illustration, one point in the Pareto curve was chosen to demonstrate the optimal set of values. In this study, the results obtained show the ability of the algorithm to achieve Pareto optimal solutions.

Keywords: NSGA II, Evolutionary Multi-Objective Optimization, Activated Sludge.

1. Introduction

Many real problems involve multiple objectives that are often conflicting and non comparable, and their solutions are associated to the multi-objective optimization problems, that represent the trade-off among the objectives, named Pareto optimal solutions. Most techniques to solve these problems comprise two phases: a multi-objective phase, which is more mathematical, and a subjective phase, namely decision-makers problem.

There are many methods to solve multi-objective problems. Very recently, the heuristic methods, such as those based on genetic algorithms, show increased application. Evolutionary algorithms are becoming an alternative to the classical optimization methods, because they permit to work with great search spaces, can generate the best compromise solution among several objectives using an unique run of optimization algorithm, besides not needing extra information, as function derivative. The evolutionary algorithms reproduce the natural evolution principles to drive the search towards an optimal solution. Such algorithms refer to a class of stochastic optimization methods that simulate the natural evolution process. According to Deb [1], the ability of an evolutionary algorithm to find multiple optimal solutions in one single simulation

run makes an evolutionary algorithm unique in solving multi-objective optimization problems. The first implementation of a multi-objective evolutionary algorithm occurred in the mid 1980s with the J. Schaffer studies, but only in the mid-1990s that the evolutionary multi-objective optimization take shape, with the development of several important algorithms [2]. Among them, the multiobjective genetic algorithm (MOGA), developed by Fonseca and Fleming [3]; non-dominated sorting genetic algorithm (NSGA), by Srinivas and Deb [4]; and the niched Pareto genetic algorithm (NPGA) elaborated by Horn et al. [5]. Also Deb et al. [6] proposed the elitist non dominated sorting genetic algorithm (NSGA II) so as to alleviate three principal criticisms of the NSGA: computational complexity, non-elitist approach and the need of specifying a sharing parameter.

2. Evolutionary Multi-Objective Method

The elitist non dominated sorting genetic algorithm, denoted as NSGA II, incorporates the concept of elitism and use an explicit diversity preserving mechanism to make it more powerful than the previous NSGA algorithm. The NSGA II was proposed to alleviate three difficulties presented in NSGA: computational complexity, non-elitist approach, and the need for specifying a sharing parameter. The flowchart for NSGA II used here is shown in Figure 1. The blocks on right of this figure indicate our Matlab functions implemented for this work.

SBX – *Simulated Binary Crossover*

Create two solutions $x_i^{(1,t+1)} = 0{,}5\left[\left(1+\beta_{qi}\right)x_i^{(1,t)} + \left(1-\beta_{qi}\right)x_i^{(2,t)}\right]$ and $x_i^{(2,t+1)} = 0{,}5\left[\left(1-\beta_{qi}\right)x_i^{(1,t)} + \left(1+\beta_{qi}\right)x_i^{(2,t)}\right]$ after two parental solutions $x_i^{(1,t)}$ e $x_i^{(2,t)}$. This operator simulates recombination operator in one point for binary variables: first it generates the random number $0 \le u_i \le 1$. The parameter $\beta_{qi} = \left(2u_i\right)^{1/(\eta_c+1)}$ *when* $\mathrm{u_i} \le 0.5$ and $\beta_{qi} = \left(2(1-u_i)\right)^{-1/(\eta_c+1)}$ *when* $0.5 < \mathrm{u_i} \le 1$ is calculated so that its probability area is the number value u_i. The η_c is the recombination distribution index. The more elaborated version of this SBX was implemented in the present paper.

Polinomial Mutation

Apply the distribution of the polynomial probability to create a solution $y_i^{(1,t+1)} = x_i^{(1,t+1)} + \overline{\delta}_i \Delta_{\max}$. The parameter is estimated as $\overline{\delta_i} = \left(2u_i\right)^{1/(\eta_m+1)}$ *when* $\mathrm{u_i} \le 0.5$ and $\overline{\delta_i} = 1 - \left[2(1-u_i)\right]^{1/(\eta_m+1)}$ *when* $0.5 < \mathrm{u_i} \le 1$

where η_m is the index for the mutation probability distribution and $\Delta_{\max}$ is the maximum allowed perturbation. We implemented the more elaborated version of this parameter estimation

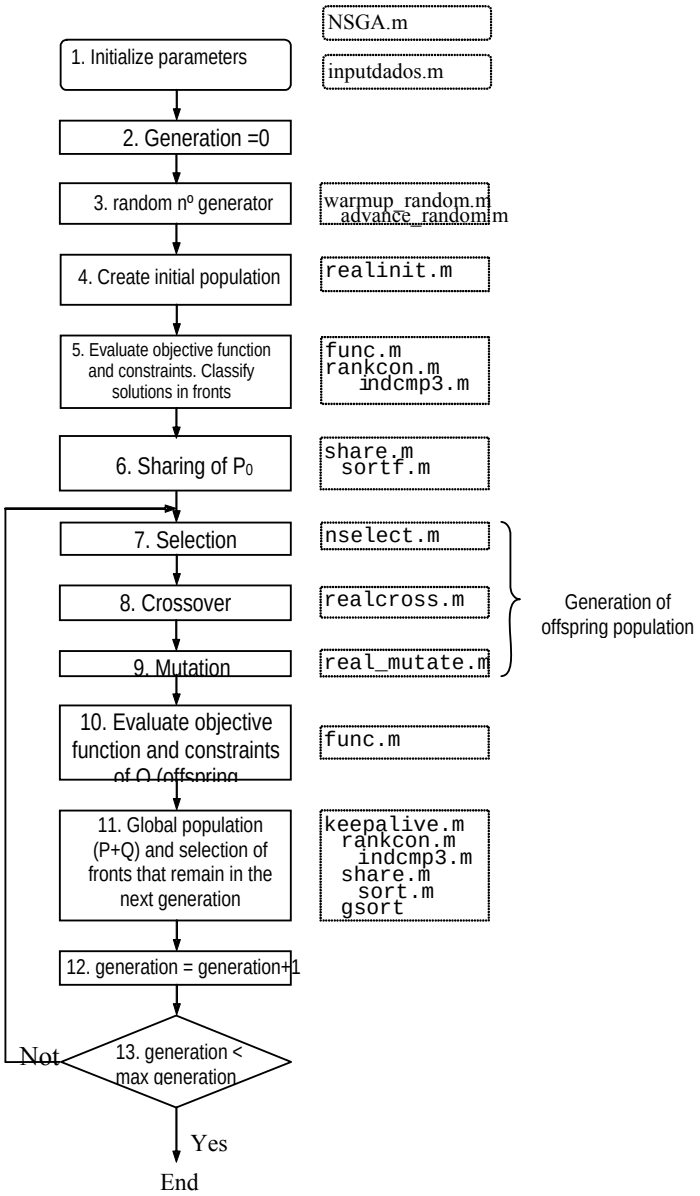


Figure 1: Flowchart for NSGA II

3. Evolutionary Multi-objective Optimization of the Activated Sludge Process

The wastewater treatment system dynamics was modeled in Simulink, see Sotomayor et al. [7-9] for details. The multi objective functions to be minimized are the Kjeldahl nitrogen EQ_{NK} , the nitrate EQ_{NO} and the energy required for aeration AE .

$$\min \begin{cases} EQ_{NK} = \frac{1}{T}\int S_{\mathrm{NK,e}}(t)\cdot Q_e(t)dt \\ EQ_{NO} = \frac{1}{T}\int S_{\mathrm{NO,e}}(t)\cdot Q_e(t)dt \\ AE = \frac{1}{T}\int \sum_{i=2}^{3}\left(0{,}4032(K_{\mathrm{LA}})_i^2 + 7{,}8408(K_{\mathrm{LA}})_i\right)dt \end{cases} \tag{1}$$

with the restraints

$$\begin{aligned}
SS_{\max} - SS &\geq 0 \\
DQO_{\max} - DQO &\geq 0 \\
N_{total,\max} - N_{total} &\geq 0 \\
DBO_{\max} - DBO &\geq 0 \\
\frac{dX_{\mathrm{B,H}}}{dt} &= \frac{F}{V}\left(X_{\mathrm{B,H}_{entra}} - X_{\mathrm{B,H}}\right) + r_1 + r_2 - r_4 \\
\frac{dX_{\mathrm{B,A}}}{dt} &= \frac{F}{V}\left(X_{\mathrm{B,A}_{entra}} - X_{\mathrm{B,A}}\right) + r_3 - r_5 \\
\frac{dX_{\mathrm{S}}}{dt} &= \frac{F}{V}\left(X_{\mathrm{S}_{entra}} - X_{\mathrm{S}}\right) + (1 - f_p)(r_4 + r_5) - r_7 \\
\frac{dX_{\mathrm{P}}}{dt} &= \frac{F}{V}\left(X_{\mathrm{P}_{entra}} - X_{\mathrm{P}}\right) + f_{\mathrm{P}}(r_4 + r_5) \\
\frac{dX_{\mathrm{ND}}}{dt} &= \frac{F}{V}\left(X_{\mathrm{ND}_{entra}} - X_{\mathrm{ND}}\right) + (i_{\mathrm{XB}} - i_{\mathrm{XP}} f_{\mathrm{P}})(r_4 + r_5) - r_8 \\
\frac{dS_{\mathrm{S}}}{dt} &= \frac{F}{V}\left(S_{\mathrm{S}_{entra}} - S_{\mathrm{S}}\right) - \frac{1}{Y_H}(r_1 + r_2) + r_7 \\
\frac{dS_{\mathrm{ND}}}{dt} &= \frac{F}{V}\left(S_{\mathrm{ND}_{entra}} - S_{\mathrm{ND}}\right) + r_8 - r_6 \\
\frac{dS_{\mathrm{NH}}}{dt} &= \frac{F}{V}\left(S_{\mathrm{NH}_{entra}} - S_{\mathrm{NH}}\right) - i_{\mathrm{XB}}(r_1 + r_2) - \left(i_{\mathrm{XB}} + \frac{1}{Y_{\mathrm{A}}}\right)r_3 + r_6 \\
\frac{dS_{\mathrm{NO}}}{dt} &= \frac{F}{V}\left(S_{\mathrm{NO}_{entra}} - S_{\mathrm{NO}}\right) - \frac{1 - Y_{\mathrm{H}}}{2{,}86Y_{\mathrm{H}}} r_2 + \frac{1}{Y_{\mathrm{A}}} r_3 \\
\frac{dS_{\mathrm{O}}}{dt} &= \frac{F}{V}\left(S_{\mathrm{O}_{entra}} - S_{\mathrm{O}}\right) - \frac{1 - Y_{\mathrm{H}}}{Y_{\mathrm{H}}} r_1 - \frac{4{,}57 - Y_{\mathrm{A}}}{Y_{\mathrm{A}}} r_3 + K_{\mathrm{LA}}\left(S_{\mathrm{O}_{sat}} - S_{\mathrm{O}}\right) \\
\frac{dS_{\mathrm{ALK}}}{dt} &= \frac{F}{V}\left(S_{\mathrm{ALK}_{entra}} - S_{\mathrm{ALK}}\right) - \frac{i_{\mathrm{XB}}}{14}(r_1 + r_2 + r_3) + \frac{1 - Y_{\mathrm{H}}}{40{,}4Y_{\mathrm{H}}} r_2 - \frac{1}{7Y_{\mathrm{A}}} r_3 + \frac{1}{14} r_6
\end{aligned} \tag{2}$$

For the details of this activated sludge ASM1 model with the explanation of the reaction rates $r_1,...,r_8$, see Tomita et al. [10]. The parameters applied for this problem were: the size of the population 100, number of iterations 50, recombination probability 0.9, mutation probability 0.1, recombination distribution index 1, mutation distribution index 1, $NO^{sp}_{\min} = 0\ g.m^{-3}$, $NO^{sp}_{\max} = 3\ g.m^{-3}$, $OD^{sp}_{2,\min} = 0\ g.m^{-3}$, $OD^{sp}_{2,\max} = 6\ g.m^{-3}$, $OD^{sp}_{3,\min} = 0\ g.m^{-3}$, $OD^{sp}_{3,\max} = 6\ g.m^{-3}$ For the control of this activated sludge wastewater treatment, see Sotomayor et al. [11]. The simple and partial results of the evolutionary multi-objective optimization are shown in Figure 2.

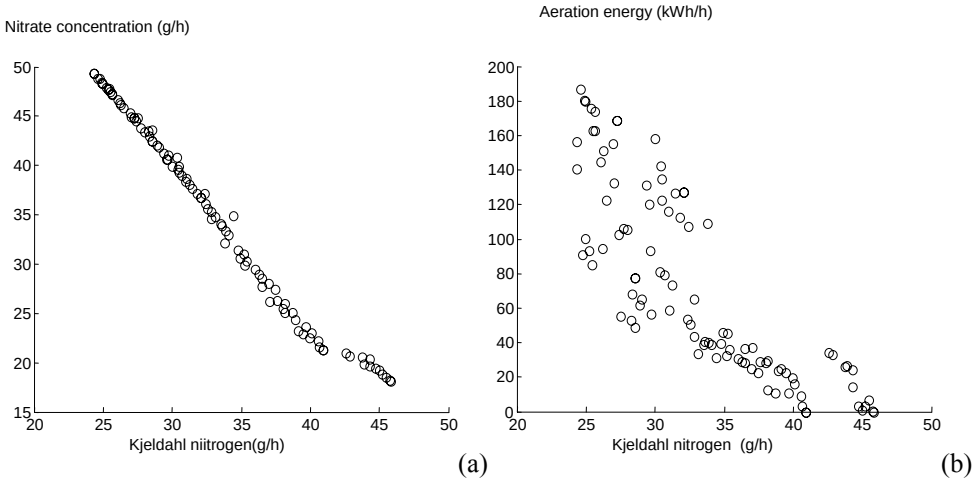


Figure 2. (a) nitrate concentration versus Kjeldhl nitrogen (b) aeration energy Kjeldhl nitrogen

In the ASM1 activated sludge model the process is highly nonlinear due to the kinetics similar to Monod. In Figure 3-b, it is observed that the convergence of the evolutionary optimization is not finished. However, these results are good enough to illustrate the methods of multi-objective optimization. In our complete work [12] the sensitivity analysis of the optimization results was performed due to parameters of the NSGA II algorithm, but this is not the main scope of the present paper.

4. Concluding remarks

The main contribution of the paper is the Matlab implementation of the NSGA II with more efficient and flexible programming for chemical engineering problems. Here, it can be integrated with a dynamic process model. This is illustrated with an ASM1 model implemented in Simulink. The usual implementation of NSGA II considers the maximum iteration as one of the stopping criteria, with no guarantee of optimization convergence. More efficient convergence criteria for chemical process applications have to be developed. Another issue for further development is the presentation of the clear decision-maker criteria for the upper layer of the multi-objective applications. The evolutionary algorithm also deserves more developments. Finally, after this experience,

the authors prefer to work, in the future, with differential evolutionary studies rather than adaptive and co-evolutionary algorithms.

References

1. Deb K (2001). *Multi-objective optimization using evolutionary algorithms*. John Wiley & Sons: New York
2. Coello CAC (2003). Guest editorial: special issue on evolutionary multiobjective optimization. *IEEE Transactions on Evolutionary Computation* **7**(2) 97-99.
3. Fonseca CM and Fleming PJ (1993). Genetic algorithms for multiobjective optimization: formulation, discussion and generalization. In: Kaufmann M (ed) Proceedings of the 1993 International Conference On Genetic Algorithms. San Mateo, pp. 416-423.
4. Srinivas N and Deb K (1994). Multiobjective optimization using nondominated sorting in genetic algorithms. *Evolutionary Computation* **2**(3): 221-248.
5. Horn J, Nafpliotis N and Goldberg DE (1993). *Multiobjective optimization usingthe niched pareto genetic algorithm***.** Urbana-Champaign, Illinois: Genetic Algorithms Laboratory, University of Illinois. (IlliGAL Report, 930051993).
6. Deb K, Pratat A, Sameer A and Meyarivan T (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* **6** (2): 182-197.
7. Sotomayor, O.A.Z.; Park, S.W.; Garcia, C (2003). Multivariable identification of an activated sludge process with subspace-based algorithms. *Control Engineering Practice*, v. 11, n. 8, p. 961-969
8. Sotomayor, O.A.Z.; Park, S.W.; Garcia, C (2002). Software sensor for on-line estimation of the microbial activity in activated sludge systems. *ISA Transactions* v. 41, p. 127-143.
9. Sotomayor, O.A.Z.; Park, S.W.; Garcia, C (2001). A Simulation Benchmark to Evaluate the Performance of Advanced Control Techniques in Biological Wastewater Treatment Plants. *Brazilian Journal of Chemical Engineering*, v. 18, n. 01, p. 81-101.
10. Tomita RK, Park SW and Sotomayor OAZ (2002). Analysis of activated sludge process using multivariate statistical tools – a PCA approach. *Chemical Engineering Journal* **90**: 283-290.
11. Sotomayor, O.A.Z.; Park, S.W.; Garcia, C (2002).A reference model for evaluating control strategies in activated sludge wastewater treatment plants. *Revue Des Sciences de L'eau*, v. 15, n. 2, p. 543-556.
12. Rosana Kazuko Tomita (2004). *Multi-objective Optimization of Wastewater Treatment Systems*.PhD Thesis. University of São Paulo. 207p. Brazil. (in portuguese) Otimização Multiobjetivo de Sistemas de Tratamento de Efluentes . 2004. 207 f. Tese de Doutorado em Engenharia Química - Universidade de São Paulo, Orientador: Song Won Park.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Evaluation of Alternative Methods for the Determination of Gasohol Fuel Properties when Compared to Standard Methods

J. F. Padilha[a]; R. W. S. Pessoa[b]; J. G. A. Pacheco[c] ; P. R. B. Guimarães[a]

[a]*Department of Engineering and Architecture – DEAR, Salvador University – UNIFACS, Av. Cardeal da Silva, n.132, CEP 40.231-250, Salvador – BA, Brazil, E-mail: paulorbg@unifacs.br*
[b]*Department of Chemical Engineering, Polytechnic School – Federal University of Bahia – UFBA, CEP 40.210-630, Salvador – BA, Brazil, E-mail: robsonpessoa2003@hotmail.com.br*
[c]*Department of Chemical Engineering, Federal University of Pernambuco – UFPE, Recife PE, Brazil, E-mail: Geraldo.ufpe@gmail.com*

Abstract

This study had the objective to evaluate alternative methods for determining properties of commercial gasoline, based on models that establish correlations between their properties and mid infrared spectra (Mid-IR) for spectra obtained with medium resolution portable equipment or obtained with high resolution equipment. Different methods of identifying outliers and pre-processing spectral data were analyzed. Methods for sample classification of the training set of data, model validation and model prediction, based on either Mahalanobis distance criteria or soft independent modelling of class analogy method (SIMCA), showed similar results. Models based on multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS) showed similar deviations for prediction of gasoline properties. Techniques for data pre-processing by derivative and smoothing did not improve robustness of the models. Models based on spectra obtained in high resolution infrared equipment showed acceptable deviations within the limits of reproducibility of the primary method, with less than 10% error. Models based on spectra obtained in medium resolution equipment showed errors higher the limits of reproducibility of the primary method in 30% of the samples.
Keywords: gasoline; infrared spectroscopy; octane number

1. Introduction

In order to ensure the quality of automotive gasoline offered to the final consumer, frequent collecting of samples is needed for analyzing their quality properties according to ASTM methods. Among the parameters used for evaluating the quality of gasoline are mainly anti-knock characteristics, distillation, density, aromatics content and for Brazilian gasoline, the ethanol content. The anti-knock characteristics are expressed by indexes of octane “motor” (MON) and “research" (RON), described by the methods ASTM D-2699-94 and ASTM D-2700-94.

These properties are determined by methods of analysis that require long time of execution, relatively high costs and large amount of samples, generating waste residues. The use of spectroscopy in the range of near infra red (NIR), proposed by Kelly et al. (1990) and Litani et al. (1997), and in the range of mid infrared (Mid-IR), studied by López et al. (1996) and Andrade et al. (1997), stands out among the alternative methods to estimate the properties of gasoline in a faster and more economical way. The use of infrared spectroscopy based on the Fourier transform (FTIR) has been proposed by Brereton (1986), Williams et al. (1989) and Giancaspro (1983), using chemometric algorithms for processing data and for calibration and validation of models.

Brazilian gasoline has some peculiar characteristics because of the content of 20% to 25% v/v of bio-ethanol and different composition of hydrocarbons in each state of the

country. These characteristics lead to the need to develop models and correlations that are able to predict diversity of gasoline properties and provide monitoring of the final product quality using fast and economical analysis.
Infrared spectroscopy-based methods are acceptable for quality control in a production line, where the samples do not have wide range of composition. Even with the use of high resolution IR equipment, models based on infrared have not been recognized as reliable for classification of commercial gasoline samples to be predicted by the model. The objective of this study was to evaluate the usefulness and limitations of alternative methods of analysis based on infrared spectroscopy for estimation of properties of commercial samples of gasoline collected at random from a macro-region of Brazil. Influence of different methods for detecting outliers and calibrating models on the robustness of the models were studied, especially for portable infrared equipment with medium resolution.

2. Experimental

The wide range of variation in composition of commercial gasoline in Brazil makes difficult to obtain generalized models. It is necessary the use of fuel samples from various regions of the country and require a statistical treatment of results for sample classification and model calibration and prediction. Database included 160 samples obtained from commercial gasoline, collected randomly at different gas stations along the Northeast region of Brazil. Half of these samples contained average volume fraction of 20% of ethanol and the other half contained a volume fraction of 25%. Samples were divided randomly into two groups, 107 of them are reserved for calibration or training models and 53 for model validation. Models were applied to predict properties of all samples for evaluating errors of classification. Infrared spectra were generated by two types of equipment, with high resolution measurements and another one with medium resolution measurements. Those data were treated separately generating two sets of models for comparison performance of models. For each set of models, various methods of detection and exclusion of outliers, parameters estimation and classification of validation data were tested comparing calculated errors.
Analyses of octane index were done in a CFR engine from Waukesha Motors. The analyses of composition of the samples were performed on a Varian CP-3800 gas chromatograph with a 100 meter PONA column. Infrared spectra were obtained with a portable equipment for analysis on field from Grabner Instrumensts IROX-2000, Mid-IR range, resolution of 4 cm^{-1}. For comparison purposes, infrared measurements were done from a high resolution spectrophotometer Perkin-Elmer Spectrum GX100, Mid-IR range, resolution of 8.3 cm^{-1}.
For the pre-processing and multivariate analysis of the infrared spectra, software Unscrambler, version 8.0 from CAMO, was used. Classification was done by comparing different methods for detection of outliers for refining the calibration set of data. Verification of clustering was done by principal component analysis (PCA) of the original spectral data, using the Unscrambler software, followed by graphic analysis of residual variances and leverages of the calibration set. Classification of samples for model validation was done by the estimative of the Mahalanobis distances (MD) from data after PCA analysis, using the toolbox of Matlab, and by soft independent modelling of class analogy method (SIMCA), based on Euclidean distance, as described by Wold et al. (1981). Each method of outlier detection generated a different set of calibration data that was used to compare models. After exclusion of possible outliers in the calibration set of data, models were trained based on the original spectra with parameter estimation for three different methods: multiple linear regression (MLR), principal component regression (PCR) and partial least squares (PLS). Alternatively, it was tested the influence of pre-treatment of spectral data by techniques of either smoothing or derivatives on the models performance. In the validation step, models built for each set

of data, from medium and high resolution infrared equipment were compared by analysis the errors of prediction of properties. In order to compare models performance, their predicting errors were compared to the reproducibility limits of each primary method used for determining gasoline properties. Reproducibility were also analyzed by calculation of root-mean-square error of calibration (RMSEC) and root-mean-square error of prediction (RMSEP) to evaluate the quality of the models and the method of classification of outliers, both the for calibration set of data for validation set of data for samples selected as adequate to be predicted by the models.

3. Results and Discussion

Figure 1 shows graphical representation of predicted data by the model vs. measured data for calibrated model of octane index RON by PLS method, when using complete spectra of calibration set of data selected by SIMCA method. In this graph predicted values of octane RON by the model were plotted as a function of measured values by the primary method for accepted and rejected samples in the SIMCA selection method. It is observed that from 109 calibration samples, 19 were rejected by the SIMCA method. Figure 2 shows results for predicted RON by PLS model vs. measured values, however, when samples of the calibration set were selected by the Mahalanobis distances (MD) selection method. In this case, from 109 calibration samples, 21 were rejected in the selection of MD. It is observed from Figures 1 and 2, that results of outlier selection are similar, although the two methods are based on different statistical principles, thus resulting in numbers of rejected samples not much different. Figures also show that some samples rejected by either SIMCA or MD criterion were well represented by the model from which were excluded. On the other hand, some samples recommended for use in the calibration were not well represented by the model from which they were used to calibrate. Based on these results, 19 samples were rejected, leaving 88 samples for the calibration set. Calibration set of data was examined for detection and exclusion of outliers by Mahalanobis distance method (MD), principal component analysis (PCA) and SIMCA, using the matrix of original 962 spectral data.

Figure 3 shows Chi-Square distribution of Mahalanobis distances of the refined calibration set of 88 samples. The distribution can set the critical value of 13.57 for the Mahalanobis distance that contains 95% of the samples. Validation samples with MDs above this value of 13.57 were selected as poorly represented by the models calibrated with this set. Figure 4 shows the comparison of three methods for selection of the calibration set. In this figure it is plotted residuals variance in function of the leverages for the first principal component PC1 from the PCA analysis of 88 samples of calibration set of data and 9 samples of validation set that were rejected by any of the three procedures. Among the three methods tested, 9 samples were rejected for the prediction, but there was no coincidence and only 7 of them were both disqualified by all methods. It is concluded that the Mahalanobis distance is not much different from the method based on either SIMCA or the graphical analysis method.

PLS and PCR models for prediction of RON index were tested with 53 samples for validation set of data, as shown respectively in Figures 5 and 6. These figures show results of prediction of RON on the basis of measured values for validation set for samples selected by the Mahalanobis distances, with 8 samples rejected and 45 samples selected as representative, from which 22 samples with a volume fraction of 20% of ethanol and 23 samples with a volume fraction 25% of ethanol in gasoline.

These subgroups are distinguished especially in the PCR model, both in calibration and in model prediction, as shown in Figure 6, which indicates that good results could be obtained with separate models for each of these groups. The performances of the models PLS and PCR are very similar.

Table 1 shows the overall comparison of different options for modeling of octane index RON, with data generated by portable infrared equipment and also the consequences of

using either the Mahalanobis distance or the SIMCA method for selecting samples of the prediction. It is also shown the influence of transforming the spectra by derivative treatment. For the spectra generated with portable equipment, the quality of models is very similar for any of the methods of sample selection (SIMCA or MD) or method of modeling (MLR, PCR or PLS). Pre-treatment of spectra by derivation or smoothing does not alter the quality of prediction. The values of Root Mean Squared Error of Calibration Set (RMSEC) and Root Mean Squared Error of Prediction Set (RMSEP) were generally above 0.7. These values are relatively high when compared to the limits of reproducibility of the primary method. The limits of reproducibility cited in ASTM D-2700 and D-2699 are 0.9 to 0.7 for RON and MON.

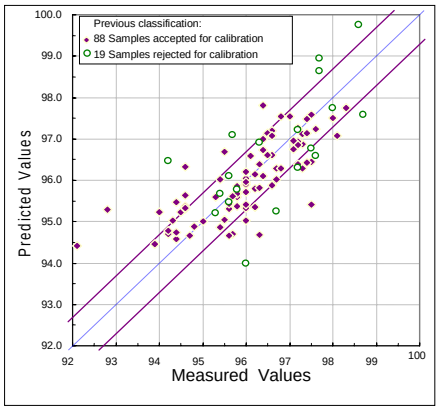


Figure 1 – RON, by PLS model from 962 spectra, 3 PCs, RMSEC = 0.75

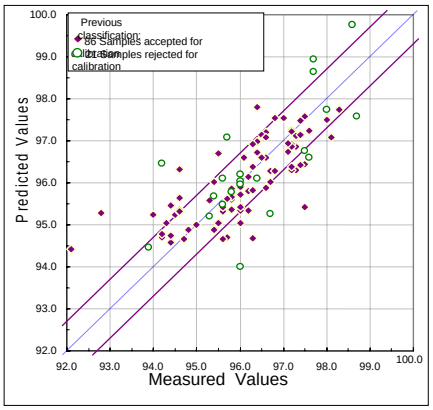


Figure 2 – RON, by PLS model from 962 spectra, 3 PCs, RMSEC = 0.76

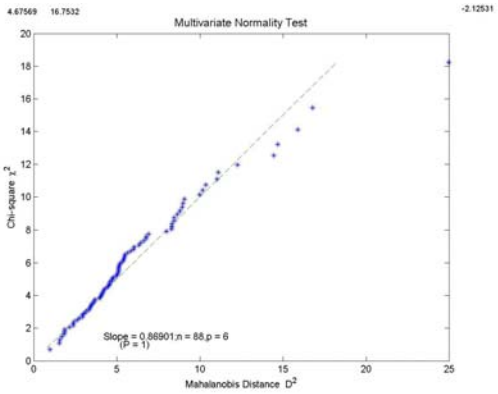


Figure 3 – Chi^2 Distribution for 88 refined samples of the calibration set.

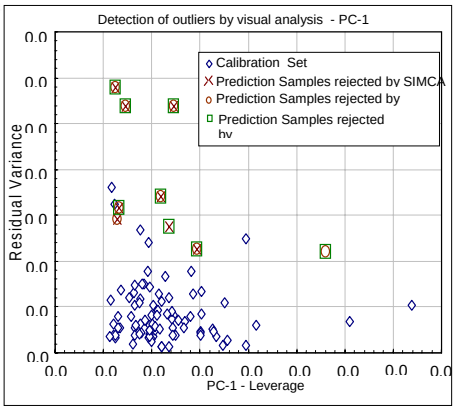


Figure 4 – Residual variances in function of Leverages for PC1.

The errors in the classification of samples are also considerable. Regardless of the method used, 15 to 30% of samples accepted in the calibration as suitable for fitting the model were represented by the model with errors above the limit of reproducibility of the primary method. From 20 to 30% of samples of the validation set and accepted as represented by the model were also predicted with values of RON with errors above the limit of reproducibility of the primary method.

For the models of MON prediction, results are similar to the RON, but the values of RMSEC and RMSEP from models of MON, ranging from 0.6 to 0.7, are lower than values for models of RON. The percentage of wrongly classified samples as represented by the models decreases slightly models of MON, reaching values of around 20%. These values, however, may be considered high for prediction purpose because the

uncertainty on the acceptance of prediction of octane MON data with this portable equipment can be up to 30% of cases.

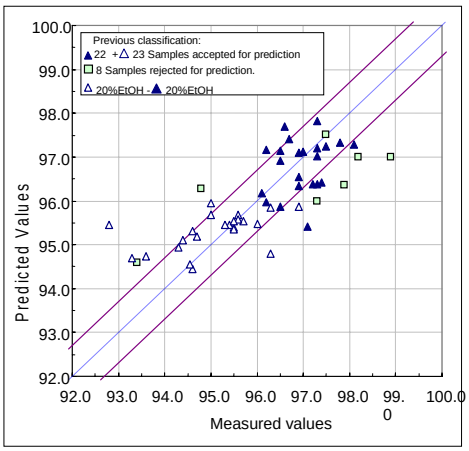


Figure 5 –RON, by PLS model from 962 spectra, 3 PCs, RMSEP = 0.79

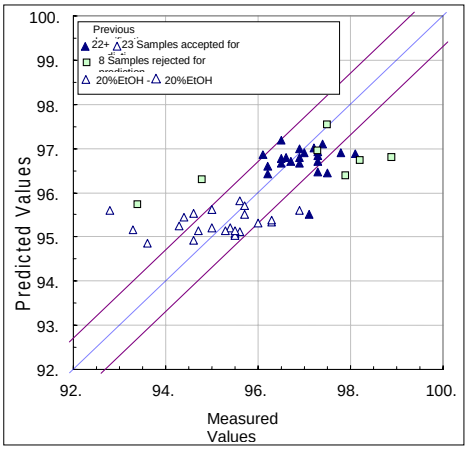


Figure 6 – RON, by PCR model from 962 spectra, 3 PCs, RMSEP = 0.82.

Table 1 - Comparison of models and methods of classification of samples for prediction of octane index RON. Infrared spectra via portable medium resolution equipment.

Spectra	**26 pts.**	**Complete**		**962**	**points**
Pre-treatment	**Max.**	**No**	**No**	**Deriv.**	**Deriv.**
Calibration Method	**MLR**	**PLS**	**PCR**	**PLS**	**PCR**
Classification Method	**SIMCA**	**SIMCA**	**SIMCA**	**SIMCA**	**SIMCA**
RMSEC	0.65	0.75	0.89	0.74	0.8
RMSEP	0.81	0.80	0.83	0.88	0.91
%errors of classification Calibration SET	14.8	19.3	27.3	14.8	18.2
%errors of aceptance Pediction SET	27.3	25	29.5	20.5	29.5
Classification Method	**MD**	**MD**	**MD**	**MD**	**MD**
RMSEC	0.65	0.75	0.89	0.74	0.8
RMSEP	1.07	0.79	0.82	0.9	0.92
%errors of classification Calibration SET	14.8	19.3	27.3	14.8	18.2
% errors of aceptance of Prediction SET	28.9	24.4	28.9	22.2	31.1

Analyzing spectra generated by the equipment of high resolution corresponding to medium infrared (Mid-IR), the classification of samples led to the identification of only 5 possible outliers among the 107 samples of the initial set of calibration. Figures 7 and 8 show the results of prediction of RON by PLS and PCR, with classification by SIMCA. Table 2 shows the comparison of different options for modeling RON.

Table 2 - Comparison of models and methods of classification of samples for prediction of octane index RON. Infrared spectra via high resolution infrared equipment.

Spectra	**23 pts.**	**Complete**		**3212**	**points**
Pre-treatment	**Max.**	**No**	**No**	**Deriv.**	**Deriv.**
Calibration Method	**MLR**	**PLS**	**PCR**	**PLS**	**PCR**
RMSEC	0.34	0.40	0.40	0.39	0.38
RMSEP	0.45	0.42	0.42	0.39	0.40
%errors of classification Calibration SET	4.9	4.9	6.9	4.9	5.9
% errors of acceptance of Prediction SET	6.1	10.2	10.2	6.1	8.2

Comparing the results shown in Figures 5 and 7, which refer to models based on spectra obtained with different equipment but with the same methods of calibration of models and classification of samples, and Tables 1 and 2 for the other methods, it can be checked improved quality of the models when the data of the same spectral range are generated by high resolution equipment.

For the prediction of RON, for example, levels of RMSEC and RMSEP with the use of models based on spectra generated by the equipment with a high resolution of 8.3 cm-1

is reduced to about 0.4, about half of those values obtained from the application of models based on spectra generated with the medium resolution equipment of 4 cm-1. These values are acceptable when compared with the limits of reproducibility established in primary method ASTM D-2699. Ssamples wrongly classified are reduced to less than 10% for high resolution equipment, and 6% in the case of the MLR model.

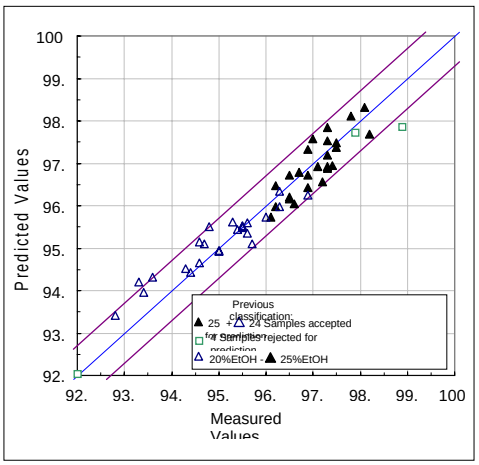


Figure 7 –RON, by PLS model from 3222 spectra, 3 PCs, RMSEP = 0.39.

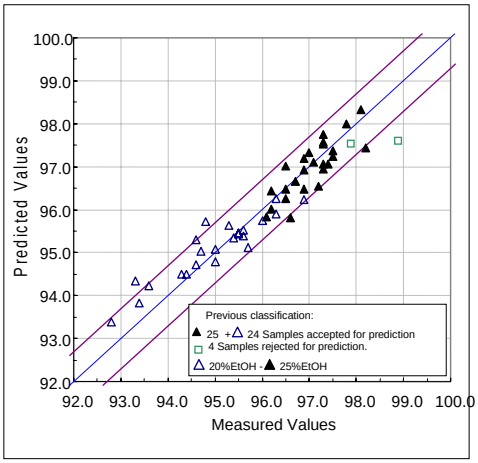


Figure 8 –RON, by PCR model from 3222 spectra, 3 PCs, RMSEP = 0.42.

4. Concluding Remarks

The methods tested for identification of outliers for calibration and for classification of samples of prediction set were similar when the spectral data have the same instrumental origin. Results did not show considerable differences in robustness of linear models for different methods of calibration or for the pre-treatment of spectra by either derivation or numerical smoothing. This suggests that, among the alternatives considered, the best method is the more convenient and more compatible with the computational resources available. Models based on spectra generated by the equipment with 4 cm^{-1} medium resolution were useful only for low accurate prediction of RON and MON. In this case, high values of RMSEP and probability of error exceeds the limits reproducibility of the primary method for more than 30% of samples (prediction set).
Models of octane indexes generated by the equipment with 8.3 cm^{-1} high resolution showed good results and is useful in the analysis of commercial gasoline. These models showed RMSEP of around half the levels of reproducibility of the primary method. At least 90% of samples were correctly classified by the models. However it should be noted that even with the use of high resolution IR equipment, the usefulness of this method is limited to situations in which the unknown samples belong to the same population used in calibration of models, although classification of unknown samples based on their spectra by one of the methods tested in this work is still needed.

References

ANDRADE, J. M.; MUNIATEGUI, S.; PRADA, D., *Fuel*, v.76(11), 1035-1042, 1997.
ASTM Standards D-2699 - Standard Test Method for Knock Charact.of Motor Fuel by Research Method.
ASTM Standards D-2700 - Standard Test Method for Knock Characteristics of Motor Fuel by Motor Method.
BRERETON, R. G. Lab. Sys., v.1, 17-31, 1986
GIANCASPRO, C.; COMISAROW, M. B., *App Spec.*, v.37(2), 153-166, 1983.
KELLY, J. J.; CALLIS, J., *Ana. Chem.*, v.62(14), 1444-1451,1990.
LITANI-BARZILAI, I.; et al., I., *Ana. Chim. Acta*, v.39(1-2), 193-199, 1997.
LÓPEZ-ANREUS, E.; GARRIGUES, S.; GUARDIA, M., *Ana. Chim. Acta*, v333, 157-165, 1996.
WOLD; JOHANSSON.; JELLUM; BJORNSON and NESBAKKEN, Anal. Chim. Acta, v133, 251-259, 1981

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Correlation of PVR, Octane Numbers and Distillation Curve of Gasoline with Data from a Thermal Wave Interferometer

Leonardo S.G.Teixeiraa, Michelle S. G. Dantas[b], Paulo R. B. Guimarães[b], Wagner Teixeira[b], Helion Vargas[c] and José A. P. Lima[c]

aDepartment of Analytical Chemistry – DQA, Chemistry Institute – Federal University of Bahia – UFBA, Campus Universitário de Ondina, CEP 40.170-280, Salvador – BA, Brazil, E-mail: lsgt@ufba.br
bDepartment of Engineering and Architecture – DEAR, Salvador University – UNIFACS, Av. Cardeal da Silva, n.132, CEP 40.231-250, Salvador – BA, Brazil, E-mail: paulorbg@unifacs.br
cCentre of Science and Technology, Norte Fluminense State University – UENF, Av. Alberto Lamengo, n.2000, CEP: 28.015-620, Campos dos Goytacazes – RJ, Brazil E-mail: vargas@uenf.br

Abstract

The ASTM standard methods for experimental determination of gasoline properties, such as octane numbers and Reid vapour pressure (PVR) are, in general, expensive, time consuming and cumbersome. Therefore, the study and development of faster and cheaper methods is of great industrial and scientific interest, particularly when one thinks of large numbers of gasoline samples to be tested. However, gasoline is a mixture of hydrocarbons and presents very complex chemical properties that require the use of several analytical techniques, and so the development of any new method will need to be validated through a series of statistical techniques. This contribution is aimed at the validation of a Thermal Wave Interferometer (IOT), conceived and developed by Vargas et al. at the Norte Fluminense State University – UENF, for the determination of octane numbers (MON and RON), distillation curve and PVR of gasoline and gasohol. The IOT determines these properties through a mathematical correlation of the thermal diffusivity of the components present in the gasoline, or gasohol, sample. This IOT validation study involved the use of a multivariate regression technique, namely SIMCA (Soft Independent Modelling of Class Analogy) for the classification of sample data and identification of outliers and the development of new correlation models for the prediction of gasoline octane numbers. In this study 97 samples of various types of commercial gasoline and gasohol, the latter containing up to 30% v/v of ethanol, were prepared and had their corresponding octane numbers (MON and RON) determined through the CFR motor method and their thermal diffusivities determined through the IOT. These experimental data were then correlated and a PLS (partial least squares) model, based on 89 thermal diffusivity variables, was developed for the prediction of the AKI, a parameter which is the mean value of the octane numbers (MON and RON) of a sample, i.e. AKI = (MON+RON)/2. Better results were obtained when the model was built for distillation curve points (Initial boiling point, 10% evaporated) and PVR data, and the thermal diffusivity data (30-870s) were centred at the mean value and smoothed, giving an RMSEC =1.832 and an RMSEV=2.270. These results indicate that the IOT is a promising, fast and cheap method for the prediction of AKI, provided the

user takes into account the inherent errors of the experimental method. Thus, the IOT may be very useful as a tool for undertaking screening analysis, i.e., selection of gasoline, or gasohol, samples that must be submitted to the more expensive and time consuming CFR motor method.

Keywords: gasoline, gasohol, octane numbers, interferometry

1. Introduction

The quality of gasoline with respect to its resistance to detonation is evaluated through its octane numbers (MON and RON) or Anti Knock Index (AKI), i.e. (MON+RON)/2. Evaluation of octane numbers is justified by the need to guarantee that the fuel meets engine demands at the time of compression and beginning of expansion (when pressure and temperature rise) without auto-ignition. The values of MON and RON are obtained through a special engine, the CFR motor (Cooperative Fuel Research), which is the only test legally recognized by the Brazilian Oil Regulatory Agency (ANP) for that purpose.

Other important tests for determining fuel quality are distillation and Reid Vapour Pressure (PVR). The former gives a measure, in terms of volatility, of HC distribution in a gasoline sample. In Brazil, ANP establishes specific temperatures for evaporation of T10%, T50%, T90% of the fuel, as well as its Initial Boiling Point (IBP) and Final Boiling Point (FBP) and the amount of residue left after distillation. The latter evaluates the tendency of gasoline to evaporate, and, therefore, the higher the PVR the easier it is for a fuel to evaporate.

Although well established and reliable, several of the standard methods used for fuel quality evaluation are tedious, cumbersome and involve the use expensive equipment and highly qualified personnel. Therefore, there is a need for developing techniques and correlations between other physical-chemical properties and gasoline quality parameters[A – B]. This research work is dedicated to the study and evaluation of use of a Thermal Wave Interferometer (IOT)[C – F], which takes photometric measures of diffusivity and relates them to gasoline quality parameters. According to Lima et al[E – F], the thermal diffusivity, defined as the ratio of the thermal conductivity to the sample heat capacity per unit volume, measures essentially the samples thermalisation time. The photothermal techniques are essentially based upon sensing the fluctuation in temperature of a given sample due to non-radiative de-excitation processes following the absorption of modulated light[G]. The importance of thermal diffusivity as a physical parameter to be monitored is due to the fact it is unique for each material. This can be appreciated from their tabulated values published in literature[H] for a wide range of materials.

Application of some multivariate calibration techniques to the modelling of gasoline properties have been reported in the literature[I – K]. Aiming at evaluating the potential of thermal diffusivity for predicting the antiknock properties of gasoline a commercial software has been used (UNSCRAMBLER®, v. 8.0) for chemometric studies. This allowed the development of linear multivariate PLS regression models through the SIMCA (Soft Independent Modelling of Class Analogy) method. The results obtained for pure gasoline (Gasoline A) and gasoline containing 20% to 30%(v/v) of ethanol (Gasoline C) have been used for the multivariate calibration that was later applied in the modelling of gasoline properties, in particular octane numbers.

2. Methodology

Correlations between experimental diffusivity data obtained through the IOT and gasoline properties obtained through the standard methods for PVR, MON, RON, Distillation (IBP, T10%, T50%, T90% and FBP) have been studied. The distillation tests have been undertaken using ISL® and HERZOG® automatic distillation units, in accordance with ASTM D86. For PVR, a Grabner Instruments Minivap-VPS, in accordance with ASTM D5191 has been used. Finally, for octane numbers (MON and RON), a CFR engine has been used, in accordance with ASTM D2699 and D2700.
The reproducibility limits of the standard methods are shown in Table 1, in accordance with their corresponding ASTM methods. These values have been considered in the analysis of the results obtained from the correlation models developed in this research study.

Table 1. Reproducibility Limits of the standard methods gasoline characterisation

Property	**Standard Method**	**Reproducibility Limits**
Specific gravity (g/cm^3)	ASTM D4052	0.0005
IBP (ºC)	ASTM D86	7
T 10% (ºC)	ASTM D86	7
T 50% (ºC)	ASTM D86	8.5
T 90% (ºC)	ASTM D86	9.5
FBP (ºC)	ASTM D86	10.5
MON	ASTM D2700	0.9
RON	ASTM D2699	0.7
Paraffins (%)	chromatography	3.90
Aromatics (%)	chromatography	2.35
Naphthenes (%)	chromatography	1.17
Olefins (%)	chromatography	1.94

These same values have been used for evaluation of the performance of the correlation models to both the calibration data set (RMSEC), Equation 1, and the prediction data set (RMSEP), Equation 2. The latter was also used to evaluate the classifying method used for detection of outliers.

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n} (\overline{y}_{i,cal} - y_{i,cal})^2}{n}} \quad (1)$$

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n} (\overline{y}_{i,prediction} - y_{i,prediction})^2}{n}} \quad (2)$$

Where,

- $\boldsymbol{n}$ is the number of samples accepted by the classifying method;
- $\overline{y}_{i,cal}$ is the predicted property;
- $y_{i,cal}$ is the experimental value obtained from a standard method.

The effectiveness of the methods for the prediction of outliers have been compared through the deviations of the refined calibration data set, the classified prediction data

set and the number of correctly identified outliers in the latter by each classifying method.

3. Results and Discussion

The model developed for prediction of AKI involved 89 thermal diffusivity variables obtained through the IOT and the use of the optimal principal components (PC) as determined by the minimum of residual variance of the dependent variable. Projections in orthogonal directions of the thermal diffusivity variables were obtained so that the correlation between the set of projections for these samples and the dependent variable resulted in linear coefficients for each thermal diffusivity variable, allowing the calculation of the partial components in orthogonal axes, which in turn made it possible to predict the value of the dependent variables.

The AKI prediction through IOT's internal model has shown a mean squared error of 3.267 for samples of the training data set and 3.295 for the validation set, while the acceptable reproducibility for this parameter is 0.8. This great discrepancy in the prediction of AKI may be observed in Figure 1, in which the mean squared error exceeds the reproducibility limits, indicating that there is room for much improvement in the model used by the IOT.

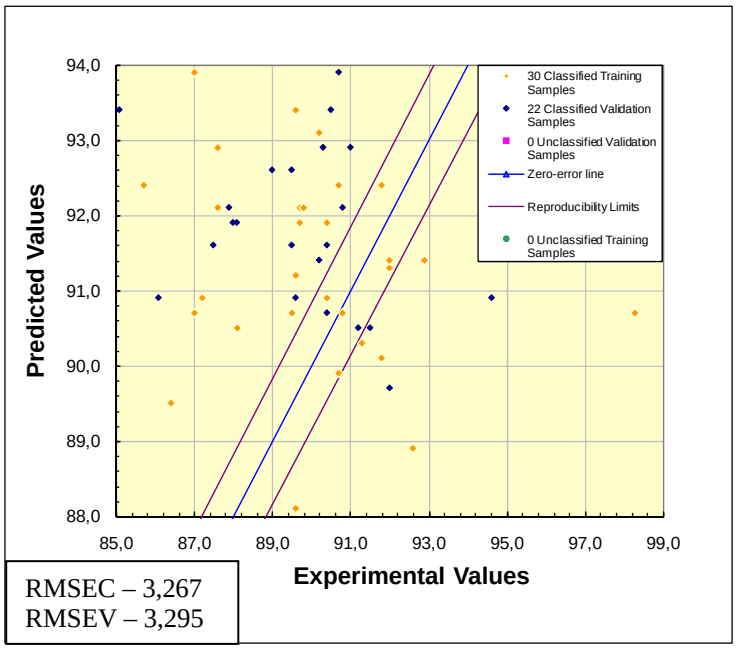


Figure 1 – AKI prediction through IOT's internal model

Figure 2 shows the results obtained from thermal diffusivity and PVR data for 30 samples. The resulting correlation shows to be unsatisfactory, since the error limits have been shifted so that the samples from the validation set would lie within the allowed area.

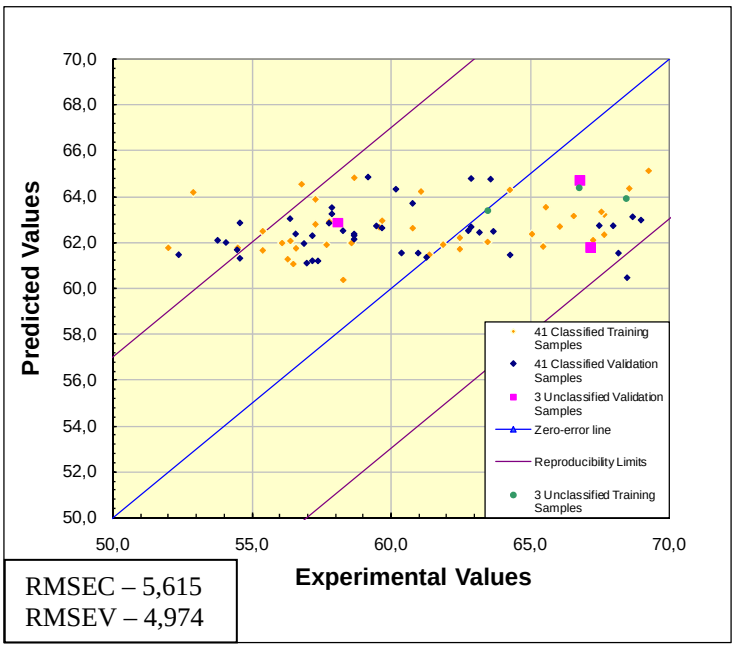


Figure 2 – PLS model for PVR data for Gasoline A and C samples

Better results, however, have been obtained when some distillation points were used. Figure 3 shows data of T10% for 44 samples, and it can be observed that these are evenly distributed along the diagonal and prediction errors are within the ASTM reproducibility limits. Therefore, the correlation thus obtained is in accordance with those obtained using thermal diffusivity.

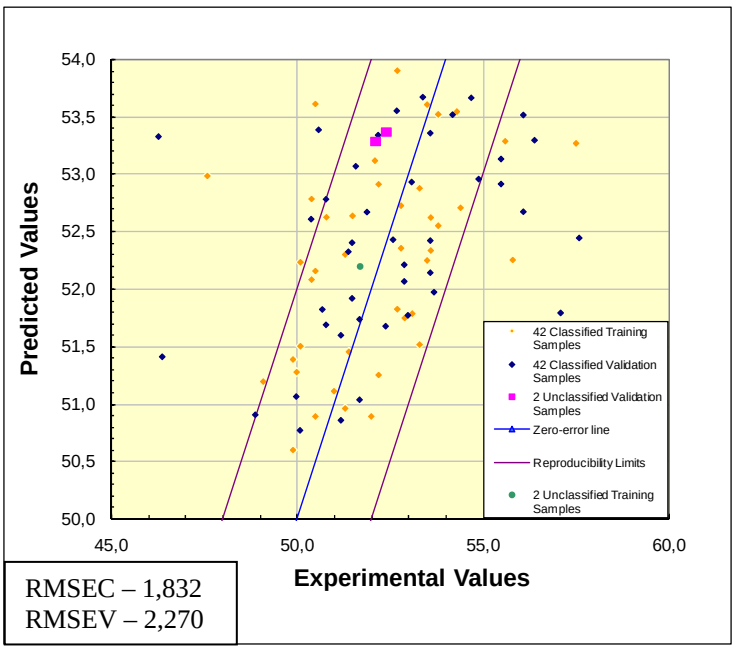


Figure 3 –PLS model for T10% data for Gasoline A and C samples

4. Concluding Remarks

In this research study PLS linear multivariate regression has been used for the development of models correlating thermal diffusivity data and gasoline properties obtained through their corresponding standard methods. Among the properties analysed, the best correlation was obtained for T10%. With respect to octane numbers, IOT has shown to be a valuable asset for screening samples that should be submitted to the standard CFR engine test.

References

A. L. S.G. Teixeira, F. S. Oliveira, H. C. dos Santos, P. W. L. Cordeiro, S. Q. Almeida, 2008, Multivariate calibration in Fourier transform infrared spectrometryas a tool to detect adulterations in Brazilian gasoline, Fuel, 87, 346-352;

B. F. S. de Oliveiraa, L. S. G. Teixeira, M. C. U. Araujo, M. Korn, 2004, Screening analysis to detect adulterations in Brazilian gasoline samples using distillation curves, Fuel, 83, 917-923;

C. I. A. Esquef, A. P. L. Siqueira, M. G. Silva, H. Vargas, 2006, Photothermal characterization of nature gas automotive fuel, Meas. Sci. Technol., 17, 1385-1389;

D. I. A. Esquef, A. P. L. Siqueira, M. G. Silva, H. Vargas, 2006, Photothermal gas analyzer for simultaneous measurements of thermal diffusivity and thermal effusivity, Anal. Chem.,78, 5218-5221;

E. J.A.P. Lima, M. G. Silva, S. L. Cardoso, M. S. O. Massunaga, H. Vargas, 2003, Monitoring of gas diffusion in air using the TWI technique: Thermal diffusivity measurements made easy, Rev. Sci. Instrum., 74, 1, 842-844;

F. J. A. P. Lima, S. L. Cardoso, M. G. Silva, H. Vargas, 2001, Thermal diffusivity as an automotive fuel characterization parameter: Correlation with motor octane number, Ind. Eng. Chem. Res., 40, 6207-6212;

G. J. A. P. Lima, E. Marín, O. Correa, M. G. da Silva, S. L. Cardoso,.C. Gatts, C. E. Rezende, H. Vargas, L. C. M. Miranda, 2000, Measurement of the thermal properties of liquids using a thermal wave interferometer, Meas. Sci. Technol., 11, 1522-1526.

H. Y. S. Touloukian, R. W. Powell,1973, Thermal Diffusivity;

I. J. J. Kelly, J. Callis, 1990, Nondestructive analytical procedure for simultaneous estimation of the major classes of hydrocarbon contituents of finished gasoline, Anal. Chem. 62, 14, 1444-1451;

J. F. Parisi, L. Nogueira and H. Prieto, 1990, On line determination of fuel quality parameters using near-infrared spectrometry with fiber optics and multivariate calibration. Anal. Chim. Acta, 238, 945-100;

K. I. Litani-Brazilai, I. Sela, V. Bulatov, I. Zilberman, 1997, On line remote prediction of gasoline properties by combined optical methods, Anal. Chim. Acta, 39, 193-199;

L. E. López-Anreus, S. Garrigues, M. Guardia. 1996, Vapor generation-fourier transform infrared spectrometric determination of benzene, toluene and methyl tert-butil ether in gasoline, Anal. Chim. Acta, 333, 157-165;

M. J. H. Alivas, 1997, Data sets of near infrared spectra. Chemometr Intell Lab Syst, 37, 255-259.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Reactive Scheduling Approach Based on Domain-Knowledge

Juan M. Novas, Gabriela P. Henning
INTEC (UNL-CONICET), Güemes 3450, Santa Fe, CP 3000, Argentina

Abstract

Real world industrial environments frequently face unexpected events that generally disrupt in-progress production schedules. This contribution presents advances in the development of a support framework to address the repair-based reactive scheduling of industrial batch plants. When facing an unforeseen event, the framework is capable of capturing the current operational plan and its status. Based on this information, a rescheduling problem specification is developed. Tasks to be rescheduled are identified and, for them, the set of the most suitable rescheduling action types (e.g. shift, reassign, etc.) is specified. For a given specification, many solutions to the problem could exist. Then, the second step of this approach relies on the generation of a constraint programming (CP) model to address the rescheduling problem just specified. To create such model each rescheduling action type is automatically transformed into different constraints. In addition, a search strategy based on domain knowledge can also be developed. Finally, the solution of the CP model and its associated search strategy will render the repaired schedule in which the repair action types that were suggested will be instantiated. A case study of a multiproduct multistage batch plant is presented, where an event of unit failure is considered.

Keywords: Reactive Scheduling, Decision Support Systems, Batch Plants.

1. Introduction

Usually, predictive scheduling techniques generate production plans which assume stationary operating conditions along the whole scheduling horizon. However, real industrial environments are dynamic in nature; unforeseen events frequently disrupt the in-progress schedule. Reactive scheduling, or rescheduling, is then performed in order to update the current agenda. This task can be periodically executed or when an unexpected event, such as a unit breakdown, order cancellation or arrival, occurs, leading to periodic versus event-driven rescheduling, respectively. To be adopted, reactive scheduling systems must provide immediate responses to disruptions. Besides, minimum changes to the original schedule are desired to maintain a smooth plant operation. Hence, repair-based or partial rescheduling is preferred to a full-scale one.

Several works about reactive scheduling of batch plants have been reported. Janak et al. (2006) presented a review of approaches tackling this kind of problem and other papers in the field have recently been published (Ferrer-Nadal et al., 2007). Even though the review of Vieira et al. (2003) focused on discrete manufacturing plants, most of the issues described in it can be found in process industries too.

This contribution tackles the reactive scheduling problem by developing a support environment based on both, an explicit representation of the domain knowledge and a CP approach. The next section presents an overview which addresses these two components. Finally, Section 3 discusses an illustrative example.

2. Reactive Support Environment

When addressing a rescheduling situation most of the objectives and basic constraints that define the original problem still apply; however, the partially executed schedule and the perturbation or triggering event has also to be taken into account. This contribution presents a support framework able to represent this knowledge and use it in the development of a solution. It has been envisioned to operate under an event-driven policy. The framework explicitly captures the status of an in-progress schedule and characterizes the disrupting event in order to specify the rescheduling problem to be faced and, afterwards, solve it. The environment integrates different modules, as it is shown in Fig.1, all based on an explicit domain representation.

As the first step of the proposed approach, tasks to be rescheduled are identified, and for each of them, the most suitable rescheduling action type is specified. An action type not only prescribes what to do (e.g. shift, reassign, etc.) on a processing task, but also a range of possibilities, i.e. feasible equipment for a reassignment, shifting interval, etc. For a given specification, corresponding to the set of action types associated to the tasks to be rescheduled, many solutions could exist. Then, the second step of this approach relies on the generation of a constraint programming (CP) model to address the rescheduling problem just specified. To create such model each rescheduling action type is automatically transformed into different constraints. A search strategy based on domain knowledge can be employed, too. Finally, the solution of the CP model and its associated search strategy will render the repaired schedule, in which the proposed repair action types will be instantiated.

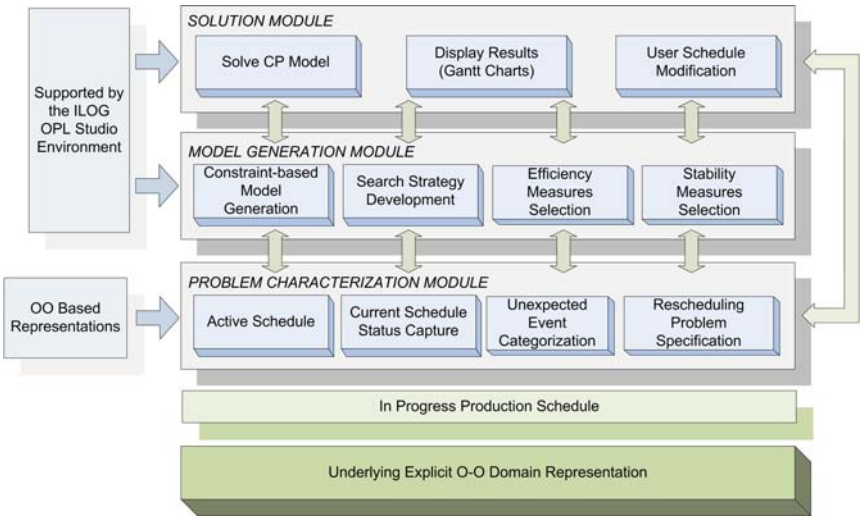


Fig.1. Components of the Reactive Scheduling Support Framework

2.1. Domain Knowledge Representation and Problem Characterization

Scheduling is a knowledge intensive activity in terms of domain knowledge. The representation of this type of knowledge is a critical issue in any scheduling support system. Fig. 2 presents a simplified version of the conceptual model of multiproduct batch plants scheduling knowledge included in the proposed framework. This information is organized into different conceptual levels. Generic concepts represent engineering or plant information that is used by entities which model a given scheduling scenario. Because of the fact that scheduling is an evolving activity, different versions

of a schedule need to be captured which are related among themselves. A schedule version is a snapshot of an active schedule, which is valid at a given moment, and is represented by means of scheduling domain entity versions (i.e., versions of the instances of *Task, ResourceSchedule, Schedule,* etc.). The handling of versions is not discussed here due to space limitations.

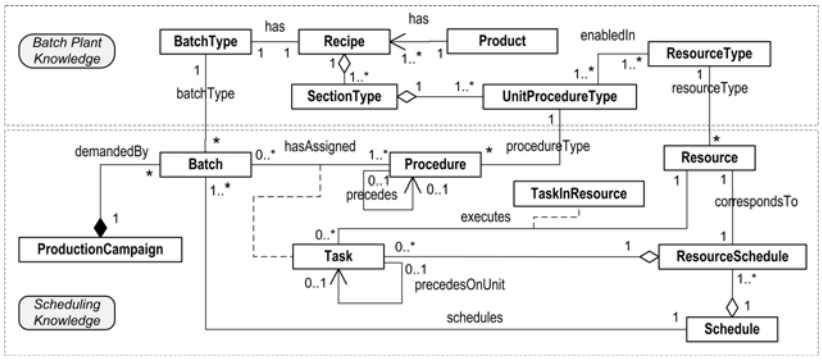


Fig.2. Partial view of batch scheduling domain concepts

Since the domain knowledge is explicitly captured, the problem characterization module fulfills the following functional requirements.

- *Capture the current operational plan and its status at the rescheduling point*. Instances of entities represented at the scheduling knowledge level, their relationships, as well as the values of their attributes, allow modeling each problem to be tackled. When addressing a rescheduling situation, the status of each task involved in the disrupted schedule must be known. At a rescheduling point *t*, a decision-making activity about tasks to be included in the reactive process has to be made. Hence, all the tasks involved in the disrupted agenda, T , are classified into the following sets, depending on their status at time point *t*: (i) already executed tasks, T_t^{AE}, (ii) non-executed tasks, T_t^{NE}, and (iii) in-progress tasks, T_t^{IP}. In addition, if the disrupting event is a batch cancellation, its associated tasks are included in the set T_t^{CE}. In turn, new tasks to be considered due to a rush batch arrival, belong to the set T_t^{NW}. These sets are related in the following way: $T = \{T_t^{AE} \cup T_t^{IP} \cup T_t^{NE}\}$; $T_t^{CE} \subseteq T_t^{NE}$; $T_t^{NW} \not\subset T$. Then, the status of a batch can be inferred from the condition of its associated tasks. Moreover, equipment status is also domain information which is taken into account.
- *Identification of the disruptive event type and its specific characteristics.* The characterization and modeling of events is another critical aspect of the problem too. The framework is able to handle equipment breakdowns, new batch arrivals and batch cancellations as disrupting events. Other types of events will be considered in the future. Based on task classification and event information, activities to be drawn in the rescheduling process are identified. Thus, RT_t is the set of tasks that needs to be considered due to the unexpected event taking place at time *t*. The remaining tasks are those not directly involved in the reactive process, referred as NT_t.

The problem characterization module uses information from the object instances pertaining to the sets previously described and from the entity representing the *UnexpectedEvent* in order to determine: (i) the updated equipment ready times, (ii) the earliest start times of those tasks that pertain to in-progress batches. This information is also employed to properly specify the reactive scheduling problem, as discussed below.

2.1.1. Rescheduling Problem Specification

For each task belonging to RT_t, its associated rescheduling action type (RAT) has to be decided. By means of the proposed methodology, each task does not have a specific and pre-defined rescheduling operation related to it; instead, it has a type of rescheduling action. Each action type prescribes a category of rescheduling operation and a range of possibilities to apply it. In this framework, the following RATs have been considered up to now: *a) Shift*, which simply "pushes" the task start time forward (right-shift) or backwards (left-shift) on the current unit. To apply such action on a task, or set of tasks, allowed left shift (*lsv*) and right shift values (*rsv*) are defined to specify a time window for the new start time. *b) Reassign*, which allows a task to be reallocated to a unit belonging to a set of feasible ones. *c) Freeze*, which is not a scheduling action itself, but is required for tasks that must keep their timings as well as their equipment assignments. In order to define the RAT to be associated with each task, they are previously classified into: (i) directly affected ones, comprising set RT_t^{DA}, (ii) indirectly affected tasks, represented by set RT_t^{IA}, and (iii) non-affected tasks, included in set RT_t^{NA}. The way tasks are categorized depends on the event type and to the extend tasks are disrupted by it. Directly and indirectly affected tasks are linked to specific rescheduling action types, which depend on the event type. On the contrary, action types associated with non-affected tasks are defined employing domain knowledge of each specific problem. To illustrate these ideas, sets RT_t^{DA}, RT_t^{IA} and RT_t^{NA} are defined for a unit failure event and their associated RATs are discussed. Similar definitions can be made for a batch cancellation (not shown due to lack of space) or other disruptions.

Task Classification for a Unit Failure Event

RT_t^{DA}. Tasks in T_t^{NE}, which are assigned to the broken-down unit, and start later than *t* but earlier than the unit recovery time. Besides, if a task is interrupted by the unit failure and cannot be recovered, its batch is canceled and all its associated tasks are inserted as a new batch (*NewBatchArrival* event) to be included in the rescheduling process.

RT_t^{IA}. These indirectly affected tasks are linked with, at least, one of the directly affected activities. They are the ones located downstream in those batches having at least a directly disrupted task. Activities located upstream can be considered as indirectly affected too, as will be shown in the example presented in the next section.

RT_t^{NA}. Includes tasks which are not affected, neither directly nor indirectly, by the unit breakdown.

RATs for Tasks Associated with a Unit Breakdown

Tasks in RT_t^{DA} are associated with a reassign action type that forces them to find another processing unit. Similarly, members of RT_t^{IA} are associated with a reassign action type that allows them to retain their current unit allocations (with a possible change in timings) or to choose a different one. Finally, different criteria were established to define the RATs to be associated with tasks included in set RT_t^{NA}. Two criteria were proposed in this work: the scheduling horizon-based (*SHB*) and a batch information-based (*BIB*) criteria.

SHB criterion: It takes into account the moment the disruptive event takes place in relation to the length of the whole scheduling horizon. Depending on the scheduling scenario and the costs of changing the current task allocations, it can be decided to: (i) freeze tasks belonging to set RT_t^{NA}, whose current start times are very close to the rescheduling point, (ii) enable shift movements over those tasks from set RT_t^{NA} having a start time not as close to the rescheduling point, but not located at the end of the

horizon, (iii) enable reassign action types over those tasks placed at the end of the scheduling horizon. A parameter V is employed to identify these three intervals.
BIB criterion: Information about the status of the current schedule is employed to define the RATs over those tasks in set RT_t^{NA}. For instance, batches having an important progress at the rescheduling point are not allowed to have their non-executed tasks reassigned, (ii) tasks associated with batches having a small positive or negative values of their slack times, are allowed to be reallocated to avoid violating their due dates.
In this way, a rescheduling problem specification gives rise to a set of rescheduling alternatives for every task, and many solutions to the problem could exist. Then, the second step of this approach relies on the generation of a constraint programming (CP) model to address the rescheduling problem just specified. To create such model each RAT is automatically transformed into different constraints.

2.2. Model Generation
In order to create a repaired schedule, the model generation module is in charge of setting up a constraint programming (CP) model, complying with the syntax of the OPL language from ILOG OPL Studio (ILOG, 2002). This model (not shown due to lack of space) comprises: (i) the basic constraints set (assignment, precedence, timing constraints, etc.), which are part of any scheduling problem. They are generated for all tasks in RT_t and (ii) event-dependent constraint set. These last ones depend on the disrupting situation been faced, and on the adopted RATs for tasks included in sets RT_t^{DA}, RT_t^{IA} and RT_t^{NA}. The objective function can be chosen among several efficiency or stability measures. Despite just one of them is optimized, the values reached by the rest are taken into account to comprehensively evaluate the scenario.

3. Example

This proposal was tested with various case studies. One of them, introduced by Pinto and Grossmann (1997), is presented in this section. The plant involves 4 processing stages and 10 units. During the scheduling horizon, 20 orders have to be processed. Fig. 3.a shows an in-progress schedule, with tasks labeled by batch name. Let us assume that unit *u4* shuts down at t = 31.00 h and it will be unavailable till a *recovery time* = 134 h.
The problem was solved employing three different approaches in order to compare results: by resorting to a full-scale rescheduling (*FSR*) and to two partial ones. In the first case all tasks involved in the rescheduling process were allowed to be reassigned. The second approach was a partial reschedule, identified as *PR1*, which employed the *SHB* criterion to handle tasks in RT_t^{NA}. Tasks in sets RT_t^{DA} and RT_t^{IA} were associated with a reassign rescheduling action type. For tasks in set RT_t^{NA}, V = 24 h was adopted (justification not discussed due to lack of space), meaning that tasks having their planned start time within the rescheduling point, and 24 h after it, were associated with freeze action type. Tasks in RT_t^{NA}, starting at least 24 h after the rescheduling point, were associated with a shift action type. The new schedule, obtained under these conditions, which is not shown due to lack of space, was not a good quality one due to the topological constraints that had to be considered. Because of them, batches assigned to units *u1* and *u2* in the first stage, which were in set RT_t^{NA}, could not be assigned to unit *u5* in the second stage and had to wait for the recovery of *u4*. On the contrary, the other partial reschedule approach (*PR2*) employed topological knowledge. The criterion used was also *SHB*, but now upstream tasks were considered in RT_t^{IA}, too. So, in order to give more flexibility to tasks assigned to units *u1* and *u2*, they were associated with

reassign action types. The solution that was obtained is shown in Fig. 3. Schedules were assessed by means of a Makespan performance measure, plus equipment and temporal stability functions that allow quantifying a smooth plant operation, among others. The solution obtained with the *PR2* approach (Mk = 259.9) is of better quality than those obtained with the *PR1* (Mk = 301.9) and *FSR* (Mk = 262.8) strategies.

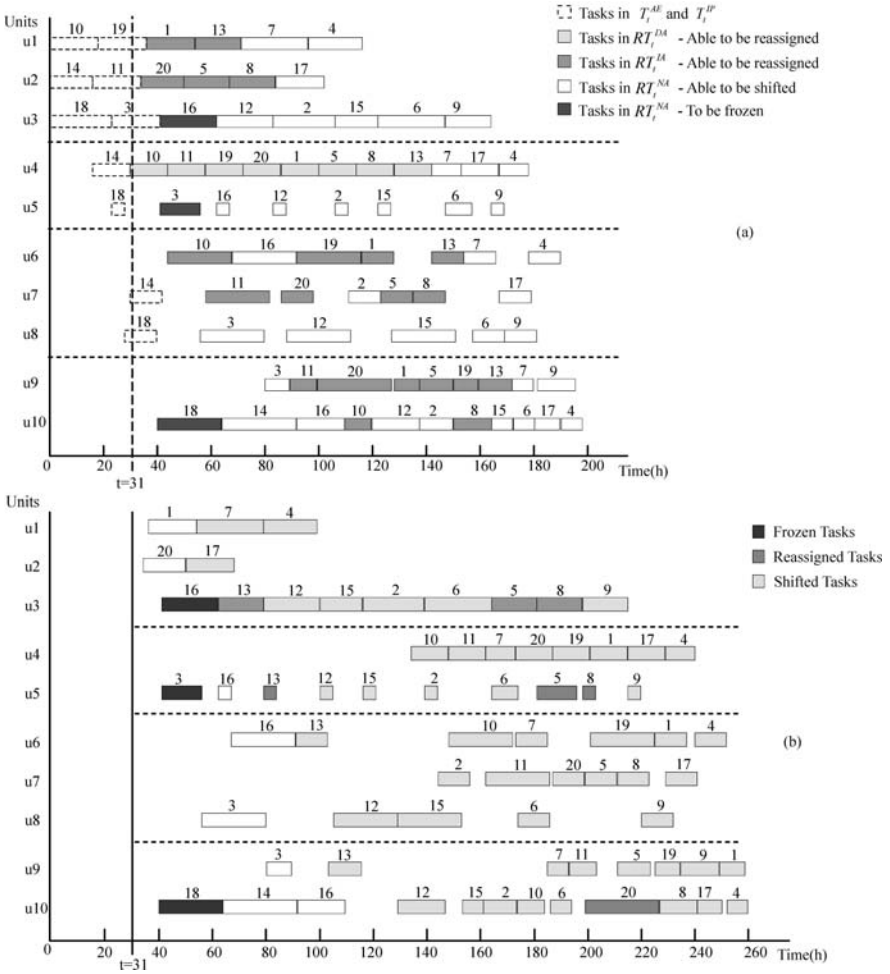


Fig.3. a) In-progress schedule (UIS policy). Classification of tasks according to *PR2* is depicted with grey shades. b) New schedule obtained after applying the *PR2* approach.

References

S. Ferrer-Nadal, C. A. Méndez, M. Graells, L. Puigjaner, 2007, Optimal Reactive Scheduling of Manufacturing Plants with Flexible Bath Recipes, Ind. Eng. Chem. Res., 46, 6273-6283.

ILOG: ILOG OPL Studio 3.5. User's Manual, France, 2002.

S.A. Janak, C.A. Floudas, J. Kallrath, N. Vormbrock, 2006, Production Scheduling of a Large-Scale Industrial Batch Plant. II. Reactive Scheduling. Ind. Eng. Chem. Res., 45, 8253-8269.

J. M. Pinto, I. E. Grossmann, 1997, A logic-based approach to scheduling problems with resource constraints, Comp. Chem. Eng., 21, 801-818.

G. Vieira, J. Herrmann, E. Lin, 2003, Rescheduling Manufacturing Systems: A Framework of Strategies, Policies, and Methods, J. Sched., 6, 39-62.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Separating asphaltenes from lube oil through supercritical deasphalting considering experimental and virtual plants and thermodynamic analysis

Maria R. W. Maciel[a], Rubens Maciel Filho[a], Viktor O. C. Cárdenas[a], Erika T. Koroishi[a], Florencia W. R. Rivarola[a], Filipe A. B. Quirino[a], Cesar B. Batistella[a], Lilian C. Medina[b]

[a]*Laboratory of Optimization, Design and Advanced Control, School of Chemical Engineering, University of Campinas P.O. Box: 6066, Campinas, São Paulo, Zip Code13083-970, Brazil.*E-mail address: wolf@feq.unicamp.br

[b]*CENPES/PETROBRAS, Cidade Universitária, Q.7, Ilha do Governador, Rio de Janeiro,Zip Code 21949-900, Brazil.*

Abstract

Studies of petroleum asphaltenes have gained considerable attention in the past decades due to the increasing production of heavy crude oils. The reduction of light crude oil reservoirs and the increasing of light oil demand, forced the petroleum industry to develop upgrading processes for raw materials and residues. Due to the limited global oil reserves, more and more heavy crudes are being processed. These crudes contain large amounts of asphaltenes and resins. Petroleum must be processed in order to have the major quantity of higher aggregated compounds. All types of petroleum are colloidal systems. In a dispersion medium consisting mainly of hydrocarbons, which can be classified into alkanes, naphthenes and aromatics, there are two groups of dispersed colloidal particles in solution: asphaltenes and petroleum resins. Reduction of asphaltenes and metal content can be achieved by disturbing the solvatation equilibrium via addition of suitable solvents, e.g., propane. The Residuum Oil Supercritical Extraction (ROSE™) process is the premier deasphalting technology available in industry today. This state-of-the-art process extracts high-quality deasphalted oil (DAO) and asphaltenes from atmospheric or vacuum residues and other heavier feedstocks. The ROSE process operates near to the critical point of the solvent and applies thermodynamic fundamentals and high-pressure phase equilibrium principles in order to provide an energy-efficient process. Depending on the solvent selection, the DAO can be an excellent feedstock for catalytic cracking, hydrocracking, or lube oil blending. The energy consumption in supercritical extraction is considerably lower than in conventional extraction, which requires a higher ratio of solvent to crude. The use of supercritical fluid has some advantages such as higher yield and improved quality of the valuable DAO and asphaltenes product, and recovery of the supercritical solvent, reducing significantly operating costs compared to other solvent deasphalting processes. This work presents a new method for petroleum deasphalting. The proposal involves the extraction of the residue fractions from molecular distillations using supercritical propane or pentane as solvent. The thermodynamic calculation and representation of the ternary system is shown. The calculation is very complex since the asphaltenes have to be represented as a set of molecules whose properties have to be predicted using special

methodology and the UNIFAC method. Our research group at State University of Campinas (UNICAMP) has developed a real supercritical extraction pilot plant and a virtual one which are optimized and validated. This work is carried out in association with Petrobras (Brazilian Oil Company).

Keywords: deasphalting process, ternary system, pilot plant

1. Introduction

In developing the technology of supercritical extraction, solvent selection is a basic problem. The efficiency of the process depends mainly on the solvent selection. The development of the deasphaltation technology using supercritical fluid (SCF) appears as a solution to improve the separation of DAO from asphaltenes. The use of supercritical fluid has some advantages like: the difference of densities between the extraction phase and the rafinate phase is larger than that obtained by the conventional liquid-liquid extraction, becoming the separation between the phases easier; the mass transfer is faster using the supercritical fluid; the quantity and the quality of the DAO can be easily controlled adjusting the temperature and the pressure of the extraction system and the efficiency to recover the oil is a function of the density of the supercritical fluid (Rodrigues, 2006 and Mendes *et al.*, 2005).

The knowledge of a proper thermodynamic modeling for mixtures of hydrocarbons in high pressures is extremely important for diverse applications in the petroleum engineering, such as reservoir simulation at high pressures and evaluation of extraction techniques of oil in deep waters.

In this work, it was studied the behavior of ternary diagrams for two systems and a supercritical deasphalting pilot plant was developed in order to reach the operational conditions desired.

2. Thermodynamic Modelling of Supercritical Phase Behavior

As it is well known, pressure and temperature conditions, composition of the mixture and molecular classes which the components belong to affect the phase equilibrium. However, different unities of a supercritical process work in a wide range of pressure, temperature and composition, so the modelling tools should be capable to predict and correlate situations where exist equilibrium such as solid-SCF, liquid-SCF and liquid-liquid-SCF.

Fortunately, in many applications, high-pressure phase behavior can be reasonably represented with a single equation of state (EOS) if the components in the mixture do not differ substantially in size, structure or shape. The most commonly used cubic EOS are the Soave-Redlich-Kwong (SRK) and Peng-Robinson (PR) equations.

In this work, the model used is the Predictive Soave-Redlich-Kwong equation of state (PSRK), which is an extension of the SRK equation of state. This model uses the Holderbaum-Gemehling mixing rules; these rules can predict the binary interactions at any pressure. Using UNIFAC, the PSRK method is predictive for any interaction that can be predicted by UNIFAC at low pressure. The main advantage of using PSRK equation of state is that it is more accurate in prediction of the binary interaction parameters and it gives more satisfactory results for mixtures of non-polar and polar components, as the case of the asphaltenes-oil mixture.

The Holderbaum-Gemehling mixing rules use a relationship between the excess Helmholtz energy and equation-of-state. The goal of these mixing rules is to be able to use binary interaction parameters for activity coefficient models at any pressure. They

do not use a relationship between equation-of-state properties and excess Gibbs energy, as in the Huron-Vidal mixing rules.

3. Development of supercritical deasphalting pilot plant

The supercritical deasphalting pilot plant has been developed at the LOPCA/LDPS Laboratories at State University of Campinas (UNICAMP). The system is equipped with an extractor, a separator, a propane reservoir, valves, pumps, a refrigerating system, temperature controllers and instrumentation for measuring pressure. The extractor has an internal removable column with a 3000ml capacity, 700bar can be reached in this vessel. The extraction unit is equipped with a heating system, temperature controllers and pressure measurement devices. The solvent is charged through the bottom of the extractor, using a booster that is also used to reach the desired work pressure. The solvent feeding line is set with a retention valve, in order to prevent reflux from the oil previously charged in the extractor. The unit has independent pumps for each solvent that will be used in the experiments. The boosters were designed for the equipment in order to attend the process. The separator consists in a 360 stainless steel column and 200bar can be reached. This unit is equipped with independent heating devices for top and bottom, with the purpose of creating a temperature gradient and consequently, enhancing the separation yield. Finally, the propane reservoir consists in a pressure column and a closed cylinder in which circulates a cooling bath, to maintain the propane in the liquid state. This reservoir is used to store the propane that will be used in the experiment, and also to receive the propane recycled from the previous run. The pressure column is made from a pressure resistant glass (up to 20bar can be reached) in order to allow a visual evaluation on the level of the propane.
Modifications were carried out in a conventional equipment design. The lines and valves were modified due to the fact that the equipment presents pressure drop. Besides, the pumps were carefully designed because they easily can present cavitation. After the modifications, the pilot plant has been shown to be able to attend the operating conditions of the process. These modifications/details have largely contributed to the equipment performance.

4. Results and Discussion

4.1 Ternary phase behavior

Simulations were carried out in *Aspen Plus* simulator in order to generate the ternary diagrams. The ternary phase behaviours were analyzed to the asphaltene-oil-pentane mixture and asphaltene-oil-propane mixture at three different conditions. As it is known, crude oil asphaltenes represent a solubility class, not a specific molecular species or family. For this, there is not a specific molecular structure for asphaltenes, since different chemical structures exist for asphaltenes, depending on the origin of the crude oil; different molar mass among several types of asphaltenes are also reported.
The characterization of the asphaltene component in the process simulator was done based on the molecular structure. For the system asphaltene-oil-pentane, it was used the molecular structure presented in Figure 1. For generating the ternary diagrams of the mixture asphaltenes-oil-propane, it was considered the asphaltene molecule of a heavy Canadian residue. The molecular structure of the heavy Canadian asphaltene is shown in Figure 2. The molecules were characterized with UNIFAC group contribution method. The thermodynamic equilibrium model used in the simulation was PSRK with Holderbaum-Gemehling mixing rules. The binary interaction parameters k_{ij} were well predicted using UNIFAC for all pairs of components.

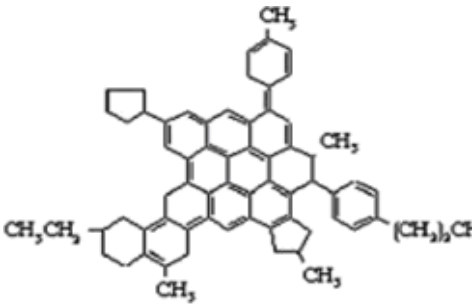


Figure 1. Molecular structure of asphaltene

(Adapted of Loh *et al*., 2007).

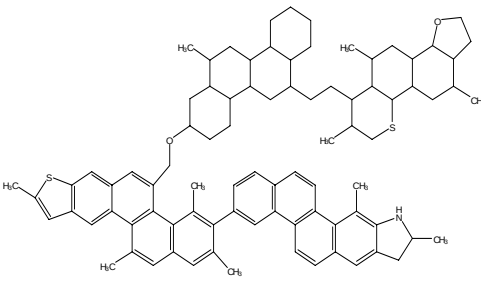


Figure 2. Molecular structure of an asphaltene from a heavy Canadian vacuum residue (Siskin *et al.*, 2006).

The ternary diagrams generated with *Aspen Plus* for the systems asphaltenes-oil-pentane and asphaltenes-oil-propane are shown in Figures 3 and 4, respectively. The principles governing the phase behavior shown in these diagrams are similar to those described qualitatively by Wilson et al. (1936).

For the system asphalt-oil-pentane, at 10 bar (subcritical conditions), pentane is almost entirely miscible with the oil, the oil is entirely miscible with the asphalt too, but the pentane is only partially miscible with the asphalt, therefore the asphalt falls out of the solution. This represents the first step of the lube oil refining process. Figure 3 (B) shows that for a pressure slightly below the critical pressure of pentane (33 bar) the two-phase region of the diagram expands considerably, so liquid pentane is in a much more expanded state than at 10 bar. According to Figure 3 (C), it can observed that the degree of separation appears to have improved at this condition for asphalt-oil feed mixtures. As the solubility power of pentane decreases it is possible to separate resins and lighter components of the oil phase. This represents the last step of separation in the oil deasphalting process.

For the system asphalt-oil-propane, Figure 4 (A) shows (15 bar) subcritical conditions. The oil is totally miscible with asphalt and the propane is partially miscible with asphalt, as it is shown by the region of immiscibility through the curve in the propane-asphalt axis, however propane presents complete miscibility with the paraffinic phase (oil), therefore the asphaltenes are dropped out of the solution. This represents the first stage of the refining process of lube oil. Figure 4 (B) shows that in a pressure slightly below to the critical pressure of the solvent (41 bar), there is an expansion in the two-phase region of the diagram, thus the liquid propane is in a more expanded state than at 15 bar. If the pressure is increased to above the critical pressure (Figure 4 (C)), the system will present a immiscibility between oil and propane, being possible to separate the resins and the lightest fractions of the residue from the propane-oil mixture, due to the reduction of the density and to the solvent power of propane in these conditions, allowing a separation of the extract stream free of asphaltenes in DAO and propane, which after some refining operations can be recycled to the process. This condition corresponds to the last stage of the ROSE process.

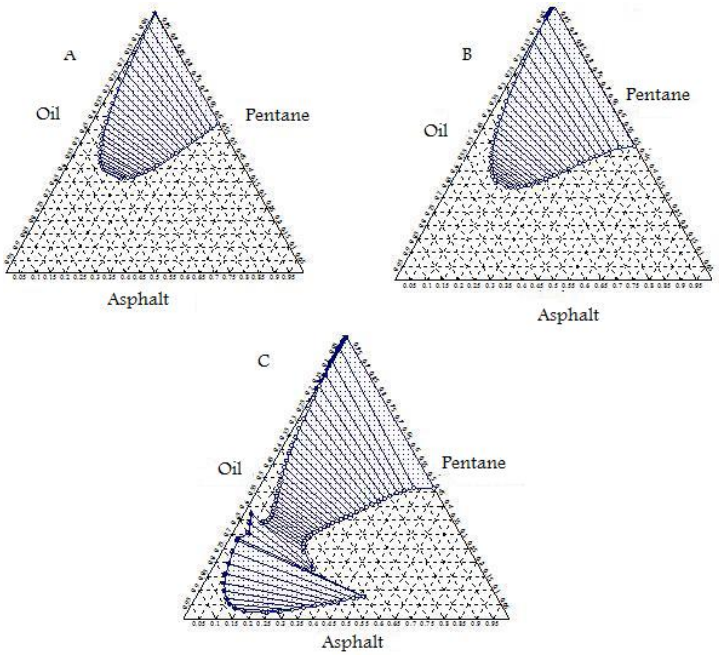


Figure 3.Ternary diagram of the asphalt-oil-pentane mixture at (A)10 bar, (B)33 bar and (C)40 bar.

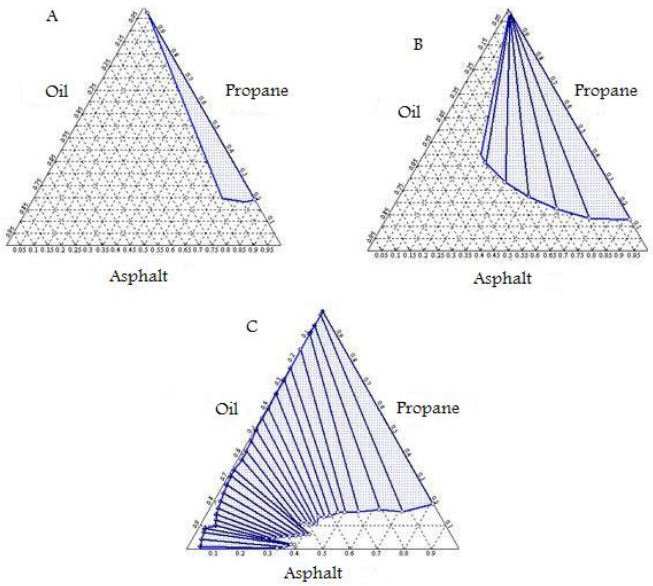


Figure 4. Ternary diagram of the asphalt-oil-propane mixture at (A)15 bar, (B)41 bar and (C)60 bar.

The simulations carried out in the Simulator ASPEN Plus have as aim to know the behavior of the system asphaltene-solvent-oil which is influenced by important parameters, such as: pressure, temperature, component compositions, among others. The knowledge of the system behavior and of the influence of each parameter becomes easier to understand the extraction process. Firstly, the simulation of ternary diagrams provided us an idea of the solubility degree among the components and how this solubility behaved in different operating conditions. To take into account this information, the assembly and operation of virtual plant become easier to be understood, helping to estimate the appropriate parameters in the experiments in the pilot plant developed for our research group.

5. Concluding Remarks

The evaluation of the phase equilibrium through the ternary diagrams is important in order to verify the operational conditions that could be use in the deasphalting extraction column, for different values of temperature and pressure and also for different kinds of solvent used, and to evaluate through these graphs which are the best separation conditions regarding to the liquid-liquid extraction, near to the critical extraction and to the supercritical extraction of asphaltenes from a vacuum residue, in order to obtain lubricant oil. Global understanding of the equilibrium and phase behavior of the system was necessary to develop the supercritical deasphalting pilot plant. This equipment has a complex operation and running, so the development and adaptation of the plant design is worth to carry out experiments in supercritical conditions in a safe and flexible way.

6. Acknowledgement

The authors acknowledgement the financial support from FAPESP
(Process: 06/59978-6 and 07/00031-2), CAPES, CNPq, PETROBRAS and FINEP.

References

1. Loh, W., Mohamed, R.S., Santos, R.G. Crude Oil Asphaltenes: Colloidal Aspects. In Press, 2007.
2. Mendes, M.F., Ferreira, C.Z. and Pessoa, F.L.P., 2005, Deasphaltation of Petroleum Using Supercritical Propane. 2nd Mercosur Congress on Chemical Engineering, 4nd Mercosur Congress on Process Systems Engineering, Enpromer, Costa Verde – Brazil.
3. Rodrigues, J.C. (2006). Internal Report. LDPS, Campinas – Brazil.
4. Siskin, M., Kelemen, S.R., Eppig, C.P., Brown, L.D. and Afeworki, M., 2006, Asphaltene Molecular Structure and Chemical Influences on the Morphology of Coke Produced in Delayed Coking. Energ Fuel, 20, 3, 1227-1234.
5. Wilson, R.E.; Keith, P.C.; Haylett, R.E., 1936, Liquid-propane Use in Dewaxing, Deasphalting, and refining heavy oils. Ind. Eng. Chem., 28, 1065.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Data Reconciliation Practice at a Petroleum Refinery Company in Brazil

José Plácido,[a] Alexandre Almeida Campos,[a] Daniel Ferraz Monteiro,[a]

[a]*Soteica do Brasil Ltda, Rua Dr. Cid de Castro Prado 108, São Paulo 04064-040, Brazil*

Abstract

In spite of being prone to both random and gross errors, measurements are vital for process monitoring, modelling, optimization and control, as well as for plant accounting. However, poor data quality may in general lead to both bad corrective actions and significant deviations between estimated and actual process state, product and feedstock storage and daily Key Performance Indicators (KPI). Data reconciliation, a widely recognized technique that can aid to improve data quality in industrial processes, has being applied since August 2005 at several PETROBRAS petroleum refineries in Brazil. With data reconciliation one is able to find out process leaks, unexpected flow deviations, product losses, errors in storage measuring devices and flowmeter biases, so these findings may be combined with suitable maintenance programs in order to get the desired data quality. In this work we report some results from our daily experience as users of Sigmafine, a data reconciliation package from OSI Software Company. Major common sources of gross errors are illustrated using actual cases taken from one of PETROBRAS refineries, and improvements in data quality and in process performance are stressed. Particularly, potential financial gain are commented, and daily actions performed to circumvent gross error sources provide not only better plant accounting, but also reduce nominal losses and suggest actions to deal to or even eliminate these otherwise unknown trouble sources.
Keywords: data reconciliation, gross errors, plant accounting, petroleum refineries.

1. Introduction

The basic idea in data reconciliation is to adjust each redundant measurement in order to obtain reconciled values which satisfy restrictions like material balances around process units. The smaller the meter tolerance, the smaller will be its correction. A meter *mj* is said redundant when its value can be obtained (e.g. by means of material and energy balances) by the set of all other *mk* measurements, $k \neq j$.

The problem here is to minimize an objective function, e.g. the sum of the squared adjusts pondered by their reciprocal variances, subject to linear restrictions (material balances), and the solution can be found in the literature (Crowe et al, 1983; Mah et al., 1976). Reconciled values constitute a reference set which can be used in data quality analysis and in gross error detection schemes (Serth et al, 1986; Iordache et al., 1985).

One can say that many process improvements certainly result when this kind of analysis is performed in petroleum refineries. We hope the results presented here, taken from our experience working on Sigmafine Project at PETROBRAS, can contribute to disseminate this practice to other process industries also.

In the next section some comments are made about Soteica's Sigmafine Project. In section 3 some common sources of gross errors are described. Section 4 presents some cases with potential financial gains, and then the conclusions follow to close the work.

2. The Sigmafine data reconciliation project at PETROBRAS

The process studied consists of several units, varying from atmospheric distillation and product treatment up to LPG (Liquefied Petroleum Gas) and solvent units. Additional considerations are also taken for the storage park and its interactions with process units, receipts and shipment measuring systems. Water and energy do not enter the dry-basis mass balances considered here.

Reconciliation runs generally consist of 24h periods where all data are appropriately averaged. Because of the dynamic nature of storage tank/sphere alignments, in Sigmafine process models these product/feedstock *transactions* are dynamically created daily at the data reading step. A transaction from a given process unit may have its target tank/sphere changed several times in a reconciliation period, and so it's important having adequate tracking of these changes in order to avoid misleading results.

Data fed into Sigmafine environment have different origins, e.g.:

(a) BDEMQ, a corporate oracle data base containing information concerning storage device measurements, product/feedstock transactions and quality lab analysis;

(b) PI, Plant Information system from OSI Software with information on process measuring devices such as temperature meters, pressure gauges and flow meters;

(c) Truck weights in shipment operations;

(d) Manual changes supplied by the user in order to keep model-to-process data adherence.

In order to reduce the manual user interventions and to automate, whenever possible, tasks such as data treatment and filtering, a new tool called SALA was developed. The SALA interface gets its information from BDEMQ and PI systems, performs tank/sphere volume balances in order to solve for simultaneous transactions and then obtains transferred masses for all transactions and the final stocks for all tanks/spheres. These are the primary data to be used in the subsequent Simafine runs.

An important advantage of the reconciled data over the measured one here concerns to consistency: one always get the same information no matter the data source used. In other words, information conflicts are eliminated by data reconciliation and all users have the same information in all data basis. Another advantage is that values for all non-measured streams are also obtained.

In the next section comments are done about major gross error sources one can expect to find in refineries.

3. Gross error sources in petroleum refineries

We have grouped the error sources into three categories: (1) errors in storage devices; (2) errors in product transactions; (3) errors in flow meters. These are explained in the next sections.

3.1. Errors in storage devices

The "measuring" of amounts of hydrocarbon mass in tanks is not a trivial task. To illustrate this, consider the following formulas:

$$Mass = V_{20} \, x \, d_{20} = (V_T \, x \, F_{T\text{-}20}) \, x \, d_{20}$$

$$V_T = V_G - V_{F\text{-}W} - V_{BSW\text{-}W} = (V_G - V_{F\text{-}W}) \, x \, (1\text{-}P_{BSW} /100)$$

where $V_{F\text{-}W}$ is a function of ($H_{F\text{-}W}$, strapping table) and V_G is a function of (H_T , strapping table, floating roof weight); V_{20} = hydrocarbon volume at 20oC (m^3); d_{20} = hydrocarbon density at 20oC (t/m^3); V_T = hydrocarbon volume at the product temperature, T (m^3); $F_{T\text{-}20}$ = tabulated conversion factor defined as $F_{T\text{-}20} = V_{20} / V_T$ (is a

function of both T and d_{20}); V_G = total fluid volume (m^3); V_{F-W} = volume of free water (m^3); V_{BSW-W} = volume of BSW water (m^3); P_{BSW} = percentage of BSW water (%); H_{F-W} = height of free water (mm); H_T = total height measurement (mm).
Hence, in order to calculate the hydrocarbon mass in tanks one needs measurements of total height H_T, height of free water H_{F-W}, temperature T, percentage of BSW P_{BSW} and density at 20oC d_{20} (and sphere *pressure*). As an additional requirement one needs to know the strapping table and (for floating roof tanks) the roof weight. Strapping tables convert height readings into tank volumes. In case of floating roof tanks, these tables discount the volume displaced by the roof according to the actual height measurement.
Hence, concerning mass determination in storage devices, one can observe that:

- Tank level readings: there exists a problem common to radar-metering systems in tanks that occurs when the product height in the tank gets values bellow the minimal operating height. Below this level the conic waves emitted by the radar can reflect at the tank side walls, generating readings that haven't correlation with the real product height in the tank. As an example, a gross error occurred in streams related to the URFCC material balance corresponding to 4.5% of its feed. This error was caused by a fault in the height measured by the radar of a RAT (Atmospheric Residue) tank.
- Temperature readings: thermocouple mal-functioning may cause false V_{20} volume gains/losses.
- Water level readings: this information is rarely available, since the automatic detection of the water-oil interface does not provide accurate results. In some tanks the zero at the height reading scale implies a great amount of dead volume below it, and then smaller free water contents cannot be appropriately accounted for.
- Product losses through valves: this is a common source of errors in spheres of intermediate propylene, and may difficult getting the correct values for *transactions*.
- Lack of homogenization: see Figure 4. This is a common error in LPG storage spheres as described in section 4.3. The thermocouple is positioned closer to the sphere shell and its readings reflect much more the environmental atmospheric variations then the real average product temperature.
- Strapping table problems: for floating roof tanks an error in the volume evaluation may result if the tank height falls bellow the position where the roof gets supported on their legs, when the position is set to the maintenance one.
- Accumulation of rain water on roof of floating roof tanks: the water weight on the roof presses the product inside the tank causing it to rise through the metering tube. The final effect is a false higher level reading in the tank, which causes several plant accounting problems depending on several factors. Several actions can be devised in order to deal with this problem, depending weather the rain occurred before, during or after the product transaction, and on the drainage efficiency also.

3.2. Errors in product transactions

Here the errors are associated mainly with cultural and human reasoning. Product transactions not always follow exactly the real alignment timing verified in the process plant. This was one of the reasons why we decided to develop the SALA interface. In order to save space, we are not going on giving more details on this topic.

3.3. Errors in flow meters

Even good flow meters can eventually fall into problems and then give bad value readings. These readings may fail because of several reasons. Some of them are explained in the topics that follow.

- Errors of zero and false measurements: these errors are more frequent on plant shutdowns. Possibly one of the pressure gauge legs keeps pressurized after the line is

blocked by the plant operator, and so a false flow reading results. In some cases this may turn difficult to find out, in searching a data base containing information from several months, the periods where the plant was really stopped. The worst case occurs when the process unit operates in a step-wise way, since the non-zero readings of the flow meter may be always present.

- Negative measurements: these measurements may appear when alignment valves are closed. The problem is that, for a period of 24 hours, the negative part contributes for the integration of flow meter readings and then a value smaller than the expected results for the whole period.
- Instrument biases: bias may have different causes, e.g., water accumulation in one of the pressure gauge legs, instrument miscalibration, mal-functioning of signal transmitters and so on. Figure 1 illustrates the latter case for a naphtha flow meter. The maintenance team tried to fix the problem several times without being successful, until someone remembered to check the signal transmitter. As can be seen in the figure, after fixing the error the measurement values started to agree with the reconciled ones, as expected

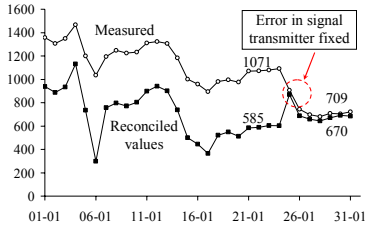


Figure 1: Bias in naphtha flow meter due to signal transmitter malfunctioning (in volume units per day).

- Design errors: in this case the flow meter is not suitable for the service and should be replaced with a new-designed one. Under-dimensioned flow meters generate excessive head losses and should not be used in order to avoid excessive erosion.

4. Some (potential) financial gain cases

4.1. LCO deviation into ODEC storage tanks

LCO (light FCC cycle gas oil), beyond having its own market value and uses, has various operational uses, e.g., as a flushing oil to clean ODEC (decanted oil) in-line filters and obstructed valves, as a make-up fluid to the bottom of equipments during process starting-ups, as a compositional oil for finished products and so on, the former being the principal one (in general the LCO used ends up entering into ODEC tanks). However, this oil consumption is not ordinarily measured. With reconciled estimations nowadays directly visible in PI system, interested people can easily verify whether LCO consumption is excessive or not, and take actions in order to minimize it if necessary.

It was realized also that sometimes there are leaks from the LCO production line directly into ODEC tanks, or, otherwise, possibly, operators may simply use the oil without taking care of the amount being expended. As an example, it can be noted in Figure 2 that initially about 50% of the LCO production was being deviated to the cited LCO sinks. This excessive consumption was then communicated to the process operation team, which, after in site verification, reported there was a leak in the ODEC tank line valve. The leak reduction "point" is marked in the figure, corresponding to about six hours of practically zero leak/consumption operation. After this period the LCO deviation increases again, probably due to operational needs, and stays at a level about 20-30% of the LCO production value. This consumption level is still historically high, but is better than the initial consumption of 50%.

Considering that LCO market price is higher than that of ODEC, it becomes clear that it's highly advisable to minimize the cited LCO deviations into ODEC tanks. It is worth

while remember that part of these LCO deviations represents operational requirements and, as so, cannot be reduced. But certainly there is enough room for improvements here!

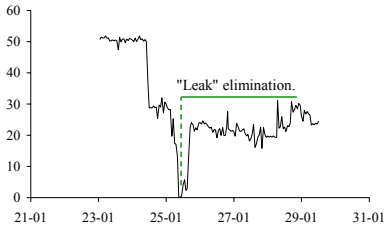


Figure 2: LCO percentage production deviation into ODEC tanks.

4.2. Solvent losses at the Aromatics Removal Unit (UDS)

This process was initially represented as a black-box unit with a feed and two production lines (rubber solvent and BTX – benzene, toluene, xylene). A little later a gross error (bias) has been found in the rubber solvent stream flow meter and, after that, reconciliation results have shown a mass leak in this process unit.

The process operation team has confirmed entrainment losses from this unit to the top drum of the URFCC fractioning tower because of an operational problem in one of the rectifying towers, which operates practically without reflux. The process model in Sigmafine was then modified in order to consider this entrainment between units.

One must note that there are no physical losses but just a product being deviated; this product ends up into gasoline tanks instead of rubber solvent tanks.

Because of the differences between rubber solvent and gasoline market prices, this deviation represents a financial loss. Total reconciled losses from June 2006 to May 2007 amount to 14% of the feed to UDS. Corresponding non-reconciled losses are 71% higher than the reconciled losses. In engineering project studies aiming at minimizing these losses, reconciled values give more realistic estimations for economic viability analysis and thus give more realistic estimations for investment return rates.

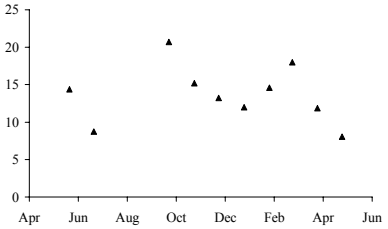


Figure 3: Percentage entrainment losses from the Aromatic Removal Unit, total amounts by month from Jun-2006 to May-2007.

4.3. Lack of homogenization in spheres of LPG

Figure 4 illustrates some typical operations concerning LPG spheres. First of all, sphere C receives LPG from external suppliers (curve 1). Thereafter a LPG transfer takes place from sphere C to sphere B (curve 2). Sphere B receives also the LPG production and so the transfer from sphere C is a simultaneous operation.

In this case, the LPG production is calculated by a V_{20} volume balance around sphere B, since the transferred volume from C to B can be obtained considering the sphere C stock variation. Any error in this transferred volume evaluation causes an error in the LPG production calculation also.

In this example, one can observe a sudden increase in the sphere C temperature at the beginning of the transference (curve 3), which reveals a lack of homogenization in the product inside the sphere. Ideally, this product homogenization should have happened before the transference operation. When calculations were repeated by using the correct temperature, considering also a small stock variation occurred before the transference,

one obtained a gain of 3.1% in the LPG production and a reduction of 1.7% in the LPG received in sphere C from the suppliers.
Taking into account that this kind of problem is not uncommon, one concludes that the annual accumulated differences may amount to a considerable value. One must consider that, with these corrections, improvements in the information quality are also obtained.
This problem has occurred because the temperature read at the only thermocouple installed in the bottom of the sphere does not reflect the actual average of the product temperature. By this reason, before any LPG selling operation the product in the sphere is recycled, a procedure that minimizes the difference between the stock reduction in the sphere and the amount of product registered at the Measuring Site (EMED). However, it's not always operationally possible recycling the products in the LPG spheres. In the case of intermediate propylene spheres, the differences caused by the lack of homogenization are much smaller since there are three installed thermocouples (bottom, middle and top) and a suitable averaged temperature is employed.

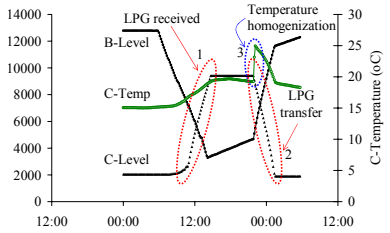


Figure 4: Lack of homogenization in LPG storage spheres.

5. Conclusion

Data reconciliation is an invaluable tool for data quality improvement in the process industry. It can help maintenance engineers to focuses attention on just the most critical instruments and this can save both time and money. Process leaks, product deviations, wrong alignment transactions, faults in stock-metering devices, product losses in general and many other sources of errors can be detected with the aid of this technique.
This work presented illustrative cases where money can be saved provided investments are done and other cases where minor daily differences may result in considerable annual accumulated savings.
In all cases, it can be said that it's worth while using this technique in daily applications to improve plant accounting and getting more reliable results for generating significant process key performance indicators and for use in plant monitoring systems.

6. Acknowledgement

We would like to acknowledge to the many people who helped us during the development of the Sigmafine Project, so many that it wouldn't be possible to name them here.

References

C. M. Crowe et al., 1983, Reconciliation of Process Flow Rates by Matrix Projection. Part I: Linear Case, AIChE Journal, 29, 6, 883-888.
R. S. Mah et al., 1976, Reconciliation and Rectification of Process Flow and Inventory Data, Ind. Eng. Chem. Process Des. Dev., 15, 1, 175-183.
R. W. Serth et al., 1986, Gross error detection and data reconciliation in steam-metering systems, AIChE Journal, 32, 5, 733-742.
C. Iordache et al., 1985, Performance studies of the measurement test for detection of gross errors in process data, AIChE Journal, 31, 7, 1187-1201.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Exploiting Modern Computing Architectures for Efficient Large-Scale Nonlinear Programming

Yu Zhu[a], Daniel Word[a], John Siirola[b], Carl D. Laird[a]

[a]*Dept. of Chemical Engineering, Texas A&M University, MS 3122, Collete Station, TX*
[b]*Sandia National Laboratories, P.O. Box 5800, Albuquerque, NM, 87185*

Abstract

While large-scale nonlinear programming (NLP) has seen widespread use within the process industries, the desire to solve larger and more complex problems drives continued improvements in NLP solvers. Because of physical hardware limitations, manufacturers have shifted their focus towards multi-core and other modern parallel computing architectures, and we must focus efforts on the development of parallel computing solutions for large-scale nonlinear programming. In this paper we briefly describe some existing and emerging architectures for parallel computing and demonstrate the potential of some of these architectures for parallel solution of nonlinear programming problems. In particular, we show the scalability of an integrated design and control problem using two techniques within a multi-core architecture.

Keywords: parallel computing, numerical computing, optimization, nonlinear programming

1. Introduction

Large-scale nonlinear programming (NLP) has proven to be an effective framework for optimization within the process industries. However, the scale of the NLP problems of interest to both industry and academia continues to grow increasingly large, often outstripping the capacity of modern workstations. While significant theoretical advancements have been made in general (serial) NLP algorithms, few parallel algorithms exist and off-the-shelf tools often are not able to provide solutions efficiently for problems larger than the capacity of a single workstation. Furthermore, computer chip manufacturers are no longer focusing on increasing clock speeds and instruction throughput, but rather on multi-core architectures and hyper-threading. This means that the "free" performance improvements that we have enjoyed as a result of advances in computing hardware will no longer be possible unless we develop algorithms that are capable of utilizing modern concurrent architectures efficiently. Furthermore, modern computing architectures are becoming more complex and the parallel computing community is seeing a shift from the use of general purpose CPUs towards specialized processors and heterogeneous architectures. In order to continue to see performance improvements, effective utilization of these parallel architectures will be necessary. This reflects a significant paradigm shift in the design and implementation of numerical optimization algorithms that can exploit emerging architectures.

All parallel architectures are not created equal. Distributed-memory Beowulf clusters are a scalable solution that can provide for many available processors, however, the communication limitations among machines may present a significant bottleneck. On the other hand, multi-core computing systems promise fast inter-process communication through shared-memory, however, memory bandwidth issues can significantly limit

parallel performance. Several fundamentally new architectures are being introduced into the scientific computing realm: Graphics Processing Units (GPUs) and the Cell Broadband Engine Architecture (Cell) provide potential for large-scale scientific computing, but their architectures are complex and present significant challenges not encountered in general computing. The Cray XMT massively threaded supercomputer represents a fundamentally new architecture that mitigates many of the memory latency problems seen in common platforms but requires a level of fine-grained parallelism heretofore unseen in parallel algorithm development.

In this paper, we contrast the different architectures available for modern scientific computing and outline their potential for efficient parallel nonlinear optimization. As a detailed case study, we consider the parallel solution of the integrated design and control of heat-integrated distillation columns under uncertainty using both serial and parallel algorithms on multi-core machines.

2. Parallel Computing Architectures for Scientific Computing

There are a number of emerging architectures for parallel scientific computing. In this section, a few of the proven and promising architectures are described in further detail.

2.1. Distributed-Memory Cluster Architecture

Beowulf clusters are an example of a distributed-memory parallel processor architecture. The cluster is built by networking affordable desktop computers through standard Ethernet or specialized networking technologies. The system consists of a primary computer functioning as a master node, controlling access to the compute nodes. Each compute node has its own local RAM. Since the memory is not shared with other nodes, communication among nodes occurs over the network. This is typically done via a Message-Passing Interface (MPI) although other technologies and paradigms exist. Communication via Ethernet is the biggest cause of latency, so for a program to run efficiently, this communication and the need for synchronization must be kept to a minimum. These architectures are most appropriate for coarse-grained problems that require limited communication (Hairgrove et al, 2001), and our previous work has demonstrated that this architecture is highly appropriate for parallel solution of large-scale structured nonlinear programs (Zavala et al., 2007). Grid computing represents an extreme case of a distributed memory architecture where compute nodes are typically heterogeneous and geographically distributed, with communication over the internet.

2.2. Multi-core Processor Architectures

Multi-core architectures allow for shared-memory and parallel computation within a single node. Each of the processing cores has access to the same local memory and these architectures promise significantly faster communication among processes (or threads) through local RAM. Nevertheless, in most multi-core systems access to local RAM is shared through a common memory bus, and the performance of individual processes may deteriorate as each process competes for access to the local memory it needs. Even with a sufficient number of processors, the memory bandwidth can become a bottleneck and deteriorate the expected benefits from parallel computing. In this current work, we demonstrate that multi-core architectures are also appropriate for parallel solution of large-scale structured optimization problems, however, they do not scale as well as distributed architectures for the class of problems studied.

2.3. GPU Architecture

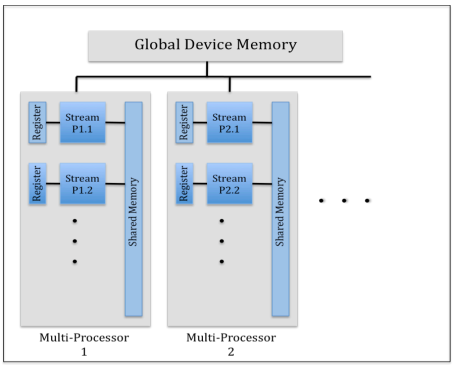


Figure 1: Diagram of GPU Architecture

Graphics Processing Units (GPUs) are another type of multi-core processor that has recently emerged within the scientific computing community. GPUs typically contain several hundred basic processing cores, however, these cores are not general purpose CPU cores. Therefore, while these systems may give access to many hundreds of cores at an affordable price, these cores are limited in their capability. For example, the general GPU architecture shown in Figure 2 is modeled after the NVidia Tesla GPU units for scientific computing. Here, each GPU device contains a number of *multiprocessors*, each with a number of single-instruction-multiple-data (SIMD) *stream processors*. These architectures have complicated memory structures that must be considered. The different types of memory in a GPU can be grouped into categories based on their scope. Registers serve individual processors; shared memory, constant cache, and texture cache serve multiprocessors; global device memory serves all cores (Richardson and Gray, 2008). While this memory hierarchy allows for very low latency at the processor level, access to global device memory has high latency. Thus, the problem must be highly parallel so that the program can break it into enough threads to keep the individual processors busy (Wolfe, 2008). In spite of these and other drawbacks, there is significant potential for these architectures in certain scientific computing tasks and examples of their success are numerous. Current efforts indicate significant potential for efficient parallel solution of linear systems. With regards to parallel nonlinear optimization, this provides a promising avenue for immediate use of these architectures since most NLP algorithms require the solution of a large linear system at each iteration.

2.4. CELL Architecture

The Cell (Cell Broadband Engine Architecture) is similar to the architecture of the GPU systems in that there are a number of non-general purpose processing cores that have shared access to various levels of memory. The Cell processor contains eight Synergistic Processing Elements (SPEs). Each SPE contains a Synergistic Processing Unit (SPU) that operates using an SIMD architecture.

Processors are mated closely to their own independent memory allowing for very low latencies between the processor and memory. However, this is a distributed memory system (Kurik et al, 2008). The Cell system has gained popularity, as it is available rather inexpensively and fully supports the Linux operating system. As with the GPU architectures, there have been significant efforts towards efficient parallel solution of linear systems on the Cell processor (Kurik et al, 2008), and the use of parallel linear solvers provides potential for parallel nonlinear optimization.

2.5. Cray XMT Architecture

The Cray XMT system is designed to give a relatively inexpensive, scalable multithreading, shared-memory supercomputing platform (Cray Inc., 2008). It is built with up to 96 processors per cabinet. Each processor accommodates 128 fine-grained hardware streams and is associated with its own memory system. Since these memory systems are linked together the system can function as a shared-memory computer

(Chavarria-Miranda, 2008). The latency associated with shared-memory systems is masked in this supercomputer through multithreading. Multiple threads give processing efficiency by skipping threads that are waiting for data from memory and running threads that have data available This is ideal for data-intensive applications requiring irregular memory access (Chavarria-Miranda, 2007).

3. Parallel Nonlinear Optimization on Modern Architectures

A wide variety of optimization formulations can result in very large-scale problems. While significant advancements have been made in nonlinear optimization algorithms (Gould et al, 2004), as the size of the problem increases, general off-the-shelf tools may not be able to provide solutions efficiently, and we may need to make use of specialized parallel solution strategies. Fortunately, very large-scale optimization problems are almost always inherently structured since they are necessarily formulated from a repeating set of mathematical expressions (Gondzio and Grothey, 2005), and approaches can be developed that exploit the structure of the problem and allow solutions in parallel.

Common approaches for decomposition of structured problems include problem-level techiques based on Bender's and Lagrangian decomposition. However, given the complicated nature of modern parallel computing architectures and the number of architectures with non-general purpose processing cores, it is important to keep the parallelism as straightforward as possible. For example, it is not possible with SIMD architectures to have different processors solving completely different subproblems since each processor must be performing the same instruction. Internal decomposition approaches for structured problems provide an alternative to problem-level decomposition techniques. Instead of modifying the optimization problem, the internal linear operations of a particular host solver are specialized to exploit problem specific structure. Parallel full-space approaches have been developed for linear and quadratic programming problems (Gondzio and Gothrey 2005, Gondzio and Sarkissian, 2003). In previous work, we developed a parallel algorithm, SCHUR-IPOPT for large-scale general nonlinear optimization problems (Zavala et al, 2007) using a parallel Schur-complement decomposition approach.

The internal decomposition approach shows significant promise for effective parallel algorithms on modern architectures since it relies on the parallelism of linear operations appropriate on SIMD systems. Furthermore, as scientific computing becomes mainstream on these architectures, linear solvers are and will continue to be the focus of development. The internal decomposition approach is an appropriate vehicle for efficient nonlinear optimization of structured large-scale problems through the use of parallel linear solvers. In this work, we demonstrate our Schur-complement internal decomposition approach (as described in Zavala et al., 2007) on multi-core systems, considering the optimal design of heat integrated distillation columns under uncertainty.

4. Case Study and Conclusions: Optimal Design of Heat Integrated Distillation Columns Under Uncertainty

The internal heat integrated distillation column has been regarded as a potential energy saving configuration for separation. The rectifying section and the stripping section in this process are separated into two columns operating at the different pressures. To adjust the pressures, a compressor and a throttling valve are installed between the two sections. Unlike the case of the conventional columns, the condenser and reboiler are not required in this process. While past research has considered conceptual design and

dynamic control, little research in this area considers potential uncertainties in the design phase. In order to minimize the total annual cost under uncertainties, a rigorous mathematical model is derived from the mass and energy balances coupled with the equilibrium relationships. Detail model information can be found in Liu and Qian (2000) ,while the design formulas for installation and operating cost can be found in Huang (2007), Olujic (2006), and Douglas (1988).

The three main design variables selected include: the diameter of distillation columns, the heat transfer area, and the brake horsepower. Two types of uncertainty are considered in this work, unmeasurable and measureable uncertainty. The relative volatility of benzene and toluene, which may vary from 2.517 to 2.117 is selected as an unmeasureable uncertainty. The range is discretized by selecting 7 points assuming a normal distribution. Measureable uncertainties include the concentration of the feed flow (Zf), which varies from 0.55 to 0.45, and the feed flow rate (F), which varies from 90 to 110 (mol/s). Both of them are measurable during operation and may be compensated by two control variables, rectifying section pressure and thermal condition in the feed flow. Here, we discretize the uncertainty in Zf and F by selecting 15 points with normally distributed assumptions. In this large-scale structured problem, there are 1575 scenarios and approximately 300,000 total variables including 453 common variables (the design and control variables common across the scenarios). In addition to the model, controllability constraints are included in the formulation.
To demonstrate the scalability of different techniques, we solve the problem with an increasing number of scenarios (225 scenarios for each of 7 blocks). We tested the scalability using the parallel internal decomposition strategy that we have implemented within SCHUR-IPOPT, which is designed to exploit the specific structure induced by the multi-scenario problem. We also tested the scalability using the general purpose shared-memory linear solver PARDISO (Schenk and Gärtner, 2006, 2006). We do not have a version of the solver interfaced with the nonlinear optimization tool Ipopt. Therefore, to test scalability, we measured the computational time for PARDISO to perform the numerical factorization and backsolve of the same linear system arising from a particular iteration of the SCHUR-IPOPT solver. It is important to note that PARDISO performs a symbolic factorization prior to performing the numerical factorization. However, within an optimization context, the symbolic factorization would only need to be performed once, at the beginning of the optimization, therefore this time was not included in the results. For this reason, coupled with the fact that PARDISO is simply solving the linear system and not performing the other serial operations required by SCHUR-IPOPT, the comparison is conservative (that is the

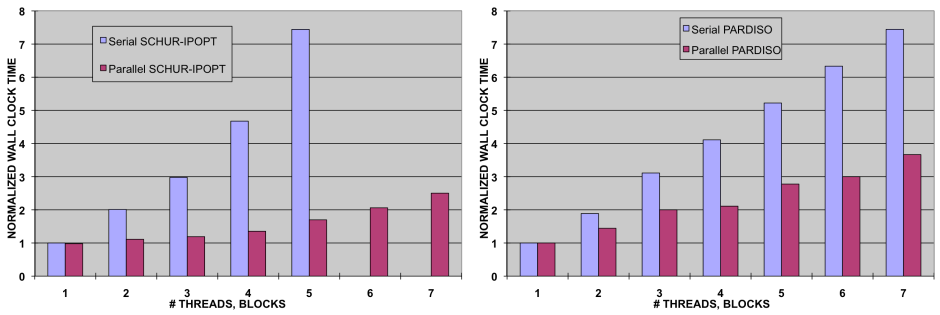


Figure 2: Parallel Scalability Results on a Multi-core System

comparison favors PARDISO somewhat). Figure 2 shows the normalized wall clock time for the two techniques. The x-axis shows the number of blocks, with an additional thread or process added along with each block. We can see that, in both cases, the scalability as we increase the size of the problem is impressive. We see a slight increase in the wall clock time as we increase the size of the problem, which is primarily due to memory bandwidth issues. In previous work, we have shown that the wall clock time is almost completely flat up to 32 processors when a distributed memory architecture is used. Furthermore, while the SCHUR-IPOPT approach is specialized for the particular problem structure, the general-purpose solver PARDISO performs comparatively well. Further tests with PARDISO integrated within the Ipopt optimization package will be necessary to draw stronger conclusions.

Clearly, the multi-core parallel architecture presents advantages in reducing the solution time of large problems. However, the multi-core architecture is limited by memory bandwidth and the number of available cores. Current research in our group is investigating the use of advanced scientific computing architectures like GPUs and the Cell processor for parallel solution of large-scale structured nonlinear programming problems.

References

Cray Inc, 2008, XMT Product Technology, http://www.cray.com/Products/XMT/Product/Technology.aspx

D. Chavarria-Miranda, A. Marquez, J. Nieplocha, K. Maschhoff, C. Scherrer, 2008, Early Experience with Out-of-Core Applications on the Cray XMT, IEEE Xplore.

D. Chavarria-Miranda, A. Marquez, J. Nieplocha, F. Petrini, 2007, Unconventional Architectures for High-Throughput Sciences, IOP Publishing, SciDAC Review.

Douglas, JM. 1988, Conceptual design of chemical processes. New York, *McGraw-Hill.*

J. Gondzio and A. Grothey. Exploiting structure in parallel implementation of interior point methods for optimization. Technical report, School of Mathematics, The University of Edinburgh, July 2005. Technical Report MS-04-004.

J. Gondzio and R. Sarkissian. Parallel Interior Point Solver for Structured Linear Programs. Mathematical Programming, 96(3):561–584, 2003.

N. I. M. Gould, D. Orban, and Ph. L. Toint. Numerical methods for large-scale nonlinear optimization. RAL-TR-2004-032, Rutherford Appleton Laboratories, UK, Nov. 2004.

W. Hargrove, F. Hoffman, T. Sterling, K. Chernush, 2001, The Do-It-Yourself Supercomputer, Scientific America, 285, 2, 72-80.

Huang, K, 2007, Design and control of an ideal heat-integrated distillation column system separating a close-boiling ternary mixture, Energy, 32, 2148.

J. Kurak, A. Buttari, P. Luszczek, J. Dongarra, 2008, The PlayStation 3 for High-Performance Scientific Computing, Computing in Science & Engineering, 10, 3, 84-87.

Liu, X., and Qian, J., 2000, Modeling, Control and Optimization of Ideal Internal Thermally Coupled Distillation Columns. Chem. Eng.Technol. 23 (3), 235-241.

Olujic, Z., 2006, Conceptual design of an internally heat integrated propylene-propane splitter, Energy, 31, 3083.

A. Richardson, A. Gray, 2008, Utilisation of the GPU architecture for HPC, HPCx Consortium.

O. Schenk and K. Gärtner, Solving Unsymmetric Sparse Systems of Linear Equations with PARDISO, Journal of Future Generation Computer Systems, 20(3):475--487, 2004.

O. Schenk and K. Gärtner, On fast factorization pivoting methods for symmetric indefinite systems, Elec. Trans. Numer. Anal., 23:158--179, 2006.

M. Wolfe, 2008, Compilers and More: GPU Architecture and Applications, HPCwire, September, 1-3.

V.M. Zavala, C.D. Laird, and L.T. Biegler. “Interior-point decomposition approaches for parallel solution of large-scale nonlinear parameter estimation problems”, Chemical Engineering Science, 63 (19), pp. 4834-4845, 2008.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Entropy Production Analysis in Extractive Distillation Using Non-Equilibrium Thermodynamics and a Rate Based Model

Diego F. Mendoza, Carlos A. M. Riascos

Department of Chemical and Environmental Engineering, National University of Colombia, Cra. 30 No. 45-03 Edif. 453 Of 300, Bogotá, Colombia.
Grupo de Ingeniería de Sistemas de Procesos.

Abstract

Entropy generation (EG) in an extractive distillation column for ethanol dehydration was determined. The rate based model and non-equilibrium thermodynamics (NET) leads to quantify separately the EG due to interface heat and mass transfer and mixing. Additionally, this model allows estimating the effects of operational conditions such as reflux ratio and feed condition on the EG. The results obtained show that mixing effects at the feeding points (hydrated ethanol and separation agent) and interface mass transfer are the main sources of EG with 55% and 41% of total EG, respectively. The significance of this work is the use of a rate based model and NET theory as a tool for energy use analysis in multicomponent distillation processes. EG analysis is the starting point to define new operational strategies and design characteristics that allow a more efficient energy use.

Keywords: extractive distillation, non-equilibrium thermodynamics, entropy production.

1. Introduction

Since the industrial revolution, the energy consumption has incremented vigorously, at the beginning; there was not preoccupation for an efficient use of the energy, because of unknowing on energy sources limitation and environmental effects of industrial activities. The global energy crisis, in the late 1970s, and the current energy generation costs have forced the reevaluation of the energy use (Linnhoff and Flower, 1978).

Separation by distillation is a major energy demanding operation in chemical process industries. Due to their low energy efficiency several strategies to improve it have been developed, distillation with heat pump and multiple-effect distillation (Seider et al., 2003) are useful techniques for saving energy; however, they can be applied only to mixtures with special characteristics; internally heat-integrated distillation columns has been developed as an alternative, at this configuration the energy consumption can be reduced in about 40% but the controllability of the systems requires special attention (Nakaiwa et al., 2003).

Recently, the energetic efficiency analysis has been focused on the loss of energy quality; it may be estimated by exergy analysis (Demirel, 2007) or using non-equilibrium thermodynamics (NET) (Liang et al., 2006). Exergy is defined as the maximum available work, in a source, bringing it to the equilibrium with its surrounding (Demirel, 2007). The exergy loss is proportional to the EG and it is estimated by the difference between exergy going into and out of the system. This

analysis provides a systematic method to diagnose whether and to what magnitude it is possible to increase the process efficiency. Whereas, the NET approach identifies the mechanism of EG and shows the location and causes of exergy losses (Liang et al. 2006), accordingly, it allows to predict the effect of operational conditions and equipment design on energy efficiency and to define the best conditions for maximizing the energy usage.

2. The Rate Based Model

The rate based model (Taylor and Krishna, 1993), considers liquid and vapor phases separated by an interface through heat and mass transport occur. Transport through both phases occurs due to temperature and concentration gradients located in the liquid and vapor films. Vapor and liquid compositions at the interface are related by equilibrium relationships. The equation set that describes the rate based model is presented in Taylor and Krishna, 1993.

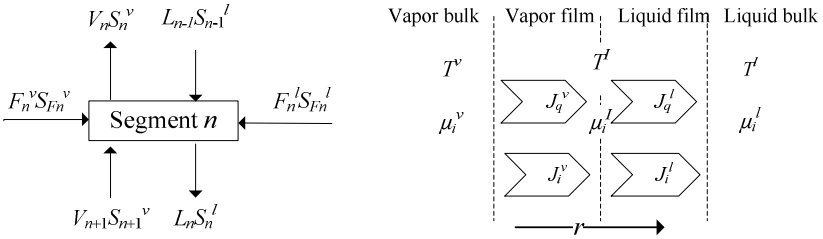


Figure 1. A segment in the distillation column (left), fluxes and forces in the system (right).

3. Entropy Production

In a distillation column three main sources of EG can be identified, namely, heat and mass transfer and mixing. EG due to transfers is calculated using NET theory in the liquid and vapor films, whereas the EG due to mixing is indirectly calculated as the difference between the global EG and the entropy generation in the films.

3.1. Global Entropy Production

The entropy balance (EB) is used to quantify the total EG in a column segment as well as in the boiler and condenser.

The EG in the *n*-th column segment, Figure 1, is:

$$S_n^p = V_nS_n^v + L_nS_n^l - V_{n+1}S_{n+1}^v - L_{n-1}S_{n-1}^l - F_{Fn}^vS_{Fn}^l - F_{Fn}^lS_n^l \tag{1}$$

Where V, L, F are vapor, liquid and feed flows; S is the molar entropy, subscripts n-1 and n+1, identify the upper and lower segments, superscripts v and l identify vapor and liquid phases. EG in the total condenser and boiler are:

$$S_C^p = V_1(S_C^l - S_1^v) - \frac{Q_c}{T_c} \qquad S_B^p = V_BS_B^v + BS_B^l - L_nS_n^v - \frac{Q_B}{T_B} \tag{2), (3}$$

Where Q_C, T_C , Q_B and T_B are heat and temperature at the condenser and boiler.

3.2. Entropy Production in the Films

According to the NET theory the local EG rate per unit of volume, σ, is the sum of the products of fluxes and forces related to heat and mass transfer in the films, Figure 1.

$$\sigma = J_qX_q + \sum_{i=1}^{C} J_iX_i \tag{4}$$

Where J_q and J_i are the diffusive flux of heat and mass, and X_q and X_i are the thermodynamic forces for heat and mass transfer, given by:

$$X_q = -\frac{\partial}{\partial r}\left(\frac{1}{T}\right) \qquad X_i = -\frac{\partial}{\partial r}\left(\frac{\mu_{i,T}}{T}\right) \qquad (5), (6)$$

Where T is the absolute temperature and $\mu_{i,T}$ is the chemical potential of species i, at constant temperature. The Gibbs-Duhem theorem leads to rewrite Equation (4) in terms of mass fluxes in the laboratory frame of reference, N_i, instead of interfacial one:

$$\sigma = J_q X_q + \sum_{i=1}^{C} N_i X_i \qquad (7)$$

The EG in the liquid film is:

$$\int_{V^{lf}} \sigma dV = \int\int J_q^l X_q^l dr dA + \sum_{i=1}^{C} \int\int N_i^l X_i^l dr dA \qquad (8)$$

Where V^{lf} is the volume of the liquid film, r is the film thickness and A is the liquid-vapor interface area. The entropy in vapor film is estimated analogously. Now, the RHS in later equation (and in the one for vapor film) must be integrated through the film thickness. For terms associated to the mass transfer we have:

$$\int_r N_i^v X_i^v dr = -N_i^v \left(\frac{\mu_{i,T}^I - \mu_{i,T}^v}{T^{vf}}\right) \qquad \int_r N_i^l X_i^l dr = -N_i^l \left(\frac{\mu_{i,T}^l - \mu_{i,T}^I}{T^{lf}}\right) \qquad (9), (10)$$

For vapor and liquid phase, chemical potentials differences are:

$$\frac{\mu_{i,T}^I - \mu_{i,T}^v}{T^{vf}} = R \ln\left(\frac{\hat{\phi}_i^I y_i^I}{\hat{\phi}_i^v y_i^v}\right) \qquad \frac{\mu_{i,T}^l - \mu_{i,T}^I}{T^{lf}} = R \ln\left(\frac{\gamma_i^l x_i^l}{\gamma_i^I x_i^I}\right) \qquad (11), (12)$$

The fugacity and activity coefficients, $\hat{\phi}_i$ and γ_i, are calculated at the average vapor and liquid film temperatures, respectively. For terms associated to heat transfer we have:

$$\int_r J_q^v X_q^v dr = \overline{J}_q^v \left(\frac{1}{T^I} - \frac{1}{T^v}\right) \qquad \int_r J_q^l X_q^l dr = \overline{J}_q^l \left(\frac{1}{T^l} - \frac{1}{T^I}\right) \qquad (13), (14)$$

Where $\overline{J}_q^v$ and $\overline{J}_q^l$ are the average heat fluxes in the vapor and liquid films:

$$\overline{J}_q^v = \frac{\int_r J_q^v dr}{\delta^v} = h^v (T^v - T^I) \qquad \overline{J}_q^l = \frac{\int_r J_q^l dr}{\delta^l} = h^l (T^I - T^l) \qquad (15), (16)$$

Here h^v, h^l are the film heat transfer coefficients at the vapor and liquid sides. Finally, EG rate due to interface heat and mass transfer for a segment (n) of the column, $S_n^{p,f}$, is given by the sum of EG in vapor, $S_n^{p,vf}$, and liquid, $S_n^{p,lf}$, films:

$$S_n^{p,vf} = \int_{V^{vf}} \sigma dV_n = \overline{J}_q^v \left(\frac{1}{T^I} - \frac{1}{T^v}\right) - R \sum_{i=1}^{c} N_{i,n}^v a_n \ln\left(\frac{\hat{\phi}_i^I y_i^I}{\hat{\phi}_i^v y_i^v}\right)_n \qquad (17)$$

$$S_n^{p,lf} = \int_{V^{lf}} \sigma dV_n = \overline{J}_q^l \left(\frac{1}{T^l} - \frac{1}{T^I}\right) - R \sum_{i=1}^{c} N_{i,n}^l a_n \ln\left(\frac{\gamma_i^l x_i^l}{\gamma_i^I x_i^I}\right)_n \qquad (18)$$

4. Simulation

The case to be analyzed is a pilot size, extractive distillation column (Table 1) which

process 50 l/h of an ethanol-water (EtOH-W) mixture, the entrainer used is ethylene glycol (EGly). Column simulations were carried out using the Ratefrac model from Aspen Plus® processes simulator.

Table 1. Specifications for simulation

Condenser		Total (segment 1)
Boiler		Partial (segment 30)
Packing	Type	Pall, random, metal
	Size	15.875 mm
Total packed height		10 m
Column diameter		76.2 mm
Number of segments		30
Reflux ratio (molar)		1
Distillate rate		0.651 kmol h^{-1}
Pressure		1 bar
Feed 1 (Entrainer)	Location	Segment 3
	State, T & P	Liquid; 333.15 K;1 bar
	Flow rates (kmol h^{-1})	EGly=0.8331; W=0.0509; EtOH=0.0
Feed 2 (W-EtOH)	Location	Segment 16
	State, T & P	Liquid, 343.15 K, 1 bar
	Flow rates (kmol h^{-1})	EGly=0.0; W=0.1149; EtOH=0.7691
Thermodynamic model	Liquid phase	NRTL
	Vapor phase	Ideal
Transport properties	Mass transfer coefficients	Onda correlation
	Heat transfer coefficients	Chilton-Colburn analogy

5. Results and Discussion

5.1. Current Operation

According to EB equation, Figure 2A, the 48% of total EG is located in the boiler, 51% comes from the packed column and only 1% is generated in the condenser. The high value of the EG in the boiler is due to heat transferred there causes the biggest changes in temperature and composition in the lower column section, Figures 2C and 2D, while in the condenser temperature and composition changes are little, because in it takes place the condensation of high purity EtOH.

In the packed column the main irreversibility source is the mixing at the feeding points, with 86% of the total EG in the packed column. Interfacial heat and mass transfer, calculated by NET, Figure 2A, generate only 14% of EG in the column. This behavior departs from works of de Koeijer and Rivero (2003), Tsirling et al. (2006) and Bandyopadhyay (2007) who claim that EG due to mass transfer is the main irreversibility in a distillation column, but agrees with the work of Tarap and Ishida (1996). The work of de Koeijer and Rivero analyzed the EG in ethanol rectification while we analyze EG in ethanol dehydration. In a rectification, the distance between operating and equilibrium lines is bigger than in extractive distillation, therefore driving forces for heat and mass transfer are higher for ethanol rectification than in our process. The relative contribution of heat (12%) and mass (2%) transfer to EG agrees with aforementioned works, where mass transfer effect is the most important irreversibility.

5.2. Feed Ratio

The Figure 3A shows that EG is reduced when feed ratio (FR) decrease, and this is due mainly to reduction on mixing effects, and the heat load in the boiler, Figure 3B. However there is minimum FR, near to 1, which makes feasible the operation i.e., to get high purity ethanol (>99%) in the distillate.

5.3. Entrainer Feeding Temperature

The influence of the entrainer temperature, Figure 4A, shows that EG is sensitive to this variable, because of the irreversibility in mixing decrease when the difference between

entrainer feed and internal flows are reduced, Figure 4B. One important thing to remark is entrainer feeding temperature, at the analyzed values, does not have influence on the distillate composition which is advantageous for the process.

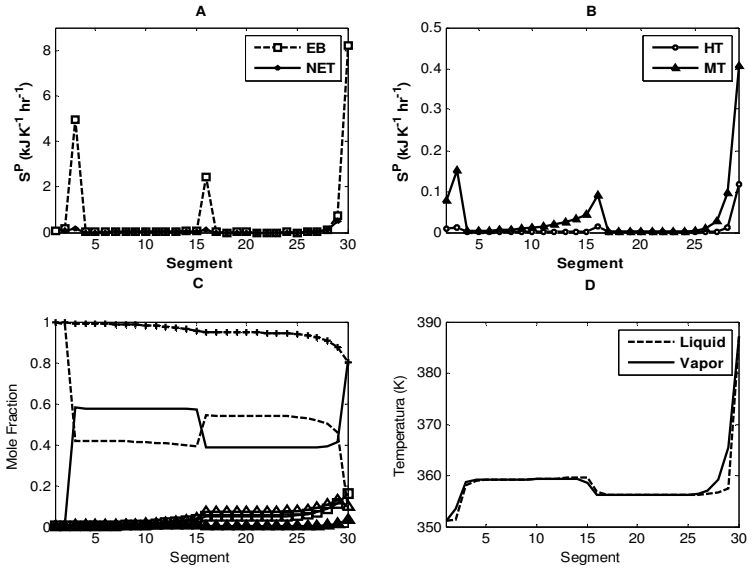


Figure 2. EG for the current operation: A. Entropy production profile calculated form EB equation and NET. B. EG due to heat (HT) and mass transfer (MT). C. Composition profile: --- EtOH(L), + EtOH(V), Δ W(L), □ W(V), − EGly(L), ▲ EGly(V). D. Temperature profiles.

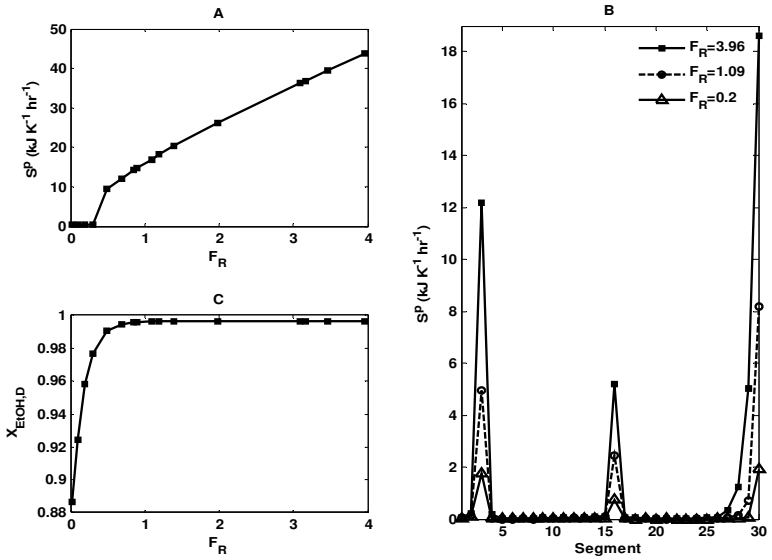


Figure 3. EG related to feed ratio. A. Total EG. B. profile of EG for different feed ratios (FR). C. Mole fraction of ethanol in distillate.

Other variables were evaluated (feeding temperature for W-EtOH mixture and reflux ratio) but the most important effects on the EG are those related with FR and entrainer feeding temperature.

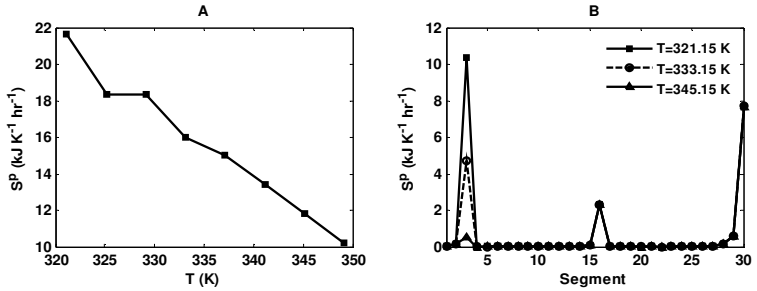


Figure 4. EG related to feed streams temperatures. A. Total EG in function of entrainer temperature. B. Profile of EG rate vs. entrainer temperature.

6. Conclusions

For Ethanol dehydration, EG take place in reboiler (48%) and column (51%) principally. Mixing effects in feeding segments have the major influence on the column`s EG (86%) followed by mass transfer (12%); heat transfer effects on EG has the least influence (2%). In segments different to feeding points, mass transfer is the major contributor to EG.

Among process variables analyzed feed ratio and entrainer temperature showed the major influence on EG. Low feed ratios and high entrainer temperatures diminish the EG due to mixing and heat load in the boiler.

EG analysis using EB and NET are valuable and complementary tools for a detailed mapping of the different entropy production sources. This analysis is useful to establishing the main process variables to be considered in the energy optimization of the distillation column.

References

S. Bandyopadhyay, 2007, Thermal integration of a distillation column through side-exchangers, Chemical Engineering Reserch and Design, 85, A1, pp. 155-166.

Y. Demirel, 2007, Nonequilibrium thermodynamics. 2 edn., New York: Elsevier.

G. de Koeijer and R. Rivero, 2003, Entropy production and exergy loss in experimental distillation columns, Chem. Eng. Sci., 58, pp. 1587-1597.

Y.C. Liang, Z. Zhou, Y.T. Wu, J. Geng and Z.B. Zhang, 2006, A nonequilibrium model for distillation processes, AIChE Journal, 52, 12, pp. 4229-4239.

B. Linnhoff and J.R. Flower, 1978, Synthesis of heat exchanger networks: I. Systematic generation of energy optimal, AIChE Journal, 24, 633.

M. Nakaiwa, K. Huang, A. Endo, T. Ohmori, T. Akiya and T. Takamatsu, 2003, Internally heat-integrated distillation columns: a review, Trans IChemE, Vol 8, Part A, pp. 162-177.

S.K. Ratkje, E. Sauar, E.M. Hansen, K.M. Lien and B. Hafskjoild, 1995, Analysis of EGRs for design of ditillation columns, Ind Eng Chem Res, 34, pp. 3001-3007.

W.D. Seider, J.D. Seader and D.R. Lewin, 2003, Product and process design principles 2e, John Wiley & Sons, pp. 346-351.

R. Taylor and R. Krishna, 1993, Multicomponent mass transfer, New York: John Wiley & Sons.

R. Tarap and M. Ishida, 1996, Graphic exergy analysis of processes in distillation column by energy-utilization diagrams, AIChE Journal, 42, 6, pp. 1633-1641.

M. Tsirling, D. Zubov and A. Barbot, 2006, Consideration of irreversibility factors for binary distillation, Theoretical foundations of chemical engineering, 40, 3, pp. 245-252.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Generative Modeling of Holonic Manufacturing Execution Systems for Batch Plants

Milagros Rolón[a], Mercedes Canavesio[b], Ernesto Martinez[a]
[a]*INGAR (CONICET/UTN), Avellaneda 3657, S3002GJC Santa Fe, Argentina*
[b]*UTN, Lavaise 610, S3004EWB Santa Fe, Argentina*

Abstract

Most batch plants are constantly subjected to unplanned disruptive events such as arrivals of rush orders, raw material delays or equipment breakdowns along with a multitude of interactions in the supply chain demanding automatic on-line rescheduling and execution control. For responsiveness and agility at the shop-floor, a holonic batch control architecture based on autonomic units is proposed. A generative modeling of the proposed holonic manufacturing execution system (MES) is presented in order to evaluate its emerging behavior and macroscopic dynamics in a multiproduct batch plant. A simulation model of the proposed holonic MES for a case study was implemented in *Netlogo*. Different scenarios are considered to assess disturbance rejection capability by the holonic MES. Results obtained are very promising.

Keywords: Generative Modeling, Agent-Based Simulation, Batch Control, Manufacturing Execution System.

1. Introduction

Centralized and hierarchical control approaches are good for optimizing static schedule models but lacks agility and responsiveness mainly due their rigid and centralized nature of traditional control structures. To trade-off *a priori* schedule optimization and industrial practice in the face of disturbances most common MESs in use today (Kletti, 2007) heavily resort to a given schedule so as to focus control only on handling details and contingencies in task execution. One such manufacturing control architectures is the holonic MES implemented by a multi-agent system along the PROSA architecture (Valckenaers et al, 2007) where the MES tries to follow a given schedule and make informed choices when it cannot follow the schedule exactly (Verstraete et al, 2008). Along similar ideas, the ADACOR architecture (Leitao et al, 2005) alternates between stationary states, where system control relies on supervisors and coordinator levels, and transient situations, triggered by the occurrence of disturbances where the MES switches its decision-making policy to a *heterarchical* structure. In this work, for agility and responsiveness decentralized schedule generation and execution control are tightly integrated through holons designed as autonomic units that managed orders and resources at the shop floor level. An interaction mechanism has been designed using the Prometheus-Hermes design methodology and a simulation model (North and Macal, 2007) was used to allow an objective evaluation of the emergent dynamic behaviors.

2. Holonic Manufacturing Execution Systems

Despite some sophisticated control methods, traditional production planning and control has serious weak points in the planning and scheduling of production orders. In that respect Manufacturing Execution Systems (MES) have the aim to make the value-

adding activities transparent and to react quickly and cost-effectively to meet new requirements (Kletti, 2007).

The holonic MES designed in this paper is the result of ongoing interactions among decentralizing decision-making agents. These agents are conceived as autonomic manager units (IBM Corp., 2006) and when integrated to a physical or abstract object (managed element) they conform the holons (HMS Consortium, 2009). As shown in Fig. 1, each agent playing its role implements the monitor-analyze-plan-execute (MAPE) loop which comprise both scheduling and control for a given order or resource. For the agent to be self-managing regarding its managed object, it must have an automated method to collect the details it needs from the manufacturing system (monitor function); to analyze those details to determine if something needs to be changed (analyze function); to create a plan, or sequence of actions that specifies the required changes (plan function); and to perform those actions (execute function).

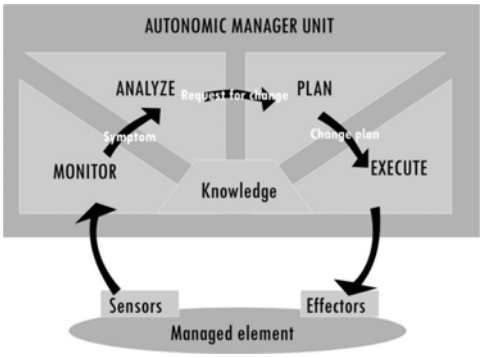


Figure 1. The autonomic agent's MAPE loop

For detailed design and specification of the holonic MES, the Prometheus-Hermes methodology was used (Padgham and Winikoff, 2004; Cheong and Winikoff, 2005, 2006). The design methodology highlights the interaction goal hierarchy, action sequences and a number of failure recovery mechanisms to provide design guidelines for multi-agent systems. An overview of the integrated methodology is shown at Fig. 2 where design artifacts for modeling interactions between agents are highlighted.

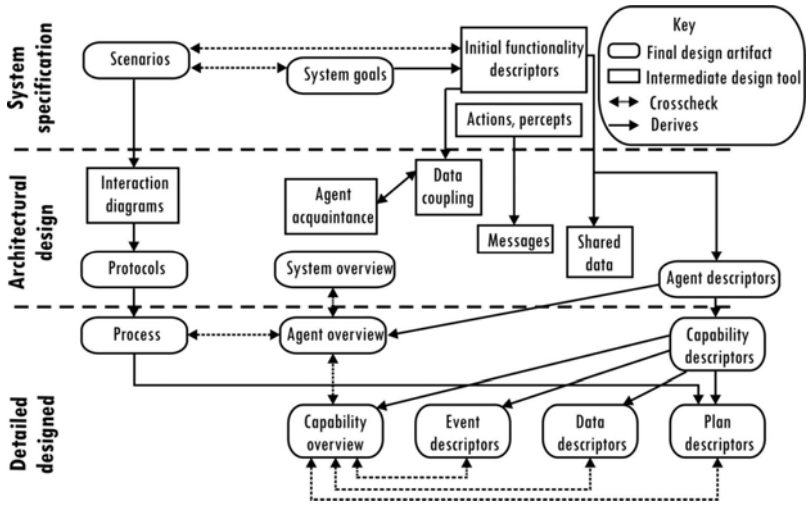


Figure 2. The phases of the Prometheus methodology integrated with Hermes

There are two different roles that can be assigned to an agent in the holonic MES, depending of the object being managed by each of them: *orders* and *resources*. The order agent (OA) is responsible for complete the order as required which determines the operations for the processing of that order. The order agent chooses the processing route to be communicated as the current order process plan (selecting one best-performing solution) and follows it moving the order between the resources. The resource agent (RA) manages the schedule for a resource and registers its usage and failure state at the Gantt chart. It is the responsible for the execution of tasks for different orders. The interaction goal hierarchy for the holonic MES is shown in Fig. 3 in which the undirected lines denote sub-goal relationships and the directed lines depict temporal dependency. The roles involved in the achievement of each goal are shown with the letter R, and it is also identified the initiator of each interaction goal with the letter I.

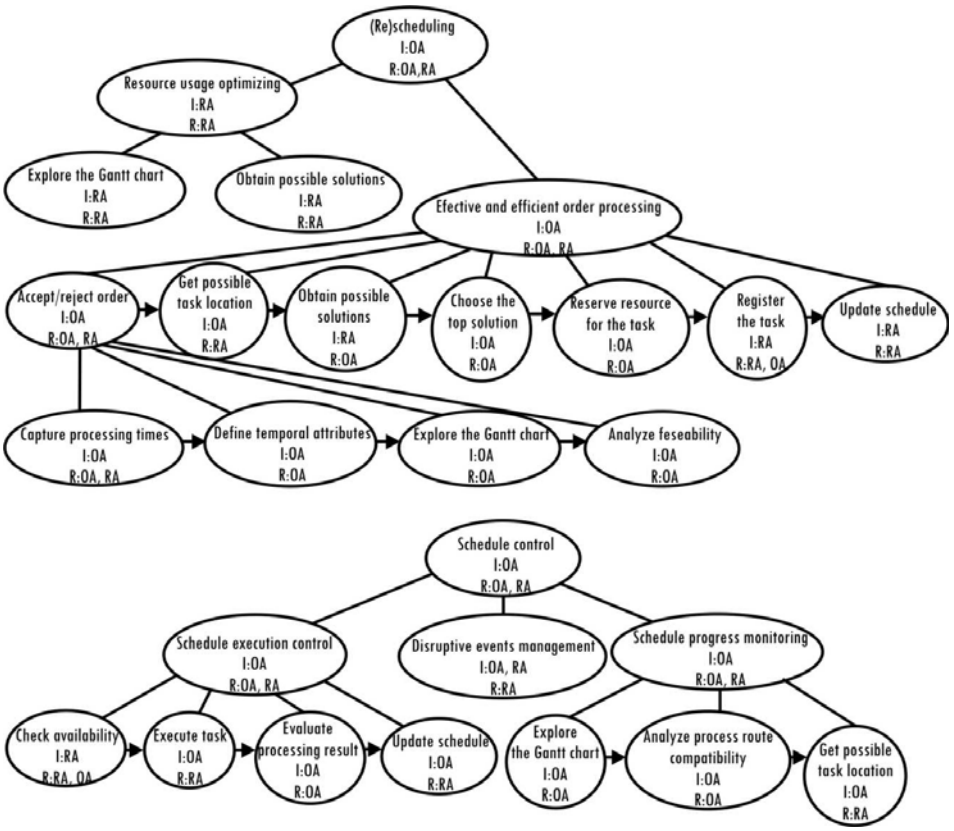


Figure 3. Goal Hierarchy in the holonic MES

When a new order arrives, the corresponding order agent looks up the Gantt in order to discover if the new order is feasible. If so, the order is accepted, and if not, it is rejected with a feasible due date. If it is accepted, the order agent asks candidate resource agents, stage by stage, different options for the probable arriving time at their queue and makes them a request about the probable finalization time at their resource. These options given by resource agents allow making a list of solutions. Given the different solutions provided by resource agents, the order agent selects the top one, with the aim of booking the time slot corresponding to the different tasks in the order. In the current

intention selection, feasible orders can be placed according to different resource agent decision rules such as SPT: shortest processing time, EDD: earliest due date and FIFO: first in first out. As a response to confirmation of resource usage by an order agent, the resource agent registers the task and updates the Gantt chart accordingly.

3. Generative Modeling

3.1. Holonic MES simulation

A computational model was implemented in Netlogo® (Wilensky, 1999), a software environment specifically designed for generative modeling of artificial agent societies. In this system, a feasible schedule for batch process emerges from the combination of individual holon' schedules. Internal coordination mechanisms between the entities in the architecture form the basis for the resultant schedules. In this simulation world, order agents (clients) negotiate with resource agents (servers) the provision of required resources to process an order and the dynamic scheduling emerging as a result of such interactions are perceived in a dynamic Gantt chart (Fig. 4).

Figure 4. A view of the resultant dynamic Gantt chart

3.2. Case Study

To exemplify the proposed approach let's consider the application of the holonic MES in a multiproduct batch plant comprising of 4 stages and 10 units to obtain 5 different products (Fig. 5). Each order has different attributes such product type and due date whereas arrival times, processing times and machine failure rates are stochastic. A batch (order) can follow many different routes through the batch plant while using different pieces of equipment. So there is a great deal of flexibility in order scheduling and it is not obvious how smoothly each order will flow through the batch plant due to a number of disruptive events. It is assumed that order agents have the objective of decreasing late deliveries, so they have decision-making logic needed to choose from the list of solutions the earliest global finalization time alternative. Resource agents resort to the FIFO dispatching rule for request processing at their resources.

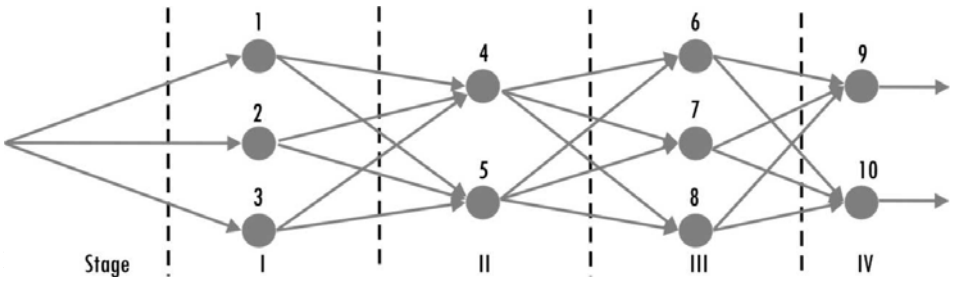


Figure 5. Multi-product plant network structure

3.3. Simulation results

Fig. 6 exhibits the dynamics of the plant for the normal operating scenario. As can be seen the total processing times of both order types shown tend to stabilize roughly after a time equal to 1000 mins. Fig. 7 shows the impact of total processing time for the order types 2 and 3 when orders of type 3 experiments and instantaneous increase in the processing time at stage II.

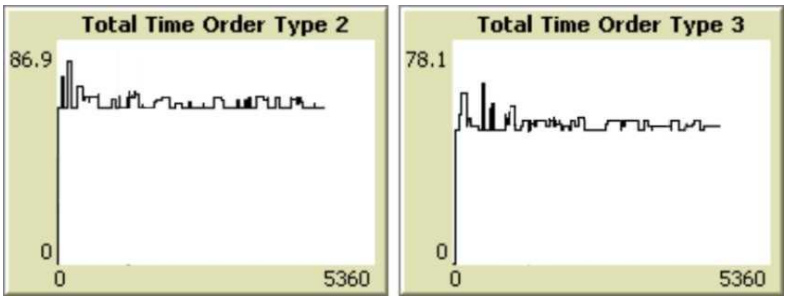


Figure 6. Total processing time for order types 2 and 3 for the normal scenario

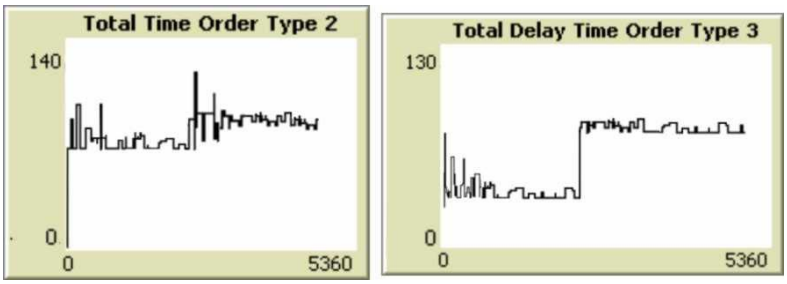


Figure 7. Total processing time for order types 2 and 3 for the scenario when order type 3 becomes more demanding of processing time at stage II since time 2000 min.

Fig. 8 describes the dynamic response of the holonic MES when the resource 1 experiments a sudden increase in its breakdown rate. Order type 3 is somewhat affected whereas type 2 orders are severely disrupted in their processing times even when this order type is not processed at this resource. At Fig. 9, the processing times for the same type of orders are shown, for the scenario where order type 2 suddenly become rush orders. Type 3 orders type are affected by this priority change but then their processing time stabilizes again, yet at a higher average whereas the type 2 orders do not suffer any sudden disruptions and their processing time stabilizes at a lower average.

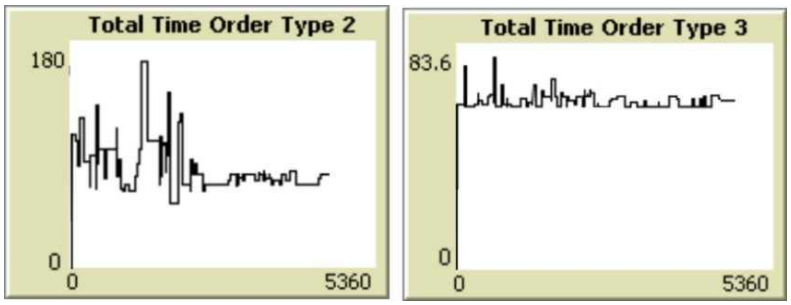


Figure 8. Total processing time for order types 2 and 3 for the scenario where an increase of resource 1 breakdown rates between minute 1000 and 2000 is simulated

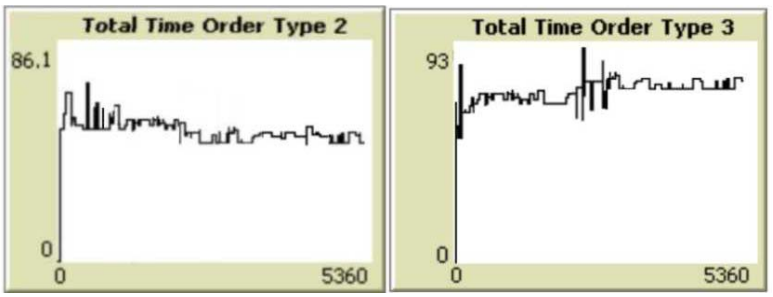


Figure 9. Total processing time for order types 2 and 3 for the scenario where type 2 orders change its priority from normal o rush orders at the 2000 min.

4. Final comments

This paper proposes a novel design of a holonic manufacturing execution system based on well-defined interactions between autonomic agents which manage two different types of objects: orders and resources. The design methodology highlights agent interactions to provide a detailed overview of a holonic MES dynamics. To asses the proposed design a generative simulation model has been implemented. Results obtained for different simulated scenarios indicate that the proposed MES is robust and stable despite the total autonomy given to order and resource agents when negotiating resource usage without resorting to a priory defined schedule.

5. Acknowledgements

The authors thanks to the ANPCyT of Argentina which partially supports the work presented here through project PICT 1099/06.

References

C. Cheong, M. Winikoff, 2005, Hermes: Designing goal-oriented agent interactions, Proceedings of the 6th Ingernational Workshop on Agent-Oriented Software Engeneering.

C. Cheong, M. Winikoff, 2006, Improving flexibility and robustness in agent interactions: Extending Prometheus with Hermes, Software Engineering for Multi-agent systems IV: Research Issues and Practical Applications, 189-206.

Holonic Manufacturing System (HMS) consortium, http:// hms.ifw.uni-hannover.de/, Last access 05/02/2009.

IBM Corporation, 2006, An architectural blueprint for autonomic computing, Available at www-03.ibm.com/autonomic/pdfs, Last access 05/02/2009.

J. Kletti (Ed.), 2007, Manufacturing Execution Systems – MES, Springer-Verlag, Berlín, ISBN 978-3-540-49743-1.

P. Leitao et al., 2005, ADACOR: A collaborative production automation and control architecture, IEEE Intelligent Systems, 20, 58–66.

M. North, M. Macal, 2007, Managing Business Complexity, Discovering Strategic Solution with Agent-Based Modeling and Simulation, Oxford University Press.

L. Padgham, M. Winikoff, 2004, Developing Intelligent Agent Systems: a practical guide, John Wiley & Sons, ISBN 0-470-86120-7.

P. Valckenaers et al., 2007, Schedule execution in autonomic manufacturing execution systems, J. of Manufact. Systems, 26, 75–84.

P. Verstraete et al, 2008, Towards robust and efficient planning execution, Engineering Applications of Artificial Intelligence, 21, 304-314.

U. Wilensky, 1999, Netlogo Modeling Environment, available at http://ccl.northwestern.edu/netlogo, Last access 05/02/2009.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Design of Sustainable Batch Processes Through Simultaneous Minimization of Process Waste, Cleaning Agent and Energy

Iskandar Halim[a] and Rajagopalan Srinivasan[a,b]

[a]*Institute of Chemical and Engineering Sciences, A*STAR (Agency for Science, Technology and Research), 1 Pesek Road, Jurong Island, Singapore 627833*
[b]*Department of Chemical and Biomolecular Engineering, National University of Singapore,10 Kent Ridge Crescent, Singapore 119260*

Abstract

The need for sustainable operation has propelled the batch process industries to consider waste minimization starting from the early stages of process development. However, the time-dependent characteristics of batch operation present challenges in minimizing, recovering and recycling of waste. We present an intelligent system that addresses these by integrating different process systems engineering methodologies – process graph (P-graph), hierarchical design strategy, water reuse method, and multi-objective stochastic optimization. We illustrate the system on a well-known literature case study.

Keywords: sustainable development; multi-objective decision making; process synthesis; simulation-optimization; simulated annealing.

1. Introduction

The drive towards sustainability has prompted the batch process industries to switch from end-of-pipe treatment to waste minimization in tackling the pollution problem. However, unlike in continuous plants, the waste profile in batch processes differs in many aspects. As the same equipment can be used for multiple operation steps, wastes differing in quantity and composition are generated intermittently with the execution of each operation step. This complicates the opportunities for waste minimization, material recycling and energy recovery within the process. Further, most of the existing methodologies for waste analysis are restricted to solving individual waste sources, such as reaction by-products (Chen and Feng, 2005), cleaning agents (Foo *et al*., 2005) and energy (Majozi, 2006) without considering the trade-off between them. This may result in incomplete waste solution at best and an incorrect one at worst. Therefore, a systematic framework to identify the various elements that affect each operation and have a potential role in the waste generation should be developed. In this paper, we introduce such a framework that amalgamates three domains: knowledge-base analysis, process simulation, and mathematical optimization for comprehensive and cost-effective analysis.

2. An Intelligent Simulation-Optimization Framework

Batch-ENVOP*Expert* has been implemented in an object-oriented framework using Gensym's G2 expert system shell. It consists of qualitative and quantitative knowledge domains. A qualitative analysis is first conducted to identify the sources of wastes and to propose heuristic solutions for eliminating or minimizing them. Simulation of process

variables responsible for the waste traits is next performed to measure the efficacy of each solution in terms of economic and environmental impacts. This is followed by multi-objective optimization to resolve the conflicting effects between the objectives.

2.1. Waste Source Heuristic Diagnosis

The first step in the quest for waste minimization in batch processes is to identify each material component present in each unit of the process. We have used Grafcet in combination with unit-based simulation to establish the material flow within the batch plant − this can be represented using P-graph model. Once the P-graph model for the process is developed, materials present in each stream and unit are fully established and then waste sources can be diagnosed. To do this, the P-graph model is decomposed into sub-graphs comprising of operation steps that contribute to the presence of each material component in the waste stream (or unit). This is done by tracing each component in the waste stream upstream through the P-graph model to detect the operations that lead to the escape of materials into waste streams. Once the waste origins are identified, the next stage of the analysis is to propose heuristic design alternatives for each of them. The reader is referred to Halim and Srinivasan (2008) for a list of design heuristics.

2.2. Functional Variables Identification

While the heuristic analysis provides an overview of the various design modifications, the actual benefit of each can only be substantiated through modification at the process variables level. For this, we classify each operation step in term of the phenomena it causes (general function) and the variables affecting the phenomena (specific function). The general function of each operation step is represented by the P-graph model, which captures the physical phenomena in the operation. The specific function of each operation is defined through a subset of variables affecting the phenomena. Once the list of functional variables representing an operation is available, the next step is quantitative assessment using a process simulator to evaluate the economic and environmental impacts arising from the modifications.

2.3. Water Reuse Method

One of the heuristic alternatives that can be proposed is reuse of cleaning waste. For this, we apply water reuse methodology of Mann and Liu (1999) to minimize the cleaning agent requirement in a batch process. Figure 1 shows the basic flow scheme of a water-using operation. A contaminant j with a mass load of $\Delta m_{i,j,tot}$ is to be washed away during an operation i using water as the cleaning agent. The maximum concentrations of contaminant j at the inlet and outlet of this operation are $C^{\max}_{i,j,in}$ and $C^{\max}_{i,j,out}$, respectively. The inlet stream consists of fresh water source f_i (i = 1, 2, …, $n_{operations}$) with zero contaminant and water from the upstream operation k, $X_{i,k}$ with the contaminant loading of $C_{k,j,out}$. The term upstream is used to describe any operation k (k = 1, 2, …, $n_{operations}$ with $k \neq i$) whose finishing time precedes the starting time of the operation i. The outlet stream from operation i consists of flowstream W_i to wastewater treatment and $X_{l,i}$ to downstream operation l (l = 1, 2, …, $n_{operations}$ with $l \neq i$).

The inlet concentration of contaminant j to operation i, $C_{i,j,in}$, is defined as a constraint with the following bound:

$$C_{i,j,in} = \frac{\sum_{k \neq i} X_{i,k} C_{k,j,out}}{\sum_{k \neq i} X_{i,k} + f_i} \quad (1)$$

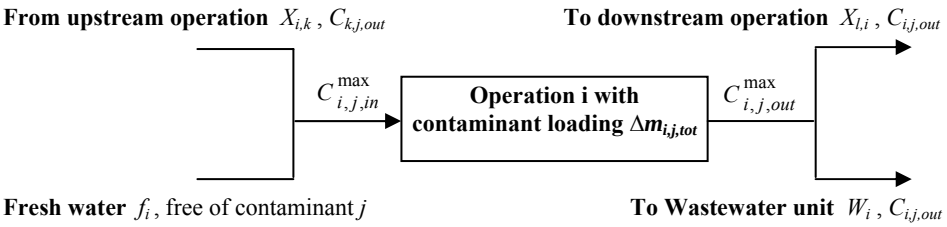


Figure 1. Schematic of water-generating operation

Likewise, the outlet concentration of contaminant j from operation i, $C_{i,j,out}$, can be described as follows:

$$C_{i,j,out} = C_{i,j,in} + \frac{\Delta m_{i,j,tot}}{\sum_{k \neq i} X_{i,k} + f_i} \tag{2}$$

Setting $C_{i,j,out}$ to be equal the maximum allowable oulet concentration gives the following equation:

$$C_{i,j,out} = \frac{\sum_{k \neq i} X_{i,k} C_{k,j,out} + \Delta m_{i,j,tot}}{\sum_{k \neq i} X_{i,k} + f_i} = C^{\max}_{i,j,out} \tag{3}$$

The following two equations can be obtained through rearranging (1) and (3):

$$\sum_{k \neq i} \left(C^{\max}_{i,j,in} - C_{k,j,out} \right) X_{i,k} + C^{\max}_{i,j,in} f_i \geq 0 \tag{4}$$

$$\sum_{k \neq i} \left(C^{\max}_{i,j,out} - C_{k,j,out} \right) X_{i,k} + C^{\max}_{i,j,out} f_i = \Delta m_{i,j,tot} \tag{5}$$

An overall mass balance around each unit operation can be written as:

$$f_i + \sum_{k \neq i} X_{i,k} - W_i - \sum_{l \neq i} X_{l,i} = 0 \tag{6}$$

The above equations form a set of constraints over which the total fresh water flow can be minimized. For a process involving multiple contaminants, an iterative optimization procedure is to be solved for each of the contaminants to obtain minimum fresh water flow that satisfies all the process constraints.

2.4. Multi-Objective Optimization Simulated Annealing-Jumping Genes

The combined outputs from the process simulation and water reuse method yield the complete set of sustainability measures for the design. They also serve as the objective function for a stochastic optimization. For this, we use the simulated annealing − jumping genes algorithm introduced by Sankararao and Gupta (2007) together with the multi-objective simulated annealing of Suppapitnarm et al. (2000) as follows:

(1) Start by specifying the solution vector x, whose elements are the decision variables of the process. To initialize, the variable values from the base-case design x' are used as the solution vector to the result of $[x', f_{env}(x'), f_{econ}(x')]$. Here, f_{env} and f_{econ} are the environmental and economic impact, respectively.

(2) Specify an initial annealing temperature T. Initially T is set to a large number $T_{initial}$ and then gradually reduced. Specify also the initial size of step change $ST_{initial}$.

(3) For N number of decision variables, a constant parameter f_b is first determined with $f_b < N$. A random number is next generated to identify the location of a decision variable p as one end of the jumping genes that lies between 0 and $N - f_b$. The other end of the jumping genes is set as $p + f_b$.

(4) Perform random perturbation to generate a new solution vector x'' in the neighborhood of x', $x'' = x' + \mathrm{R}.ST_{initial}$ (7)
where R is a random number between -1 and 1.

(5) Subsequent to the perturbation of the decision variables, jumping genes operation is next performed: $x_i''' = x_i'' + r\ (x_{i,\ max} - x_{i,\ min})$ (8)
where $i = p$ to $p + f_b$ is the length of jumping genes, r is the random number between 0 and 1 and $x_{i,\ min}$ and $x_{i,\ max}$ are the minimum and maximum values for the respective variables. Using the candidate solution vector, simulate the process to obtain new objective values $f_{env}(x''')$ and $f_{econ}(x''')$.
(6) Compare the objective values of x''' with all solutions in the Pareto set. If x''' dominates any element of the Pareto set, replace that element with x'''. If x''' is Pareto-optimal with all the elements in the set, then include it in the set.
(7) If x''' is dominated, then accept it as the current solution vector with a probability $P = \min\{1, \exp[-(f_{env}(x''') - f_{env}(x'))/T).\exp(-(f_{econ}(x''') - f_{econ}(x'))/T)]\}$ (9)
At the same time, a random number P_{rand} is generated between 0 and 1.
(8) x''' is accepted as the new solution only when $P > P_{rand}$; otherwise the earlier solution vector x' is retained.
(9) Periodically the annealing temperature T is reduced using a reduction factor R_T. The step change size is reduced also.
(10) Periodically, x' is replaced with a selected solution from the Pareto set.
(11) Repeat steps (3) to (10) for a predefined number of iterations N_{iter}.

3. Case Study: Batch Production Plant

We have tested the framework on a literature case study (von Watzdorf *et al*., 1994). Figure 2 shows the flowsheet of the process which involves reaction between chemical A and B in a cooled-jacketed vessel resulting in formation of product C and byproduct D. After the reaction is performed, the reaction mixtures are transferred to an intermediate storage tank and the reactor vessel washed with cleaning liquid. After three reaction cycles, the mixture in the storage is batch distilled to produce two main cuts of high purity A and C. The final residue from distillation is D, which is sent to a separate storage. The process requires substantial amounts of fuel due to heating operations − steam is needed in the reboiler and heater HE101. Our objective here is to minimize the environmental impact from three waste sources: process waste (T111), cleaning waste (T106) comprising mainly component A, C, and D and fuel emission due to heating process without adverse economic impact from the changes. The process currently generates an environmental impact of 1691 calculated using the WAR algorithm software (Martin and Young, 2008). The impact of the fuel emission is calculated as the sum of the total heating energy (MJ) multiplied by an equivalent environmental impact factor of 0.0043. The limiting contaminant concentration data for the three washing process is shown in Table 1. The economic objective as represented by the operating cost – including costs of waste treatment (process and cleaning waste), cleaning agent and heating and cooling energy – is $6084/hour. This is calculated using the cost data of Table 2.

Table 3 shows some of the proposed qualitative solutions for this process. Based on the solutions proposed, cooler set point, reboiler set points at different stages of operation, and purity specifications of A and C are set as the decision variables as they directly impact the waste generation. Figure 3 shows the Pareto optimal set that is obtained by manipulating the functional variables. Compared to the base-case design, the maximum reduction in the environmental impact and operating cost is found to be 1.9% and 2.8%, respectively. Figure 4 shows the proposed recycle network of cleaning agent. The savings in the amount of cleaning agent comes from recycling of cleaning effluent from the first wash operation to the second and from the second washing to the third.

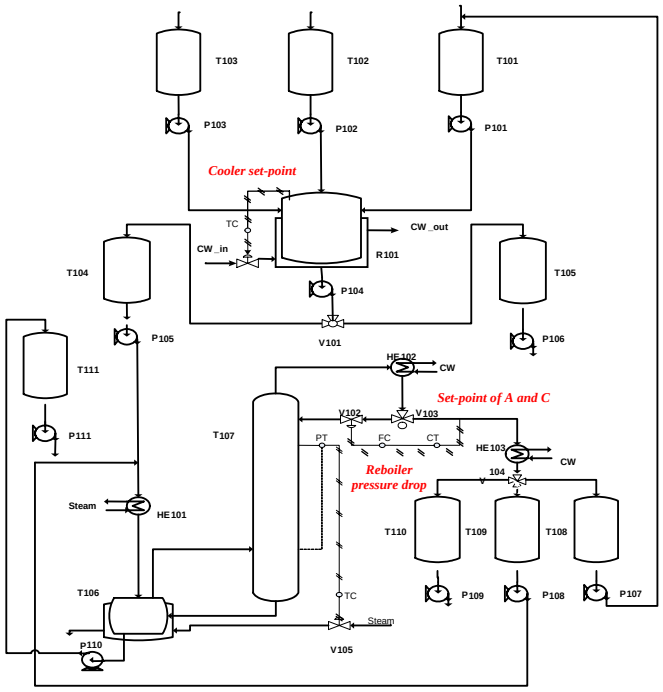


Figure 2. Batch process flowsheet

Table 1. Limiting concentration for washing process

Operation i	Contaminant j	$C^{max}_{i,j,in}$	$C^{max}_{i,j,out}$
1st wash	A	5500	7500
	C	20000	27786
	D	11000	15667
2nd wash	A	6000	8000
	C	21000	28714
	D	12000	16571
3rd wash	A	7000	9500
	C	26000	35643
	D	15200	20914

Table 2. Economic data

Unit	Cost
Washing agent ($/kmol)	2
Hot Utilities ($/MJ)	0.07
Cold Utilities ($/MJ)	1.0
Process waste ($/kmol)	5
Washing waste ($/kmol)	5

4. Conclusions

We propose an intelligent simulation-optimization framework for identifying sustainable alternatives for batch processes. We account for waste generated by the reaction-separation operation, cleaning waste and fuel emission. The knowledge-based approach is able to diagnose the waste sources and identify heuristic design alternatives. Specific variable-level manipulation is evaluated using process simulator to minimize the amounts of process wastes as well as energy. Water reuse model is applied to reduce the amount of cleaning agent through optimal design of recycle network structure. Optimal variable manipulation is next identified using a jumping gene based simulated annealing technique to evaluate both the environmental and economic objectives. The system has been successfully tested with convincing results.

Table 3. Waste minimization alternatives

Operation Step	Solution
Reaction (R101)	Optimize the reaction condition to eliminate byproduct D and fully convert material A
Washing (T105)	Recycle the liquid waste to the next cleaning process
Distillation (T106)	Use further separation after the vaporization process to avoid C from becoming waste

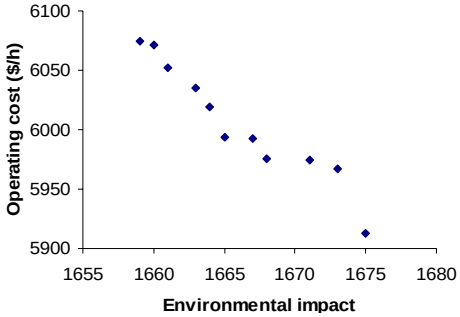


Figure 3. Pareto optimization

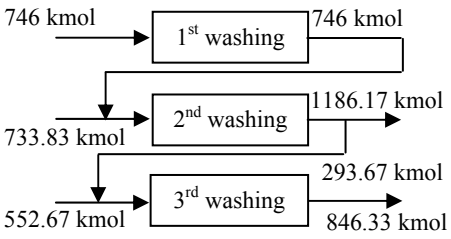


Figure 4. Proposed recycle network

References

Q. Chen and X. Feng, 2005, A pollution reduction methodology in reactor design, *Chemical Engineering and Processing*, 44, 13.

C.Y. Foo, Z.A. Manan, R.M. Yunus and R.A. Aziz, 2005, Synthesis of maximum water recovery network for batch process systems, *Journal of Cleaner Production*, 13, 1381.

T. Majozi, 2006, Heat Integration of Multipurpose Batch Plants Using a Continuous-Time Framework, *Applied Thermal Engineering*, 26, 1369.

I. Halim and R. Srinivasan, 2008, Designing sustainable alternatives for batch operations using an intelligent simulation-optimization framework, *Chemical Engineering Research and Design*, 86, 809.

J.G. Mann and Y.A. Liu, 1999, Wastewater Minimization through Mathematical Optimization, *Industrial Water Reuse and Wastewater Minimization*. McGraw-Hill. New York.

B. Sankararao and S.K. Gupta, 2007, Multi-objective optimization of an industrial fluidized-bed catalytic cracking unit (FCCU) using two jumping gene adaptations of simulated annealing, *Computers and Chemical Engineering*, 31, 1496.

R. von Watzdorf, U.G. Naef, P.I. Barton and C.C. Pantelides, 1994, Deterministic and stochastic simulation of batch/semicontinuous processes, *Computers and Chemical Engineering*, 18(supplement), 343.

A. Suppapitnarm, K.A. Seffen, G.T. Parks and P.J. Clarkson, 2000, A simulated annealing algorithm for multiobjective optimization, *Engineering Optimization*, 33, 59.

T. Martin and D. Young, 2008, Waste Reduction Algorithm Graphical User Interface, US Environmental Protection Agency. Cincinnati.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Plant Design Project Automation Using an Automatic Pipe Routing Routine

Ewerton E. S. Calixto,[a] Paula G. Bordeira,[a] Hugo T. Calazans,[a] Cesar A. C. Tavares,[a] Marco T. D. Rodriguez[a]

[a]*Chemtech Serviços de Engenharia e Software LTDA, Rua da Quitanda, 50, andar 21, Rio de Janeiro – RJ, CEP 20011-030, Brazil*

Abstract

The front-end engineering and design (FEED) phase of an industrial plant project aims to determine the feasibility of the enterprise, and to produce accurate cost estimates. Successful FEED studies guarantee that provisional decisions on investments can be made more quickly by the owner or investor. Thus, the stakeholders are aligned with the project purposes, minimizing the risks involved for the acquisition of materials. The routing of pipes is a crucial phase in the plant design field, representing approximately 20% of total project costs for the industry as a whole. This work discusses the state-of-the-art concerns for plant design projects, and presents an effective software product for automating the pipe routing process. The results presented herein indicate that the developed solution is able to reduce man-hours and costs involved in the FEED phase of a project by providing, in quantitative terms, a satisfactory pipe routing.

Keywords: Automatic Pipe Routing, Front-End Engineering and Design, Automation, Plant Design, Materials Take Off.

1. Introduction

Specific literature on project management indicates that approximately 70% of all projects executed during the last five years were not finished on time or within budget (Saputelli et al., 2008). Therefore, the optimization of time, cost, materials and human resources are important aspects to be taken into consideration during the project's life-cycle.

A front-end engineering and design (FEED) study is a Process Engineering approach applied during the early planning stages of an industrial plant project. The strategy of this type of study consists of two main activities: (1) determining the feasibility of the enterprise, and (2) producing more accurate cost estimates. During the FEED stage of a project, about 80% of the total development costs are determined (Peters and Timmerhaus, 1991). By separating the two activities mentioned above and performing them earlier in the project, it is possible to take better advantage of scope definitions, reducing project risks and costs.

One important concept for understanding the FEED study activities is that of materials take-off (MTO). MTO is the list of all piping material to be used for the execution of a certain project. Successful FEED studies guarantee that provisional decisions on investments for the construction of an industrial plant can be made more quickly by the owner or investor. Thus, the stakeholders are aligned with the project objectives, minimizing the risks involved for the acquisition of materials. In this context, it is important that MTOs provide reliable estimates (Burdorf et al., 2004).

Due to the complexity and the time-intensive nature of the evaluation, the modeling of pipework items represents a significant portion of the overall costs of industrial plant projects. The main goal of the study presented here was to develop an integrated system which could reduce design time and maintain the accuracy of the quantitative analysis during the FEED pipe routing activities of plant design projects.
The designing principle adopted for the design and development stages of the proposed system aimed at the possibility of creating a software environment that could communicate with third-party, commercial computer-aided design (CAD) software products, such as PDMS[1] (Plant Design Management System).
This paper proposes a rule-driven pipe routing tool for decreasing the time needed for the execution of FEED projects. The results presented herein indicate that the developed system is able to reduce man-hours and costs involved in the FEED phase of a project by providing, in quantitative terms, a satisfactory pipe routing.

2. Pipework for Industrial Plant Projects

The importance of pipework in the industrial plant field is enormous. All industries have pipe networks which are responsible for conveying different types of fluids all over the industrial plant layout. Each part of the pipe network has a specific role, and based on this role, is essential for the industry to run well (Telles, 2001).
In the process industry context, the pipe networks have a great influence in the overall project, connecting equipment, such as pressure vessels, reactors, tanks, pumps and heat exchangers. The industrial plant layout study analyzes the most economical spatial configuration for the industrial plant, meeting constraints regarding construction, safety, operation and maintenance factors (Mecklenburgh, 1985).
According to Telles (2001), the pipework value for the chemical processing industry represents between 20% and 25% of the total industrial plant construction costs. Additionally, pipework assembly reaches, on average, between 45% and 50% of the equipment layout expenses, while the pipework design is worth approximately 20% of the project industry as a whole. Therefore, the pipework design and other correlated activities are often considered the most challenging efforts during a plant design project (Kentish, 1982). Piping design must take into consideration not only economics and design simplicity, but also flexibility, support requirements, and several other mechanical aspects (Guirardello and Swaney, 2005).
The pipework engineering activities depend on several rules and principles which should be met during the design, construction and maintenance phases of the project. Piping projects involve an intricate network of components including pipes, flanges, supports, gaskets, bolts, valves, strainers and expansion joints. These elements are available in a variety of materials, types and sizes, and are manufactured under national, international or proprietary standards. Other aspects such as safety needs should also be considered during the pipework design cycle. Pipe networks containing hazardous fluids, for instance, must be protected from damage by external mechanical impacts, leaks from adjacent pipes and external sources of heat. In addition to these considerations, the designing principles must also provide the unit plant employees with safe access conditions for performing operation, inspection and maintenance tasks, as well as guarantee that the pipe supports and bridges are designed with sufficient traffic impact resistance and mechanical strength to support the loads exerted on them (Holmes, 1973).

[1] AVEVA PDMS is a three-dimensional computer-aided design (3-D CAD) software package for engineering, design and construction projects distributed by AVEVA Solutions Ltd.

The front-end loading (FEL) methodology is an engineering project approach used in order to ensure that exhaustive alternative capital analyses will be performed during the life cycle of industrial plant projects. FEL helps oil and gas operators integrating business and technical goals to reduce project costs, improve operational efficiency and maximize profitability (Saputelli et al., 2008). The FEL Definition phase of a project corresponds to the FEED engineering stage. This stage involves basic engineering, operations plan, risk plan, contracting and sanctioning activities. The FEL's fundamental thesis is that the anticipation of the plant construction phase can effectively advance the return on investment (ROI) of a given industrial project.

3. Automatic Pipe Routing in Computer-Aided Design

In order to decrease the amount of effort required for performing process and plant design activities, a great variety of computer-aided design (CAD) software products has been proposed. One of these products is AVEVA PDMS, which provides PDMS Router, a built-in tool for the automatic routing of pipes. The environment offered by PDMS routes the pipes orthogonally, using graph's theory principles (Guibbons, 1985) and a shortest-path algorithm (Wang and Crowcroft, 1992) to minimize the amount of piping material needed for the routing.

The purpose of the PDMS Router is to minimize material costs while avoiding clashes with other plant elements. The plant designer is able to choose the pipes intended for the routing process, and to create constraints based on spatial planes and points. The routing algorithm used by PDMS Router has three different operational modes, which are called Level 1, Level 2 and Level 3. At first, PDMS Router searches for a route using the Level 1 mode. If no clash-free route is found after the first operational mode, a search is made using Level 2. Similarly, if no free route is found after Level 2, then Level 3 is used.

As shown in Figure 1, the Level 1 operational mode searches for an orthogonal route between the head point (PH) and tail point (PT) of a pipe, using the minimum number of bends and elbows.

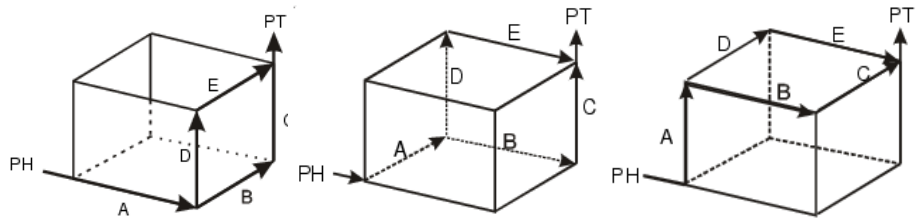


Figure 1. Level 1 mode (AVEVA Solutions Ltd., 2007).

If all Level 1 routes are blocked, PDMS Router will attempt Level 2 routes. In Level 2 mode, PDMS Router will attempt to withdraw the route into the box by a distance that could enable the pipe to bypass the obstruction. Then, PDMS Router tries the same routing patterns as those used in Level 1 mode. An example of a Level 2 route is shown in Figure 2 (a).

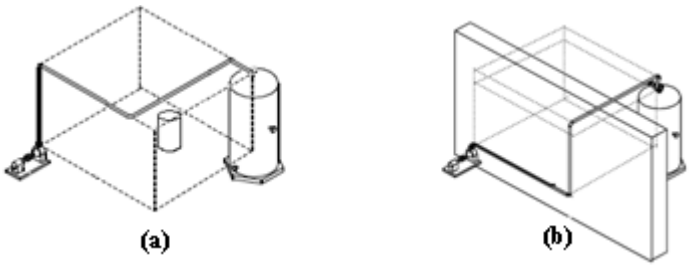


Figure 2. Level 2 (a) and Level 3 (b) (AVEVA Solutions Ltd., 2007).

If PDMS Router cannot find a clash-free route using first and second level routes, it will attempt to find a Level 3 route, as shown in Figure 2 (b). In Level 3 mode, PDMS Router extends the box outwards until it bypasses the obstruction and then attempts to route the pipe using Level 1 routing principles.

It is important to keep in mind that PDMS Router does not provide any other automation beyond the de facto routing of pipes. This means that the designer has to perform all the manual tasks needed for PDMS Router to work, such as creating the pipe hierarchy elements and determining their material specification, as well as creating and connecting the head and tail points of all their branches. The routing features of PDMS Router can only be used after all the previously mentioned tasks are completed.

4. Automatic Pipe Routing Routine

In order to reduce the human involvement in the tasks necessary for the PDMS Router to work, a software architecture environment was implemented in the PML[2] programmable language. The proposed environment should be able to perform the steps shown in Figure 3 and described further down.

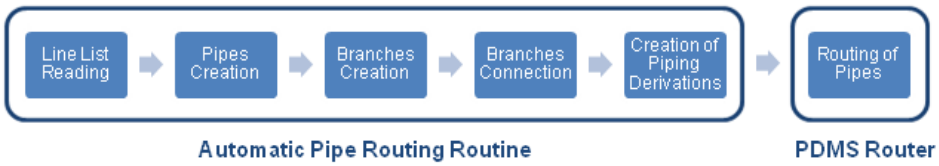


Figure 3. Automated workflow process for the automatic routing of piping elements.

The proposed workflow process consists of a chain of six activities. The first five ones are logically grouped in the implemented Automatic Pipe Routing Routine. Each one of the mentioned activities is responsible for automating a different step necessary for the effective PDMS Router processing.

Initially, the routine reads the information presented in the line list. Using this information, the routine creates the hierarchy elements of the pipes meeting the material specification constraints presented in the list. After this step, the branches are created and then connected to their connection points. By reading the line list, the routine evaluates the number of piping derivations necessary to be created for each pipe. These derivations are created and then, as the last step for the automated workflow process, the pipes are routed with PDMS Router.

[2] PML (Programmable Macro Language) is an object-oriented programming language for interacting with AVEVA PDMS.

The main advantage of using this routine in a FEED project is the possibility of performing an accurate evaluation of piping materials with minimum effort. The resulting estimates may be given in terms of relevant MTO data, such as the total evaluated piping weight and length.

4.1. A practical case study

Some evaluations were performed in a specific refinery plant unit (Figure 4) called the atmospheric distillation unit (ADU) during the FEED phase of an energy company project. The evaluations basically consisted of a comparison between the total pipe length obtained by the manual piping design and that obtained by the automatic pipe routing routine proposed herein.

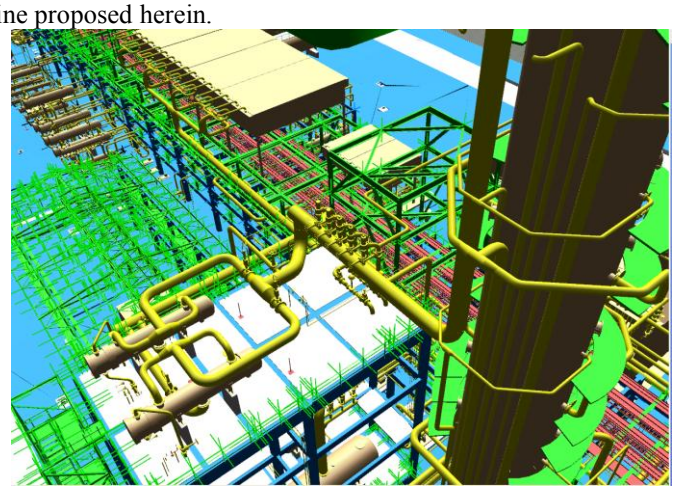


Figure 4. Set of pipes routed in the atmospheric distillation unit of the refining project used as a case study for this work.

A set of fifteen pipes described in the industrial plant engineering flowsheet diagram was chosen. The purpose was to route automatically the selected pipes and compare the results with the routing provided by the company piping design team.

As shown in Table 1, the comparison results indicate an average error of 6.5% between the total piping length obtained by the manual design and by the automatic pipe routing process. According to the literature (Telles, 2001), such an error represents a comfortable and reliable margin for a successful MTO.

Table 1. Pipe Length Comparison Results.

Total Length (With Routine) (m)	334.6
Total Length (Without Routine) (m)	312.86
Error (%)	6.496

This result gives a good margin in MTO reports and the piping materials can be purchased with more reliability.

5. Future Work

For future work, it is important to minimize the layout and measuring differences between the manual and the automatic routing of a single pipe. The impact magnitude of such differences seems to be attenuated, or at least stabilized, when a large set of pipes is being routed in a complex and dense plant unit. The current version of the

proposed integrated environment only evaluates total pipe routing lengths. However, future implementations should provide weight estimates as well, since weight is a very important parameter for MTO reports.
Another working possibility would be to improve the automatic pipe routing routine layout with respect to qualitative terms. For example, in this context, design factors associated with pipe flexibility and stress analysis could be taken into account for the layout design. Such layout considerations could be used as an appropriate strategy for increasing the accuracy of the automatic pipe routing MTO estimates.

6. Conclusions

The use of the automatic pipe routing routine presented herein can reduce the time required for the production of reliable MTO estimates at the end of FEED studies for chemical engineering projects. The results indicate that the implemented solution is able to automate manual and time-consuming stages previously required for the three-dimensional piping design process. The possible benefits provided by the routine should impact all the project stakeholders, ranging from designers to investors. Some improvements should still be considered, however, regarding increasing the accuracy given by the routine's MTO estimates.

References

AVEVA Solutions Ltd, Pipework Design User Guide, 2007.

Bondy, J.A., Murty, U.S.R., 1982, Graph Theory with Applications, New York: Elsevier Science Publishing Co., Inc.

Burdorf, A., Kampczyk B., Lederhose M., Schmidt-Traub H., 2004, CAPD – Computer Aided Plant Design, Computers and Chemical Engineering, 28, 73-81.

Guibbons, A., 1985 Algorithmic Graph Theory, Cambridge University Press.

Guirardello, R., Swaney, R.E., 2005, Optimization of Process Plant Layout with Pipe Routing, Computers and Chemical Engineering, 30, 99-114.

Holmes, E., 1973, Handbook of industrial pipework engineering, first edition McGraw-Hill.

Kentish, D., 1982, Industrial Pipework, Mcgraw Hill.

Matsui, Y., Takagi, H., Emori, S., Masuda, N., Sasabe, S., Yoshimura, C., Shirai, T., Nioh, S., Kinno, B., 1979, Automatic Pipe Routing and Material Take-Off System, Proceedings of the 16th Conference on Design automation, 121-127, San Diego, CA, United States.

Mecklenburgh, J.C., 1985, Process Plant Layout. New York: Halsted Press; John Wiley & Sons.

Park, J.H., Storch, R.L., 2002, Pipe-routing Algorithm Development: Case Study of a Ship Engine Room Design, Expert Systems with Applications, 23, 299-309.

Peters, M.S., Timmerhaus, K.D., 1991, Plant design and economics for chemical engineers, (4th ed.) New York: McGraw-Hill, Inc.

Póvoa, A.P.B., Mateus, R., Novais, A.Q., 2002, Optimal Design and Layout of Industrial Facilities: An Application to Multipurpose Batch Plants, Ind. Eng. Chem. Res, 41, 3610-3620.

Telles, P.C.S., 2001, Tubulações Industriais: Materiais, Projeto, Montagem, 10ª Edição, LTC.

Saputelli, L. Hull, R. Alfonzo, A., Front End Loading Provides Foundation for Smarter Project Execution, Oil&Gas Financial Journal, Special Report Assessing & Managing Risk, PennWell Corporation.

Wang, Z., Crowcroft, J., 1992, Analysis of Shortest-Path Routing Algorithms in a Dynamic Network Environment, ACM SIGCOMM Computer Communication Review, 22, 2, 63-71.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A continuous targeting approach for integrated solvent and process design based on molecular thermodynamic models

André Bardow, Klaas Steur, Joachim Gross

Process & Energy Department, Delft University of Technology, Leeghwaterstraat 44, 2628 CA Delft, The Netherlands

Abstract

Solvent properties and process behavior are highly interdependent. The integrated design of solvents and processes in a single optimization problem is, however, still prohibitive mainly due to the discrete nature of solvent selection. In this work, a framework for integrated solvent and process design is introduced using a continuous-molecular-targeting approach (CoMT-CAMD). Based on a detailed molecular-thermodynamic model, a hypothetical solvent is introduced in the process optimization problem. The properties of the hypothetical solvent are relaxed and continuously optimized along the process variables thus achieving the integrated design. In a second step, the hypothetical solvent is mapped onto real components. The framework is implemented using the PCP-SAFT equation-of-state and exemplified for the design of solvents for CO_2 capture.

Keywords: computer-aided molecular design (CAMD), conceptual process design, solvent design, CO_2 capture.

1. Introduction

Solvent design is one hallmark example of product design 0. Solvent-based separations are employed in almost all chemical, biochemical and pharmaceutical processes. The selection of a suitable solvent is thereby the determining step for the feasibility and the efficiency of a separation. Selectivity for specific compounds is the primary requirement for a solvent but properties such as capacity, viscosity or toxicity also have to be considered. A final choice thus requires a trade-off between these properties 0.

In practice, solvent selection is often based on experience and simple heuristics leading to suboptimal solutions. Computer-aided molecular design methods (CAMD) provide a systematic means to identify suitable solvents 0. CAMD can be interpreted as the inverse of property prediction: Given a set of target properties, a molecule is sought that satisfies the property specifications.

However, the definition of a representative set of solvent target properties is by itself an intricate problem due to the strong interaction between solvent properties and the process at hand. The ultimate goal is therefore the integrated design of the solvent molecule and the corresponding process 0.

Due to the discrete selection between molecules, the solution of this integrated design problem in a single mathematical optimization is prohibitive today. In 0, a reverse engineering approach has therefore been followed: First, constraints on property values are obtained from process optimization. Then, components are generated that meet these constraints. While this approach allows for the derivation of process-specific targets, a direct link between process performance and molecular characteristics is not achieved.

Recently, advanced thermodynamic models have been developed in the SAFT-family of equations-of-state 0. While most CAMD-approaches employ thermodynamic models applicable at low pressures only, the integration of an equation-of-state (EoS) allows for a sound description of high-pressure systems 0. More importantly, these models are based on a coarse-grained but still very detailed molecular picture. In particular, a sound integration of multipole interactions has been achieved in the PCP-SAFT EoS 000.
In this work, the advanced properties of molecular thermodynamic models such as PCP-SAFT are exploited to integrate solvent and process design. By relaxing the solvent pure component parameters as free variables during process optimization, a hypothetical target molecule is obtained in a continuous optimization problem. This continuous-molecular-targeting approach to computer-aided molecular design (CoMT-CAMD) thus circumvents the discrete selection in the integrated design problem. Only in a second step, the hypothetical target molecule is mapped onto real solvent molecules.
The proposed CoMT-CAMD approach is presented in the following section. A case study for carbon dioxide capture is discussed in Section 3 before conclusions are given.

2. Continuous-molecular-targeting approach

2.1. Proposed framework for integrated solvent and process design
Thermodynamic tools are usually developed to predict phase behavior in process applications. Given the structure of a molecule, thermodynamic properties are predicted to describe the expected process behavior (cf. upper part of Figure 1).

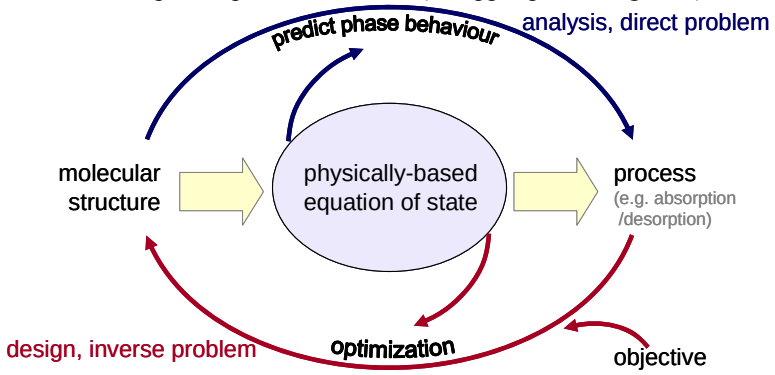


Figure 1: Integrated solvent and process design using a physically-based equation-of-state.

In the proposed design framework, the inverse problem is regarded: Given a certain objective, both the process and the molecular structure are optimized to meet the constraints and reach the optimal process performance (cf. lower part of Figure 1).
In our view, a physically-based equation-of-state is central to truly link molecular characteristics to process performance as indicated in Figure 1. In a physically-based equation-of-state, a molecule is represented by a set of pure component parameters which correspond to clearly defined molecular attributes. In contrast to earlier approaches, these parameters do not lump many effects into one coefficient. Thereby, a close correspondence between a parameter set and specific molecules is achieved. This correspondence is the key factor in the suggested approach.
The integrated solvent and process design problem is approached in two steps:

- Step one: Continuous-molecular targeting
  Both process variables and solvent parameters are simultaneously optimized in a continuous fashion by relaxing the solvent pure component parameters from their

discrete values. This optimization yields settings for optimal process performance and a hypothetical target molecule.

- Step two: Structure mapping
  The hypothetical target molecule characterized by a set of equation-of-state parameters is mapped onto real substances. Only here, discrete decisions have to be taken.

While the approach still consists of two parts, molecular attributes are now directly assessed in process design. The two steps are briefly described in the following.

2.2. Step one: Continuous-molecular targeting

Fully integrated solvent and process design is *de facto* prevented due to the discrete nature of molecule selection. By employing a physically-based equation of state, each candidate solvent is described by a discrete set of pure component parameters. In the continuous-molecular-targeting step, this discrete set is relaxed and a continuous search in the full parameter space is performed. The resulting optimization problem is then

$$\begin{aligned} &\max_{x,y} f(x) \\ &s.t. \quad h(x, g(x,y)) = 0 \\ &\qquad g(x,y) = 0 \\ &\qquad \mathrm{c}(x,y) \le 0 \end{aligned}, \tag{1.1}$$

where x denotes the process variables and y the pure component solvent parameters. The latter typically do not appear directly in the design objective $f(x)$. The process model is given by the equality constraints $h(x,g)$ where the thermodynamic quantities $g(x,y)$ enter. Process and solvent inequality constraints are summarized in $c(x,y)$.

For the integrated design problem (1.1), a standard NLP solver may be employed. Thereby, optimal process settings and pure component parameters for a hypothetical solvent are obtained. The resulting objective gives an upper bound for the process performance achievable by any molecule describable within the chosen thermodynamic framework. This information by itself is valuable in assessing the process design.

To perform the CoMT-step, three elements are required: the thermodynamic model, a process model (incl. constraints) and an objective function. These are discussed next.

2.2.1. Thermodynamic model: PCP-SAFT equation-of-state

As central physically-based equation-of-state, PCP-SAFT is employed 0000. PCP-SAFT has been shown to allow for good predictions of phase behavior in complex mixtures. In PCP-SAFT, the individual molecular contributions to the residual Helmholtz energy are resolved (see 0000). In the present context, it is important that each molecule is represented by a set of seven pure component parameters in PCP-SAFT. These seven parameters correspond to clearly defined molecular attributes:

- two geometric parameters describe size and elongation;
- three energy-related parameters describe dispersion, association energy and volume;
- two polar parameters specifying the point dipole and quadrupolar moments.

The idea of a physically-based model shows from the fact that, e.g., the polar interactions can be obtained from independent sources such as quantum mechanics 00. As an equation-of-state, PCP-SAFT provides a consistent description of all residual thermodynamic properties of a molecule. Given the seven parameters, the behavior in any process application is thus defined. Separate – and probably inconsistent – physical property models describing e.g. heat capacities and saturation pressure are not required. In the present work, PCP-SAFT provides residual enthalpies and equilibrium constants. For the ideal-gas heat capacities, a simple correlation for organic solvents is introduced.

2.2.2. Process model

The proposed method is very flexible with respect to the process model employed. Typically, process equations consist of Material balances, Equilibrium balances, Summation equations and Energy (Heat) balances (MESH). In addition, constraints on process variables (e.g., pressure limitation for safety) may be added.

The process model obtains thermodynamic input from PCP-SAFT.

2.2.3. Objective function

While classical solvent design methods had to focus on solvent properties, process-wide objective functions can be incorporated in the presented CoMT-CAMD-framework. The trade-offs between different solvent properties are thus directly addressed through their implications on process behavior.

Process-wide objective could reflect aspects of cost-, energy- or second-law-efficiency. Different objectives can be employed to study the implications on solvent and process.

2.3. Step two: Structure mapping

The continuous-molecular-targeting step yields optimal PCP-SAFT-parameters y^* of a hypothetical solvent. To implement the findings in practice, these parameters are mapped onto real substances. Thereby, discrete decisions are taken.

Group-contribution methods could provide a suitable tool to implement this mapping. For advanced SAFT-models, these methods are currently being developed 00.

Here, the structure mapping is performed by a database search. In the PCP-SAFT database, potential solvents are represented by their parameter vector y. To assess their expected performance, a Taylor approximation is performed around the CoMT-optimum y^* to predict the expected loss in the objective function $f(x)$ for all database entries

$$\Delta TA(y) = J(y^*)(y - y^*) + \frac{1}{2}(y - y^*)^T H(y^*)(y - y^*), \tag{1.2}$$

where J and H denote the Jacobian and the Hessian of the problem (1.1) available from the optimization in step one. The Taylor approximation criterion allows for a sound assessment of the database entries as it accounts for the different sensitivities with respect to different parameters. Based on the predicted performance, a ranked list of real molecules is obtained. For a final assessment, a process optimization for this limited number of solvent candidates should be performed.

3. Case study: Physical absorption for CO_2 capture

3.1. Problem description

The proposed method is exemplified for the design of physical solvents for CO_2. The study is motivated from the pre-combustion route for carbon dioxide capture and storage. To focus on the characteristics of the proposed framework, the discussion is limited to a single-stage absorption unit. It is straightforward to extend the method to the full process. A solvent is searched that provides the highest separation efficiency with respect to CO_2. Thus, the objective function is given by

$$f(x) = x_{CO2} / y_{CO2}, \tag{1.3}$$

where x and y denote liquid and vapor mole fraction leaving the absorber, respectively. The MESH equations describing the absorber are implemented. The inlet streams and temperatures are fixed. The pressure is constant (p=50 bar). The PCP-SAFT parameters are constrained to the range of values available in the database of existing molecules.

3.2. Results

3.2.1. Step one: Continuous-molecular targeting

In the following, methanol serves as a reference solvent. As initial study, a single property of methanol is varied in seven individual optimizations (Figure 2).

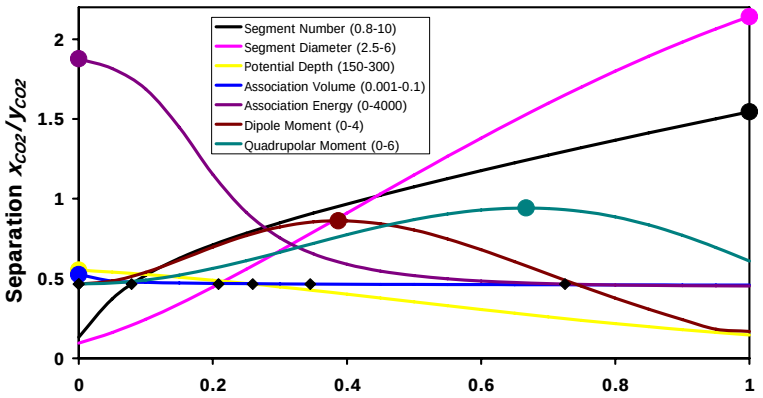


Figure 2: Separation of CO_2 as function of normalized PCP-SAFT parameter values. Circles indicate the optima, and diamonds mark the base case corresponding to methanol.

The method is clearly capable of improving the selectivity. The base case parameters for methanol are PC-SAFT values and do not account for its dipole. The true normalized dipole moment of methanol of is μ=0.425. This value corresponds very well to the maximum found. Indeed, methanol is used commercially in the Rectisol process.

The simultaneous CoMT-optimization of all seven solvent properties generates a hypothetical solvent with a separation factor of 6.31. This value is ten times higher than that of methanol. Indeed, the hypothetical solvent found is a rather extreme molecule. All but one molecular property lie at a bound of the parameter space (Table 1).

Table 1: Normalized PCP-parameters for hypothetical target molecule obtained from CoMT-optimization in comparison to initial guess methanol.

Parameter	m_i	σ_i	ε_i	κ_{iAB}	ε_{iAB}	μ_i	Q_i
Methanol	0.079	0.209	0.259	0.345	0.725	0.	0.
CoMT-optimum	1.	1.	1.	0.	0.512	1.	1.

3.2.2. Step two: Structure mapping

The hypothetical target molecule has now to be related to real substances. The PCP-SAFT database available to us contains parameters for about 120 substances. The structure mapping is performed using the Taylor approximation (1.2) criterion (Table 2).

The solvents suggested from the structure mapping yield significantly improved process performance compared to methanol. It should be noted that the hypothetical target solvent is very far from real existing substances. The average distance in the normalized parameter space is larger than 44%. This undermines the quality of the Taylor approximation as it is based on a local analysis.

Still, the results are intuitively sound as the optimization aims for highest selectivity with respect to CO_2. In an actual process design, the solvent recycling by desorption would also have to be considered. The trade-off between absorption and desorption is expected to lead to less extreme molecular properties.

Table 2: Ranked list of solvents obtained from structure mapping using Taylor approximation criterion. The separation finally obtained is also given. Methanol is included as reference case.

Compound	DMF	DMSO	BNI	PNI	MeOH (PC)	MeOH (PCP)
ΔTA	-5.3	-5.5	-5.8	-6.4	-19.9	-15
Separation	2.53	1.94	2.7	2.59	0.46	1.02

DMF – Dimethylformamide, DMSO – Dimethyl Sulfoxide, BNI – Butyronitrile, PNI – Propionitrile, MeOH – Methanol with PC- and PCP-SAFT parameters, respectively.

4. Conclusions

Integrated solvent and process design is a long-standing challenge. Due to the discrete decisions required, this problem can *de facto* not be solved with a single optimization. In the present work, a continuous-molecular-targeting approach to computer-aided molecular design (CoMT-CAMD) is proposed to address the integrated design by defining a hypothetical solvent. If the hypothetical solvent is characterized by a consistent set of physically-based attributes, a straightforward mapping onto real molecules is possible.
The approach is demonstrated for the design of physical solvents for carbon dioxide capture. Solvents are found that yield significantly higher separation than the commercial solvent methanol. These results demonstrate the potential of the method.
In future work, a more detailed process description will be employed to represent all design trade-offs and to fully exploit the advantages offered by an integrated approach.

References

Achenie, L.E.K., Gani, R. & Venkatasubramanian, V. (2003) Computer Aided Molecular Design: Theory and Practice, Elsevier, Amsterdam.

Odele, O. & Macchietto, S. (1993) Computer Aided Molecular Design: A Novel Method for Optimal Solvent Selection, Fluid Phase Equilibria 82, 47-54.

Hostrup, M., Harper, P.M. & Gani, R. (1999) Design of environmentally benign processes: integration of solvent design and separation process synthesis, Comp. Chem. Eng. 23, 1395.

Eden, M.R., Jørgensen, S.B., Gani, R. & El-Halwagi, M.M. (2004) A novel framework for simultaneous separation process and product design, Chem. Eng. Process. 43, 595-608.

Gross, J. & Sadowski, G. (2001) Perturbed-chain SAFT: An equation of state based on a perturbation theory for chain molecules, Ind. Eng. Chem. Res. 40, 1244-1260.

Pereira, F.E., Keskes, E., Galindo, A., Jackson, G. & Adjiman, C.S. (2008) Integrated design of CO2 capture processes from natural gas. In: Georgiadis, M.C., Kikkinides, E.S. & Pistikopoulos, E.N., Energy Systems Engineering, Wiley-VCH, Weinheim, pp.231-248.

Gross, J. (2005) An equation-of-state contribution for polar components: quadrupolar molecules, AIChE Journal 51(9), 2556-2568.

Gross, J. & Vrabec, J. (2006) An equation-of-state contribution for polar components: dipolar molecules, AIChE Journal, 52(5), 1951-1961.

Gross, J. & Kleiner, M. (2006) An equation-of-state contribution for polar components: polarizable dipoles, AIChE Journal 52(3), 1194-1204.

Lymperiades A., Adjiman, C.S., Galindo, A. & Jackson, G. (2007) A group contribution method for associating chain molecules based on the statistical associating fluid theory (SAFT-γ). J. Chem. Phys. 127, 234903.

Thi, C.L., Tamouza, S., Passarello, J.P., Tobaly, P. & Hemptinne. J.C. de (2006) Modeling Phase Equilibrium of H2 + n-Alkane and CO2 + n-Alkane Binary Mixtures Using a Group Contribution Statistical Association Fluid Theory Equation of State (GC-SAFT-EOS) with a kij Group Contribution Method, Ind. Eng. Chem. Res. 45, 6803-6810.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Systematic Representation and Property Prediction of Fatty Systems for Process Design/Analysis in the Oil and Fat Industry

Carlos A. Diaz-Tovar[a], Roberta Ceriani[b,c] , Rafiqul Gani[a] and Bent Sarup[d]

[a]*CAPEC - Department of Chemical and Biochemical Engineering - Technical University of Denmark.Søltofts Plads, Building 229, 2800 Kgs. Lyngby, Denmark*
[b]*Department of Food Engineering, State University of Campinas, P.O. Box 6121, 13083-862 Campinas – SP, Brazil*
[c]*Faculty of Applied Sciences, State University of Campinas,13484-350 Limeira – SP, Brazil*
[d]*Vegetable Oil Technology Business Unit, Alfa Laval Copenhagen A/S, Maskinvej 5, DK-2860 Søborg, Denmark*

Abstract

The solvent recovery section of a soybean oil extraction process has been studied with special emphasis on the effect of design variables on the performance of the process. Two different sets of simulations were carried out to analyze these effects. The first set of simulation results were analyzed with the Plackett-Burman (PB) method to identify the parameters (process variables) that make the greatest impact on the process. The second set of simulations was used to generate a reduced model relating the design variables to the identified parameters. Through the PB methodology, it was possible to obtain a reduction of 40% in the number of process variables to be considered in the performance analysis. Even tough, the response surface pointed out that the process is already within the optimum zone, optimization with the reduced model pointed further improvement in the amount of hexane recovered.
Keywords: Simulation, optimization, solvent recovery, surface response methodology, vegetable oil.

1. Introduction

Fats and oils are a source of energy and play an important role in human nutrition. The production of edible oils/fats involves a great variety of processing steps and unit operations, and state of the art in process modeling has only to a very limited extent penetrated this industry and part of the reason is the complex nature of the lipid systems involved. Lipid technology refers to products and processes (modification, purification, separation, etc.) that involve fatty acids, their derivatives, and related substances. This work develops computer aided methods and tools for the systematic design and analysis of a process that employs lipid technology, the solvent recovery section of the soybean oil extraction process.

It is known that the soybean oil is almost exclusively produced by solvent extraction and the most widely used solvent for extraction of oil from oilseeds is hexane, because it has high solubility for fats and oil, chemical stability and a suitable vapor pressure allowing for easy separation by flash and stripping operations [1]. The hexane recovery section (see Figure 1) includes four parts: An oil recovery, a condensation system, a mineral oil system and a water-solvent separation. After the "washing" of the soybean

flakes, both the miscella (mixture of oil and solvent) and the white flakes (extracted flakes wet with solvent) are heated separately to remove the solvent. The flakes go to a desolventiser–toaster–drier–cooler (DTDC) unit and the miscella enters the first part of the solvent recovery section, the oil recovery system, where the oil is concentrated. Here the solvent is removed from the oil that is subsequently concentrated nearly to 100% (crude oil). All the recovered solvent vapors mixed with steam, including those coming from the DTDC facility, are retrieved by means of three sub-systems that make part of the solvent recovery system: the condensation, the MOS (mineral oil system), and the water–solvent separation. Therefore, vapors are condensed and the residual vapors are captured in the mineral oil system. The condensates from the condensation system enter a water–solvent separation section, where the solvent is recovered and after mixing with a make-up stream, is redirected to the extractor.

The aim of this work is to optimize the performance of the solvent recovery section of the soybean oil extraction process in terms of the important design variables. This is not an easy task since the number of design variables involved in the process (temperature, pressure, and number of stages) is quite large, and optimization of the process performance with a rigorous simulation model in this case, is quite complex. An alternative approach is to use techniques such as the Plackett-Burman and the Full Factorial Designs [2], which are useful tools to identify, observe, and analyze the critical parameters of the process and their effects on the process behavior subject to changes in the design variables; the objective here is to obtain a reduced model. Two techniques are considered here: the first applies a screening procedure to discard design variables that do not affect the overall performance of the process; the second establishes the order to perturb the remaining design variables to observe their effect on the process performance defined in terms of the identified process parameters.

The solution procedure employed in this work consists of three main steps: Process simulation, process model reduction, and process optimization. The process simulation step consists of generation of the properties of the chemicals (lipid compounds: glycerides, fatty esters, and unsaponifiable compounds) present in the process and creating a database for use in process simulation models. The model reduction step applies design of experiments techniques to identify the most significant design variables, the parameters affecting the process performance, and creates a reduced "quadratic" model relating the design variables and the process parameters. The optimization step determines the design variables values corresponding to the optimal process performance.

2. Methods and tools used

2.1. Property prediction of lipid compounds

Oils and fats oils are complex chemical mixtures of compounds such as glycerides, fatty acids, fatty esters, sterols, and tocopherols. In this work, thirteen different types of the most widely used vegetable oils (soybean, sunflower, canola, etc.) and two different kinds of fats (lard and tallow) were reviewed in the literature in order to define their typical chemical composition.

The database created from this literature review includes a total of 28 fatty acids (from C6 to C24), 51 triacylglycerides (from C28 to C58), 44 diacylglycerides (from C14 to C40), 16 monoacylglycerides (from C8 to C22), 58 fatty esters (methyl and ethyl from C6 to C24), and 26 minor compounds (sterols, tocopherols, and tocotrienols). The database also includes predicted properties for the lipid compounds such as melting

point (a new model has been developed); vapor pressure [3]; densities [4]; and critical properties that were calculated by means of the software ICAS™ ProPred [5].

2.2. Process simulation model

The solvent recovery section consists of several unit operations: flash, heath exchange, mixing, splitting, stripping, and absorption. In this work, the process section was simulated using the commercial software PRO II and the database of lipid compounds integrated to PRO II.

The process model describing the solvent recovery section is shown in Figure 1 [1]. According to the flow sheet the streams where both the hexane solvent and vegetable oil can be lost are the following: Crude_Oil, Vap_Out, Solid_Out, S33 (Purge), Mos_Water, and Vap_Mos. The corresponding process design variables are listed bellow:

- Oil recovery section: Temperature and pressure of evaporator 1 (TEVAP1 & PEVAP1), temperature of evaporator 2 (TEVAP 2), and temperature and number of stages (NS) of the stripping column (TSTP1 & NSTP1).
- Condensation section: Pressure of the condenser 1 (PCOND1) and temperature of condenser 2 (TCOND2).
- Water-hexane separation section: temperature and pressure of the decanter (TDECANT & PDECANT).

Process parameters:

- Oil recovery section: Pressure of evaporator 2 (PEVAP 2), duty and pressure drop of evaporator 3 (QEVAP3 & DPEVAP3), and pressure of the stripping column (PSTP1).
- Condensation section: Temperature of the condenser 1 (TCOND1) and pressure of condenser 2 (PCOND2).
- MOS System: Temperature, pressure, and number of stages of the absorption and stripping columns (TABSPC, PTABSPC, NTABSPC, TSTP2, PSTP2, & NSTP2).

2.3. Model Reduction

The Plackett-Burman method is a technique where only the main effects are considered to be significant, achieving thereby a reduction in the number of design variables. The advantage of this technique is that it only requires a number of experimental runs that are a multiple of 4 instead of a power of 2 [2]. The full factorial design is a technique that is sufficient for evaluating many production processes because it considers not only individual effects but also non distinguishable effects on a response.

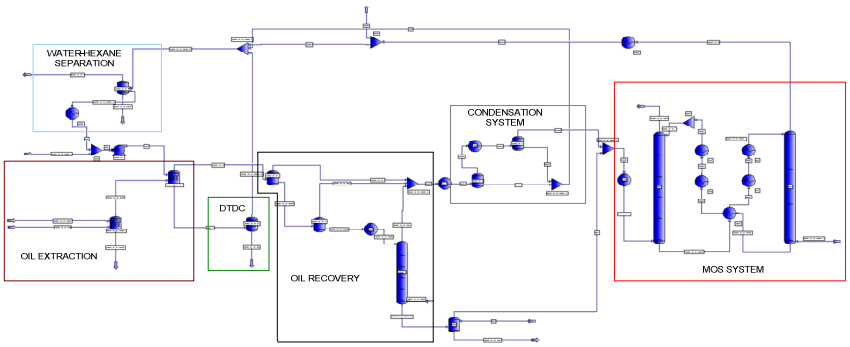


Figure 1 Solvent Recovery PRO II Model [1]

3. Case Study: Optimization of a Soybean Extraction Process

3.1. The Optimization Problem

The goal of this optimization problem is to minimize the loss of hexane solvent and vegetable oil (product) in the solvent recovery process section.

Objective Function

$$OF = Crude\_Oil(Hex) + Vap\_Out(Hex) + Solid\_Out(Hex) + S33(Hex) + Vap\_Mos(Hex) + Mos\_Water(Hex) + Vap\_Mos(Oil) + Mos\_Water(Oil) \quad (1)$$

Constraints

$26.46 \le TCOND2(°C) \le 35.80$ (2)

$25.51 \le TDECANT(°C) \le 28.19$ (3)

$53.34 \le TEVAP1(°C) \le 72.16$ (4)

$88.00 \le TEVAP2(°C) \le 132.00$ (5)

$104.50 \le TSP1(°C) \le 115.50$ (6)

$0.26 \le PCOND1(bar) \le 0.36$ (7)

$0.18 \le PDECANT(bar) \le 0.42$ (8)

$0.39 \le PEVAP1(bar) \le 0.73$ (9)

Number of Stages ≥ 1 (10)

3.2. Results

3.2.1. Process Simulation

Data related to the solvent recovery process was available, so it was possible to make a detailed analysis of the accuracy of the soybean oil extraction process simulation. The average relative deviation (ARD) between the industrial data and the simulated data for the crude oil and the overall loss of hexane was analyzed according to,

$$ARD = \frac{\sum \left| x_i^{\exp} - x_i^{\text{model}} \middle/ x_i^{\exp} \right|}{N} * 100 \quad (11)$$

Where $x_i^{\exp}$ is the plant data for vegetable oil and hexane mass compositions in stream i, and, x_i^{model} are the calculated data for the same variables for N different experimental (plant) runs. The ARD for the oil composition in the crude-oil stream is 0.31%, while the ARD for the loss of hexane in the process is 4.71%.

The analysis of the simulation results generated through the process simulation for various (perturbed) values of the design variables led to the following observations:

- Presence of hexane in the stream Solid_Out can be neglected because, according to the plant data, the DTDC process removes all hexane and water from the white flakes.
- The major loss of vegetable oil in the process is located in the solvent extraction process, where some oil may remain in the waste solid. Since a predictive model to calculate this variable was not available, the stream Solid_out was not considered in the model reduction step.
- The amount of oil in the streams S33 (Purge) and Mos_Water can be neglected as these values were found to be very small.
- Two liquid phases (water and hexane) are present in the vessel named DECANTER. Therefore, stream tagged as Vap_Mos in the flow sheet was neglected in the further solution steps.

Based on the above observations, the objective function (see Equation (1)) is reduced to

$$OF = Crude\_Oil(Hex) + Vap\_Out(Hex) + S33(Hex) \quad (12)$$

Equation (12) actually corresponds to the amount of hexane recovered in the process, which will now be maximized.

$$\text{Objective Function} = \text{Recovery of Hexane} \quad (13)$$

3.2.2. Model Reduction

Plackett-Burman Design: The ±1 levels (perturbed values) used in the 16 simulations for each of the design variables are listed in Table 1. From analysis of the simulation results it was concluded that only five variables significantly affected the amount of hexane recovered in the process.

Full Factorial Design: Based on the results of the PB method a second set of 43 simulations (2^5+2*5+1) was performed using the full factorial technique. From a statistical analysis of the results, a quadratic model for the Hexane Recovery response was obtained. The model, Equation (14), for maximizing hexane recovery is given in terms of 3 design variables (TCOND2, TEVAP2, and TSP1) as only these variables affect the hexane content (see Table 2). The ARD value for the reduced model is 0.008%. Using Equation (14) the response surface for this model was generated (see Figure 2).

Table 1 Plackett-Burman technique levels

		LEVELS	
CODE	**VARIABLE**	**-1**	**1**
V1	TCOND2	26.46	35.80
V2	PCOND1	0.26	0.36
V3	TDECANT	25.51	28.19
V4	PDECANT	0.18	0.42
V5	TEVAP1	53.34	72.16
V6	TEVAP2	88.00	132.00
V7	TSTP1	104.50	115.50
V8	NSTP1	4.00	6.00
V9	PEVAP1	0.39	0.73

Table 2 Plackett-Burman technique results

	HEXANE (Kgmol/hr)		
	Vap_Out	**Crude_oil**	**Purge**
SIGNIFICANT VARIABLES	VAR1	VAR1	VAR1
	VAR2	-	VAR2
	-	-	-
	-	-	-
	VAR5	-	VAR5
	-	VAR6	-
	-	VAR7	-
	-	-	-
	-	-	-

$$\text{Hexane Recovery} = -0.005T_{EVAP2}^2 + 0.013T_{EVAP2} + 0.0054\left(T_{EVAP2}\right)\left(T_{SP1}\right) - 0.0178T_{SP1} - 0.0047\left(T_{SP1}\right)\left(T_{TCOND2}\right) + 99.963 \quad (14)$$

3.2.3. Optimization

The optimization model of the solvent recovery process is represented by Equations (15)-(18). As stated before, the goal is to maximize the objective function

$$\textbf{Max} \quad \text{Objective Function} = \text{Hexane Recovery} \quad (15)$$

s.t.

$$26.46 \le T_{TCOND2}(°C) \le 35.80 \quad (16)$$

$$88.00 \le T_{EVAP2}(°C) \le 132.00 \quad (17)$$

$$104.50 \le T_{STP1}(°C) \le 115.50 \quad (18)$$

The optimal values of the design variables and the percentage of hexane recovered under the optimized condition are: $T_{\mathrm{TCOND2}} = 35.80°C$, $T_{\mathrm{EVAP2}} = 110.14°C$, $T_{\mathrm{STP1}} = 104.50°C$, and $\mathrm{Recovery} = 99.98\%$.

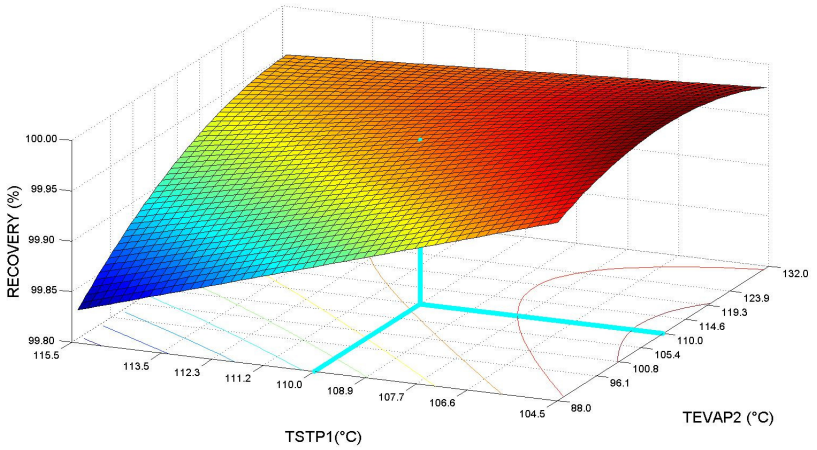


Figure 2 Hexane Recovery Surface Response

4. Conclusions

The predicted properties of the lipid compounds included in the database have been validated by comparing the simulation of the process and the plant data. The first set of operational variables considered in this work would have needed more than 500 simulations if the full factorial technique was used without any pre-analysis. Therefore, the Plackett-Burman method was helpful to significantly reduce the total number of simulations that needed to be performed. With the reduced set of design variables, the full factorial technique was applied and a reduced model correlating the selected design variables and the process parameters was obtained. The development of the reduced model allows the representation of the process in terms if the most important design variables. The response surface generated through the reduced model showed that the process is within the optimum zone, although improvements are possible. Therefore, optimal values of design variables where hexane recovery is maximum, was found. The results confirm that through a combination of property database and process modelling techniques, the operation and design of industrial lipid technologies can be optimized.

References

1 A. Martinho, H. A. Matos, R. Gani, B. Sarup, W. Yougreen, 2008, Modeling and Simulation of Vegetable Oil Processes, Food and Bioproducts Processing, 86, 87, 9.

2 D. C. Montgomery, 2005, Chapter 6: The 2^k Factorial Design, Design and Analysis of Experiments, 6^{th}, 51.

3 R. Ceriani, J. A. Meirelles, 2004, Predicting vapor-liquid equilibria of fatty systems, Fluid Phase Equilibria, 215, 2, 10.

4 J. D. Halvorsen, W.C. Mammel, Jr, L.D. Clements, 1993, Density estimation for fatty acids and vegetable oils based on their fatty acid composition, Journal of the American Oil Chemists' Society70, 9, 6.

5 J. Marrero, R. Gani, 2001, Group-contribution based estimation of pure component properties, Fluid Phase Equilibria, 183-184, 26.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

The Virtual Product-Process Design Laboratory for Design and Analysis of Formulations

Elisa Conte[a], Ricardo Morales-Rodriguez[a], Rafiqul Gani[a]

[a]*CAPEC, Department of Chemical & Biochemical Engineering, Technical University of Denmark, Soltofts Plads, Building 229, 2800 Lyngby, Denmark*

Abstract

The objective of this paper is to present a virtual laboratory for chemical product-process design (virtual PPD-lab) as a tool for the design and analysis of formulations. With the virtual PPD-lab, users can test their design ideas on model-based computer-aided tools before performing experiments to validate the designed product; also, manifold product and process alternatives matching *a priori* defined targets can be generated. The significance of this virtual laboratory is that the experimental effort in the development of new products and processes can be drastically reduced and attention can be focused on few alternatives; as a consequence, time and resources can be saved. This paper highlights a new feature of the virtual PPD-lab through which the design of formulations can be achieved. The application of this new feature in formulation design is illustrated through different case studies dealing with personal care products.

Keywords: Product Design, Mixture Design, Computer-Aided Formulation Design.

1. Overview of the Virtual Laboratory

In design of chemical products and the processes that can manufacture them, one first tries to find a candidate product that exhibits a certain desirable or targeted behavior and then tries to find a process that can manufacture it with the specified qualities. The candidate may be a single chemical, a mixture, or a formulation of active ingredients and additives. The common practice to develop these products is experiment-based trial and error approach, supplemented sometimes with model-based computer-aided tools to speed-up some of the steps. The virtual PPD-lab is an innovative and more effective approach to product and process design. It contains methods and tools to allow the modeling and simulation of the needed experimental scenarios. For this to work, the 'in-house' models need to be reliable and efficient and the architecture of the software needs to include the work-flows related to different product-process design problems. Also, interfaces for efficient data-flow between different tools need to be defined. The architecture has to be flexible to allow changes in work-flow, addition of new models and of new data for future extension of the application range.

From the virtual PPD-lab it is possible to access ICAS[1], an Integrated Computer Aided System consisting of a number of toolboxes that help to efficiently solve a wide range of problems. Some ICAS toolboxes can be launched as standalone applications for the virtual PPD-lab, such as CAPEC DataBase, the Computer Aided Molecular Design tool (ProCAMD), the Property Prediction tool (PropPred), and the Modeling Tool (MoT). These toolboxes can easily be updated with newer and more reliable models[2]. To guide users in solving specific product and process design problems, templates defining work- and data-flow for different types of design problems have been developed for the virtual

PPD-lab. Two design templates were available in its first version: design of polymeric microcapsules for the controlled release of active ingredients, and, design of pesticide formulations for plant uptake. A third template, which guides the user to design solvent based formulations, has now been added and is presented in this paper as a necessary tool in the design of formulated products.

2. Computer-Aided Formulations Design

Many chemicals-based products of everyday life such as sun lotions, shower creams, paints, insect repellents, and so on, are formulations. These formulations are consumer-oriented products, since they are directly used by the consumers and in order for them to be commercially successful products, they have to meet the consumer needs. These needs can be manifold: an insect repellent, for instance, has to give an effective protection against mosquito bites, to be long lasting, safe, easily applicable, stable, with a pleasant odor. In order to achieve all these different performance criteria, several chemicals have to be blended together, and this leads to formulations. Usually, three main categories of substances are involved: the Active Ingredient (AI), responsible for the activity of the final product (for an insect repellent, the activity is to be effective against mosquito); a solvent mixture; additives (dispersants, emulsifiers, wetting agents, usually less than 2% in volume) responsible for enhancing the end-use product properties. The solvent mixture has a central role in a formulation, since it is usually present in high concentrations and has the important functions of dissolving the AI and providing for some of the end-use product properties; therefore, a tool for designing solvent mixtures, has been developed and has been included in the virtual PPD-lab environment; the new tool is based on the MixD algorithm[3] and has been developed to design liquid mixtures.

In this paper, a methodology for the design of formulations is proposed, and it is shown in Figure 1. The first and most important task is the problem definition (Task 1), which consists of three sub-tasks. In Task 1.1 an understanding of the consumer needs has to be achieved. The consumer needs are the product performance criteria. In Task 1.2, the performance criteria have to be translated in terms of chemical properties (target properties); for instance, an insect repellent has to be long lasting and this performance criteria depends on the evaporation time T_{90}, which is a property that can be modelled and controlled during the design. Then, constraints on the target

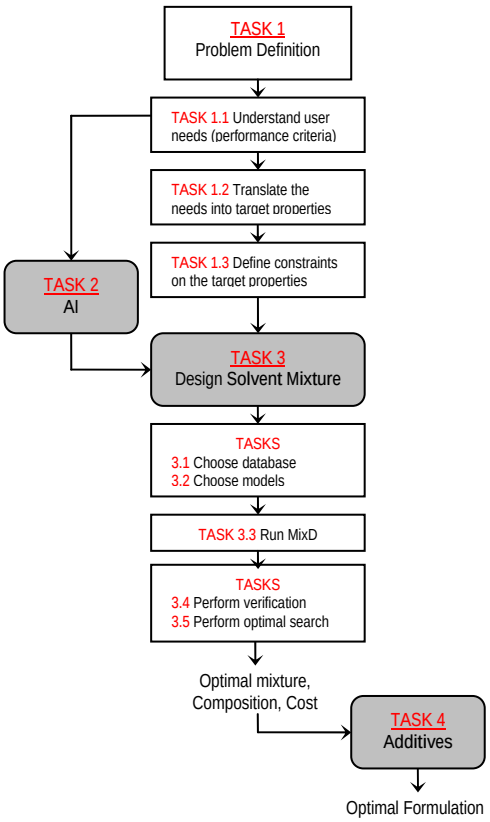


Figure 1. Formulation design methodology.

properties have to be settled in Task 1.3, in the form of lower and upper bounds. After Task 1, the desired activity of the product is known, therefore one or more Active Ingredients need to be chosen (Task 2). The AI affects all the design choices which will be taken in the next task, Task 3, the solvent mixture design. As stated above, the solvent mixture has two important functions: dissolving the AI (also supporting and enhancing its activity), and providing for specific end-use product properties. In order to ensure the AI solubility, at Task 3.1 an appropriate list of solvents needs to be chosen (for instance, if the AI is well known to be oil-soluble, the solvents considered in the database should all be miscible with oils); in addition, it is important to consider the solubility parameter of the mixture as a target property, since when two substances have similar values for the solubility parameters, they are also mutually soluble.
In Task 3.2 the property models with which the target properties have to be modelled are chosen. MixD is now launched (Task 3.3); results consist of all the single liquid phase binary mixtures matching the *a priori* defined constraints, their composition and cost. Task 3.4 involves the verification[3] of the designed mixtures, while Task 3.5 performs the optimal mixture search[3] according to a Performance Index PI.
Once the optimal mixture, its composition and cost have been calculated, the necessity of final additives is analyzed in Task 4. These additives are added in order to enhance the end-use properties of the product; in personal care products, for instance, perfumes or relieving/moisturizing agents can improve the cosmetic properties.

3. Case studies

The methodology developed for the design of formulations is highlighted through two case studies dealing with personal care products: the first formulation is an insect repellent (water-free and water-based products are presented), while the second one is a sunscreen lotion. The design of a paint formulation has previously been presented[3].

3.1. Insect repellent formulations

When consumers think about an insect repellent, they want a product which is effective against mosquito, long lasting, safe, compatible with other material, stable (it does not have to decompose in a relatively short time), and, finally, a product with good cosmetic properties, that is, nice odor and good skin feeling.
The traditionally used DEET-based products have been claimed[4] to be aggressive on surfaces like clothes, glasses, plastics, to show a high potential to irritate eyes, to be sticky and to posses an unpleasant odor. The insect repellent we are interested in aims to be a competitor of DEET-based products; the more recent picaridin has been shown to be far superior to DEET in terms of safety, toxicology, material compatibility and cosmetic properties[4], and it has therefore been chosen as the AI for our formulation.
The main product types available are alcohol-based liquids, emulsions, and sticks. For this case study it has been chosen to design a liquid solution based spray product.
Picaridin provides for the effectiveness of the product; the solvent mixture (including additives) which has to be designed has to guarantee that the final product will match all other consumer needs. The durability of the formulation depends on the evaporation rate of the solvent mixture, which corresponds to the parameter T_{90} (evaporation time: time required for 90% by weight of the mixture to evaporate); the safety can be related to the LC_{50} parameter (aqueous concentration causing 50% mortality in a fathead minnow population after 96 hours); material compatibility can be guaranteed by choosing non aggressive and non corrosive solvents; stability can be controlled by assuring that the solubility parameter of the mixture is close to that of the AI (δ_{AI} = 24.1 $MPa^{½}$), and verifying that the condition for one liquid phase has been satisfied (a step in the MixD

algorithm[3]). The chosen product aims to be a spray: the form of the formulation will be controlled with the mixture kinematic viscosity (v) and density (ρ); products with kinematic viscosity values lower than 75 cS are suitable for a spray formulation[5].
Picaridin shows high solubility in alcohols, and usually insect repellent products are alcohol-based; consequently, for this design, only alcohols (74 alcohols) are retrieved from the database. For modeling all the target properties linear mixing rules are sufficient, except for the evaporation time model[6], which employs the UNIFAC model.
The single phase solvent mixtures matching the target criteria are listed in Table 1 (ordered in terms of increasing cost, the PI). It has been decided not to verify the mixture properties calculated with the linear models, since alcohols are all polar and associating fluids, and mixtures of similar fluids usually do not show large contributions for the excess properties, that is, linear models are reliable.

Table 1. Mixtures suitable for alcohol-based insect repellent formulations. 'AllAl' stands for allyl alcohol. δ [MPa^{½}], v [cS], ρ [kg/l], LC_{50} [mol/l], T_{90} [s], Cost [$/kg].

Mixtures	x_1	δ	v	ρ	LC_{50}	T_{90}	**Cost**
methanol - 1-butanol	0.33	24.89	2.42	0.80	0.05	773	1.74
2,2,3-trimethyl-3-pentanol - AllAl	0.16	26.41	2.89	0.93	0.06	887	1.77
methanol - 2,2-dimethyl-1-butanol	0.55	26.33	2.60	0.84	0.05	805	1.78
methanol - 2-methyl-2-heptanol	0.70	26.96	2.85	0.82	0.06	1096	1.79
2-methyl-2-heptanol - AllAl	0.14	26.54	2.61	0.91	0.05	663	1.81
ethanol - 2,2-dimethyl-1-butanol	0.71	25.27	2.40	0.83	0.05	631	2.30
ethanol - 2,2,3-trimethyl-3-pentanol	0.79	25.34	2.90	0.85	0.05	1084	2.31
ethanol - 2-methyl-2-heptanol	0.82	25.47	2.52	0.82	0.05	755	2.40
methanol - 2-methyl-1-propanol	0.35	25.13	2.98	0.79	0.05	528	2.44
2,2-dimethyl-1-butanol - AllAl	0.23	26.26	2.51	0.91	0.05	598	2.75

If an optimization with respect to the cost is performed (but others or multiple performance index can also be used), the mixture methanol and 1-butanol is found to be the optimal mixture.
Usually, a fragrance is added to these products in order to cover the unpleasant odor of the AI (as in the case of DEET based repellents); picaridin does not have an unpleasant odor, but a pleasant scent could be an additional product quality; α and β-Santalol[7] are almost colorless, slightly viscous alcohols, rarely available separated. They are both soluble in alcohols and oils, so they are suitable for addition to the formulation just designed. A sweet, woody and tenacious scent makes them easily recognizable.
The formulation just designed is water-free, because picaridin has a very low solubility in water; but some commercially available products (Autan®, Bayer) are based on a water-alcohol mixture, therefore, a water-based formulation has also been designed; to ensure picaridin is highly soluble in the alcohol, it has been verified that the solubility parameter of the alcohol in the final mixtures is very close to that of picaridin.
Water is a polar and associating fluid like alcohols, and mixtures of similar fluids should be reliably described by linear models. Although, it is well know that aqueous solutions may exhibit unusual behavior; consequently, it has been decided to verify that the single phase liquid mixtures obtained with MixD indeed have the correct kinematic viscosity (using a more reliable model based on UNIFAC[8]); all mixtures were found to match the target viscosity and are reported in Table 2; these are all single liquid phase mixtures.

Table 2. Mixtures suitable for water-based insect repellent formulations. Units of measure are the same specified for Table 1.

Mixtures	x_1	δ	v	ρ	LC_{50}	T_{90}	**Cost**
2-propanol - water	0.24	42.00	1.31	0.87	0.52	661	0.92
ethanol - water	0.27	42.00	1.01	0.89	0.58	734	1.42
1-propanol - water	0.25	42.00	1.28	0.88	0.47	628	2.07

The presence of water in the formulations improves both the safety and the price. It is interesting to note that the mixture ethanol-water, with a composition of 27% (molar base) of ethanol corresponds to the one used in Autan® (29% of ethanol), but the design just performed reveals that it is not the cheapest alternative: 2-propanol with water (24% of 2-propanol) is the optimal solution with respect to the cost.

3.2. Sunscreen liquid formulations

The sun produces a wide range of electromagnetic radiation. Ultraviolet light is responsible for suntan and sunburn and increases the risk of skin damage. Sunscreens are cosmetic formulations that block UV rays. Consumers want a product which is effective against sunburns, prevents skin aging, is long lasting, is safe, has a good skin feeling and has a nice scent. Water resistant formulas are usually preferred; in addition, a spray product will be designed.

In order to have an effective protection against UV-A and UV-B rays, and prevention of skin aging, the active ingredients that may be used are many: UV-A and UV-B filters; antioxidants; insoluble pigments, since they are opaque to light so they provide a physical barrier for radiations. Since a waterproof formulation has to be designed, all chemicals involved have to be water insoluble; all the AIs chosen are oil-soluble substances. The AIs chosen are: avobenzone[9] and octyl salicylate[9] as UV-A and UV-B filters, respectively; α, β-carotene and vitamin A as antioxidants; titanium dioxide as an opacifying agent (it also acts like a thickener).

Since the performance criteria for a sunscreen are quite similar to those of an insect repellent (they both are personal care spray products), except for the activity, the target properties could be considered the same for these two products; therefore, similar constraints involving δ, v, ρ, LC_{50}, T_{90} are specified. A solvent mixture that can dissolve or disperse the AIs so that sedimentation is avoided and that can provide for the product end-use properties has to be designed; esters are largely used in personal care and in pharmaceutical applications for their interesting properties; they are oil-soluble substances so they perfectly fit the purpose of designing a waterproof product; in addition, they support the rapid and uniform spreading of the sensorially heavy UV filters. From the database, 97 esters has been retrieved, and they are employed in this case study. Considerations on the choice of the models are the same as in §3.1.

Table 3. Mixtures suitable for sunscreen formulations. 'MAcAl' and 'dimethprop-3-methbut' stand for methoxyacetaldehyde and dimethylpropyl 3-methylbutanoate. LC_{50} [mol/m^3].

Mixtures	x_1	δ	v	ρ	LC_{50}	T_{90}	**Cost**
MAcAl - 2,2-dimethylpropyl butanoate	0.89	18.95	0.53	0.83	3.63	1018	1.40
MAcAl - tert-butyl pentanoate	0.89	18.95	0.53	0.83	3.63	1018	1.40
MAcAl - isobutyl isopentanoate	0.89	18.93	0.48	0.83	3.65	879	1.40
MAcAl - 1,1-dimethprop-3-methbut.	0.91	18.91	0.52	0.83	3.86	847	1.41
MAcAl - 2,2- dimethprop-3-methbut.	0.91	18.92	0.53	0.83	3.80	940	1.41
MAcAl - isobutyl 3,3-dimethylbutanoate	0.91	18.92	0.53	0.83	3.80	940	1.41

The single phase mixtures matching the targets are listed in Table 3. Mixtures show very similar properties since methoxyacetaldehyde is always present in high concentrations and some of the second compounds in the mixtures are isomers with similar properties. Verification of the mixture properties calculated with the linear models is not necessary in this case, since esters are all polar and non-associating compounds, and their mixtures have negligible excess properties.
Since the costs of the designed mixtures are approximately the same, there is no reason to use the cost as the optimization criteria. The mixture suggested as the optimal is methoxyacetaldehyde and 1,1-dimethylpropyl 3-methylbutanoate, since it shows the highest LC_{50} value, which corresponds to the lowest toxicity value.
The necessity of additives can be now analyzed. Octocrylene can be used as a stabilizer; it also augments other UV absorption and improves their uniform skin coating[9]; it is an ester, therefore suitable for the formulation. Parabens are widely used as preservatives in the cosmetic and pharmaceutical industries; they prevent decomposition by microbial growth or by chemical changes. Also in sunscreens a fragrance can be added in order to enhance cosmetic properties; para-menth-3-yl phenylacetate[7] is a slightly viscous liquid with mild floral and extremely sweet and tenacious notes; iso-propyl salicylate[7] is an oily liquid, with a sweet ethereal herbaceous, yet quite tenacious scent with distinctly fruity character. They both are esters, therefore suitable for the formulation designed.

4. Conclusions and future work

The virtual PPD-lab has been introduced and the use of the resident models, methods and tools has been illustrated. A systematic methodology developed for formulation design has been highlighted through two case studies; this methodology employs the algorithm MixD which has been previously presented[3]. The significance of the PPD-lab is that the time and the effort in the development of formulated products can be drastically reduced. Two case studies dealing with personal care formulations have been shown. Future work is to enlarge the number and the complexity of products which can be designed and tested through the virtual PPD-lab.

References

1. ICAS Documentation, 2003, Internal report, CAPEC, KT-DTU, Lyngby, Denmark.
2. Conte, E., Martinho, A., Matos, H. A., Gani, R., 2008, Combined group-contribution and atom connectivity index-based methods for estimation of surface tension and viscosity, *Ind. Eng. Chem. Res.*, 47(20), 7940-7954.
3. Conte, E., Morales-Rodriguez, R., Gani, R., 2009, The virtual product-process design laboratory as a tool for product development, accepted for ESCAPE 19.
4. Badolo, A., Ilboudo-Sanogo, E., Patoin Ouédraogo, A., Costantini, C., 2004, Evaluation of the sensitivity of Aedes aegypti and Anopheles gambiae complex mosquitoes to two insect repellents: DEET and KBR 3023, *Am. J. Trop. Med. Hyg.*; 9 (3), 330–334.
5. Bagajewicz, M. J., 2007, On the role of microeconomics, planning, and finances in product design, *AIChE J.*, 53(12), 3155-3170.
6. Klein, J. A., Wu, D. T., Gani, R., 1992, Computer aided mixture design with specified property constraints, *Comp. and Chem. Eng.*, 16(5), 229-236.
7. Arctander, S., 1969, Perfume and flavor chemicals. Aroma chemicals. Allured.
8. Cao, W., Knudsen, K., Fredenslund, A., Rasmussen, P., 1993, Group-contribution viscosity predictions of liquid mixtures using UNIFAC-VLE parameters, *Ind. Eng. Chem. Res.*, 32, 2088-2092.
9. Ansmann, A., Gondek, H. Kawa, R., Tesmann, H., 2001, Sun screen agents, US Patent 6280712.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Catalyst deactivation in reactive distillation

Rui M. Filipe[a], Henrique A. Matos[b] and Augusto Q. Novais[3]

[a]*Departamento de Engenharia Química, Instituto Superior de Engenharia de Lisboa R. Conselheiro Emídio Navarro, 1, 1959-007 Lisboa, Portugal.*
[b]*Centro de Processos Químicos, Departamento de Engenharia Química e Biológica, Instituto Superior Técnico, Av. Rovisco Pais, 1049-001 Lisboa, Portugal.*
[c]*Departamento de Modelação e Simulação de Processos, Instituto Nacional de Engenharia, Tecnologia e Inovação, Est. Paço do Lumiar, 1649-038 Lisboa, Portugal.*

Abstract

This work addresses the effect of catalyst deactivation in reactive distillation. A rigorous dynamic model developed in gPROMS is used to assess the impact in the column of the change in the catalyst activity. A model system of the olefin metathesis, wherein 2-pentene reacts to form 2-butene and 3-hexene was used to evaluate the column behavior and some output parameters, such as product purity and energy consumption. Strategies to overcome the deactivation are addressed through the use of different feed qualities and reboil ratios.

Keywords: reactive distillation, catalyst deactivation, modeling, simulation.

1. Introduction

Reactive distillation combines reaction and separation into one physical shell leading to potential economic and environmental gains (Malone and Doherty, 2000; Taylor and Krishna, 2000). This process intensification also leads to systems with significantly greener engineering attributes (Malone *et al.*, 2003).

In the authors' previous work, a framework combining feasible regions and optimization techniques has been developed for the design and multiobjective optimization of complex reactive distillation columns and reacting systems with variable degrees of relative volatilities (Filipe *et al.*, 2008b). This led to the consideration of reactive distillation columns with distributed feeds, many of them involving the combination of superheated and subcooled feeds. This combination was found to provide a source or a sink of heat at specified trays of the columns, which favors reaction while reducing the total reactive holdup requirements. It was also found that higher conversions could be obtained with the same reactive holdup by using these feed qualities outside the traditional range, which led to the consideration of using this technique to overcome catalyst deactivation during column operation.

Catalyst deactivation represents both an operational and a design problem, and has received little attention from the research community. Wang *et al.* (2003) address the control of reactive distillation columns when the production rate changes or the catalyst deactivates, and propose a control scheme able to maintain high purity and high conversion under such conditions. Catalyst deactivation reduces the reaction conversion achieved at each tray, which may limit column performance and product specifications. If deactivation is taken into account at the design stage, the operating conditions over the catalyst life-cycle may be assessed and preventive actions taken at an early stage.

In this work, the effects of the catalyst deactivation are investigated using the optimal design specifications obtained previously regarding feasibility and performance. Additionally, the manipulation of the feed quality and reboiler ratio were used to investigate how to tackle catalyst deactivation. With this aim, a rigorous dynamic model developed in gPROMS was used and applied to an illustrative example, i.e., the olefin metathesis system, wherein 2-pentene reacts to form 2-butene and 3-hexene.

2. Dynamic model

To investigate the dynamic behavior of a reactive distillation column with catalyst deactivation, a rigorous dynamic model has been developed using gPROMS language. The model was built modularly and allows for different number of trays and feeds, and feed qualities. Mass and energy balances are used at each element of the column. Pressure drop over the column is considered and calculated from the vapor flow speed and liquid height at each tray. Deviations from phase equilibrium can be accounted for through the built-in Murphree stage efficiency equation although they were neglected in this work. The physical properties are estimated using the gPROMS included package IPPFO for ideal systems. The coefficients used within IPPFO library were regressed using data from Aspen Plus. The reaction is considered to occur only on the liquid phase at specified trays of the column. Instead of specifying the catalyst amount, a reactive holdup amount is set on these trays.

Five control loops are used. The distillate and bottoms flowrate are manipulated to control reflux drum and reboiler levels, respectively. The pressure on the first tray is controlled through the energy removed in the condenser. The purity of the heavy component in the bottoms is controlled by manipulating the reboiler duty. Reflux ratio is manipulated to control the ratio D/F through the measurement of the distillate flow leaving the column. This control configuration allows the manipulation of the feed energy content by adjustment of the feed temperature while holding constant the reboil and D/F ratios, conditions previously established in the column design and optimization stages.

The design used in this work was obtained from the previously reported design and optimization framework (Filipe *et al.*, 2008a; Filipe *et al.*, 2008b) applied to the olefin metathesis system. All the RD column design details (Table 1) were transferred from the Aspen Plus simulations performed prior to gPROMS implementation. The availability of these design details from previous Aspen Plus runs proved valuable for the initialization of the simulations, which is always a major task when dealing with complex gPROMS models.

A good agreement between both models was achieved as shown in Table 1, with the largest difference being observed in the condenser and reboiler duty (6.6% and 5.5%). Nevertheless the temperatures at these sections of the column are the same in both models: 277 K and 341 K, respectively. Product purity refers to the amount of 2-butene in the distillate stream and 3-hexene in the bottoms stream, which is the same, due to the ratio D/F used (0.5) and the stoichiometry of the reaction.

3. Catalyst deactivation

Catalyst deactivation is simulated through the inclusion of a multiplication factor in the reaction rate constant, which varies between zero (no catalyst activity) and one (full catalyst activity). With this option, the variation of the catalyst activity during his life cycle can be simulated. For convenience, the variation of this factor will be mentioned as catalyst deactivation from now on.

Table 1. Case study details and results comparison

	Aspen	gPROMS
Number of stages	14	
Feed tray	8	
Reboil ratio	3.20	
Reflux ratio	6.02	6.22
Feed rate (mol/s)	5.56	
Condenser duty (kW)	-460.68	-491.27
Reboiler duty (kW)	256.58	270.57
Purity (mol %)	97.77	98.00
Reactive trays	6-10	
Reactive trays holdup (kmol)	4.97	
Feed temperature (K)	401 (q = -0.37)	
D/F	0.5	
Column diameter (m)	0.61	
Pressure at first tray (Pa)	101325	

A negative exponential decay over time is considered for catalyst activity, which decreases to 50% after 81 days of service (Figure 1). Purity is maintained above 90% until activity decreases below 33%, after which if falls steeply. Figure 2 shows that reboiler and condenser duties remain practically constant when catalyst deactivates, for the specified conditions (fixed reboil ratio and D/F).

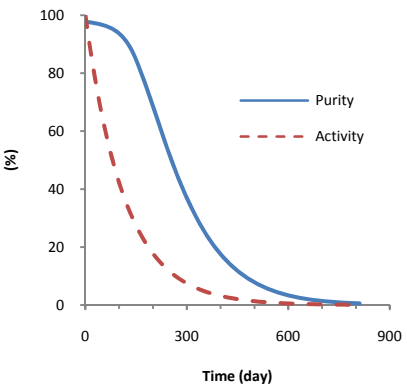


Figure 1. Variation of product purity and catalyst activity with time

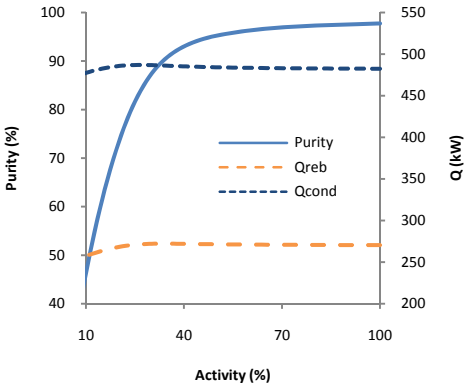


Figure 2. Variation of product purity, condenser and reboiler duty with catalyst activity

4. Influence of the feed quality

The use of feed qualities outside the traditional range has proved to be important in the reduction of the required total reactive holdup (Hoffmaster and Hauan, 2006; Filipe *et al.*, 2008b) and thus in attaining designs with lower cost indicator (Filipe *et al.*, 2009). Feed quality is related to the energy content of the feed through equation (1), where F is the feed flow, q the feed quality, and L_{n-1} and L_n the liquid entering and leaving the tray n, respectively.

$$L_n = L_{n-1} + q \cdot F \tag{1}$$

The direct specification of the feed quality is not possible in the model developed as it would involve the specification of the feed enthalpy, a dependent variable. In fact, the specification of the feed is made with values of temperature, pressure and composition which for a single component stream limit the conditions to below ($q > 1$) and above ($q < 0$) temperature's boiling point. A combination of two feeds with temperature above and below the boiling point with different contributions to the total flow could be used to obtain a mixed feed with quality comprised between $0 < q < 1$. Although possible, this strategy has not been used in this work. A variation of the feed temperature was used to assess the effect of the feed quality outside that range. The discontinuity observed in Figure 3 and 4 is due to the vaporization enthalpy contribution at the 2-pentene boiling point (310 K)

Figure 3 depicts the variation of the purity when the feed temperature varies between 273 K and 410 K. Although the gain obtained with the vaporized feed is notorious with an increase of over 2% in product purity, no apparent benefits result from overheating the vapor, as the change in product purity is marginal. On the contrary, changing the temperature under the boiling has a visible effect on purity.

Figure 4 shows that the reboiler duty remains practically unchanged, apart a small variation when the feed is vaporized. This behavior is justified by the fixed reboil ratio used and also to the feed being located near the middle of the column. The condenser has to remove all the extra energy added through the feed in order to maintain the constant pressure and, consequently, its duty increases with the feed temperature.

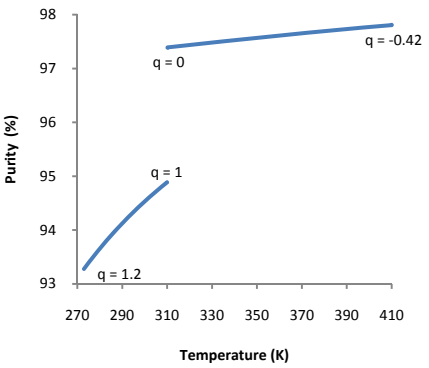


Figure 3. Variation of product purity with feed temperature

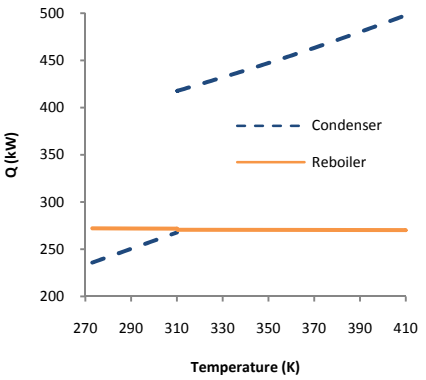


Figure 4. Variation of reboiler and condenser heat duty with feed temperature

The effect of the feed quality was also investigated for situations with partial deactivation of the catalyst. Figure 5 compares the evolution of the purity with the feed temperature for scenarios where catalyst has 50, 75 and 100% activity. The effect of feed temperature on purity is found to increase with the catalyst deactivation, highlighting the impact of the feed temperature on product purity, particularly for those situations where the catalyst has partially lost its activity. These results are also in accordance with the previous finding that reactive holdup can be reduced by manipulation of the feed quality (Hoffmaster and Hauan, 2006; Filipe *et al.*, 2008b). Furthermore, it explains the selection of the extreme feed quality values when minimizing the reactive holdup, if no other restriction is used (Filipe *et al.*, 2008b), since the purity will keep increasing.
Nevertheless, and for the present case, the effect of the feed quality is not strong enough to override large changes in catalyst activity, except when the feed boiling temperature is surpassed. The low reactive holdup associated with this design contributes to this difficulty, once it leaves a small margin to deal with further decreases in the already low catalyst availability. On-going research addresses optimal designs with higher reactive holdup, where a more marked influence of the feed quality is expected.

5. Influence of the reboil ratio

Since feed quality change is not enough to overcome the catalyst deactivation effect on the product purity, the effect of the manipulation of the reboil ratio was also studied. This variable provides an indirect way to vary the column energy supply. Temperature feed was adjusted from 401K to 320 K, with a decrease in purity of 0.54%, considered as not significant for the temperature differential involved.
Figure 6 depicts the effect of the reboiler ratio on product purity for different catalyst activities at a feed temperature of 320 K. The positive effect of increasing the reflux ratio is more noticeable for low activities. This effect has been previously noticed for the feed quality. The changes in the reboiler ratio are, therefore, found to be insufficient to achieve higher purities.

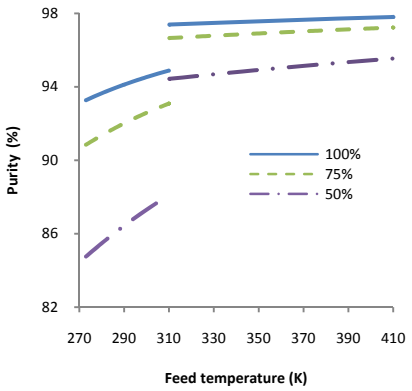


Figure 5. Variation of product purity with feed temperature for different catalyst activities

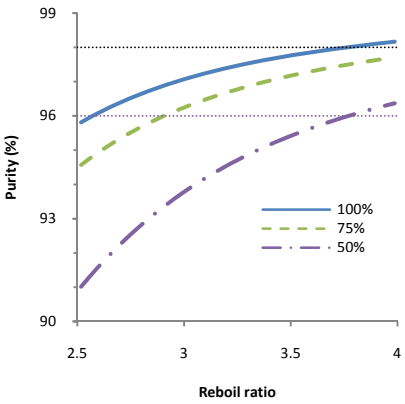


Figure 6. Variation of product purity with reboiler ratio for different catalyst activities (T_{feed} = 320 K)

However, the reboiler ratio can be used to handle catalyst deactivation for a larger period if a lower product purity specification is allowed. Figure 6 shows that 98% product purity can only be reached with full catalyst activity, within the reboiler ratios considered. Nevertheless, when a new specification of 96% is considered, all three scenarios of activity can reach the goal, although with different reboiler duties requirements, as seen in Table 2. An increase of 47% in reboiler duty is required when activity decays to 50%, indicating the need for a careful assessment of the right moment for catalyst replacement.

Table 2. Reboiler ratio and reboiler duty with catalyst activity for a 96% product purity specification

Catalyst activity (%)	Reboiler ratio	Reboiler duty (kW)
50	3.78	320.4
75	2.91	246.8
100	2.57	218.0

6. Conclusions and final remarks

This paper investigates the effect of catalyst deactivation in reactive distillation columns. A rigorous dynamic model developed in gPROMS is used to simulate the reduction of catalyst activity and assess its effects in column performance. The influence of the feed quality, i.e., the energy content of the feed, is also addressed and its positive effect verified. The impact of the feed was found to be small, especially for temperatures above the boiling temperature of the feed. Future work will investigate different design configurations for which a higher impact is expected.

Policies to overcome catalyst are also addressed with the manipulation of the reboiler ratio to enlarge the operation time of the column with the same catalyst load. It was found that for a lower product purity specification this can be an effective way to maintain product purity. Future work will focus on the optimization of the catalyst replacement policies, through the analysis of the costs involved.

References

R.M. Filipe, S. Hauan, H.A. Matos, A.Q. Novais, 2008a, A sensitivity analysis on optimal solutions obtained for a reactive distillation column. 18th European Symposium on Computer Aided Process Engineering. B. Braunschweig and X. Joulia. Lyon, France. 25: 211-216.

R.M. Filipe, S. Turnberg, S. Hauan, H.A. Matos, A.Q. Novais, 2008b, Multiobjective Design of Reactive Distillation with Feasible Regions. Industrial & Engineering Chemistry Research, 47, 7284-7293.

R.M. Filipe, H.A. Matos, A.Q. Novais, 2009, Performance indicators in reactive distillation design. Chemical Product and Process Modeling, Accepted.

W.R. Hoffmaster, S. Hauan, 2006, Using feasible regions to design and optimize reactive distillation columns with ideal VLE. AIChE Journal, 52, 1744-1753.

M.F. Malone, M.F. Doherty, 2000, Reactive Distillation. Industrial & Engineering Chemistry Research, 39, 3953-3957.

M.F. Malone, R.S. Huss, M.F. Doherty, 2003, Green chemical engineering aspects of reactive distillation. Environmental Science & Technology, 37, 5325-5329.

R. Taylor, R. Krishna, 2000, Modelling Reactive Distillation. Chemical Engineering Science, 55, 5183-5229.

S.J. Wang, D.S.H. Wong, E.K. Lee, 2003, Control of a reactive distillation column in the kinetic regime for the synthesis of n-butyl acetate. Industrial & Engineering Chemistry Research, 42, 5182-5194.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Sustainability Analysis of Chemical Processes Plants Using a Hybrid Heuristic and Indicator Model

Ana Carvalho,[a,d] Iskandar Halim,[b] Rajagopalan Srinivasan,[b,c] Henrique A.Matos,[a] & Rafiqul Gani[d]

[a] *CPQ, Instituto Superior Técnico, Av Rovisco Pais, 1049-001, Lisboa, Portugal*
[b] *Institute of Chemical and Engineering Sciences , Jurong Island, Singapore 627833*
[c] *National University of Singapore, 10 Kent Ridge Crescent, Singapore 119260*
[d] *CAPEC, Technical University of Denmark,DK-2800, Lyngby, Denmark*

Abstract

This paper presents an integrated framework comprising of two computer aided systems, ENVOP*Expert* and *SustainPro*, for sustainability assessment of chemical process plants. Analysis based on ENVOP*Expert* diagnoses the waste sources, identifies heuristic design alternatives and highlights trade-offs between the environmental and economic objectives. The Indicator model of *SustainPro* evaluates further these alternatives and screens them at an in-depth level through its sustainability metrics analysis to identify sustainable design alternatives. The application of the framework is illustrated through an ammonia production case study.

Keywords: sustainability; process design; intelligent system; metric-based analysis; indicators

1. Introduction

One of the major concerns facing the world today arises from increasing industrial activities that have lead to rapid depletion of non-renewable resources and increase in pollution. With the current emphasis on sustainability, much improvement is expected from the process industry to minimize raw material and energy usage and waste generation without compromising the economic value of the enterprise. Responding to these challenges requires a new insight into the characteristics of a sustainable system and a fundamental rethinking of how a production unit is to be designed, built and operated [1]. Various computational tools have been proposed to assist the process engineers in the decision-making to identify and evaluate sustainable design alternatives (those that improve a sub-set of sustainability metrics). An expert system, called ENVOP*Expert*, has been developed for waste minimization analysis of chemical processes [2]. Another tool is *SustainPro* which applies a set of mass and energy indicators for generation and evaluation of sustainable design alternatives [3]. In this paper, we propose an integrated framework comprising both of these for a comprehensive sustainability study. We illustrate the framework by solving a case study involving the ammonia production process based on synthesis gas route.

2. An Integrated Framework of ENVOP*Expert* and *SustainPro*

The two computer aided systems are integrated according to the framework shown in Figure 1. The procedure involves the following steps.

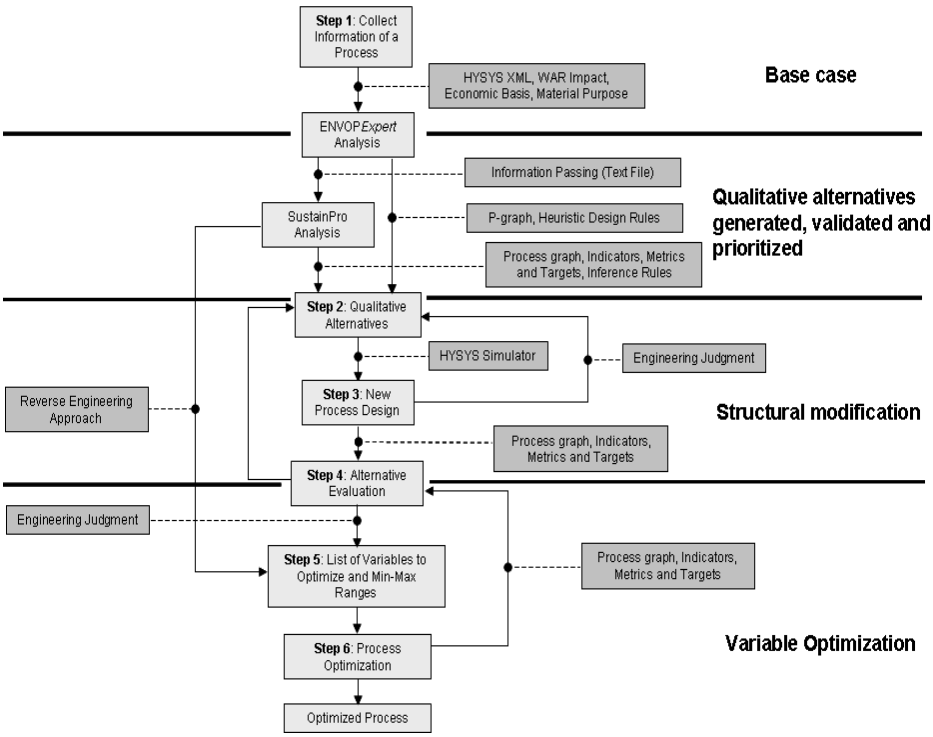


Figure 1: Integrated framework of ENVOP*Expert* and *SustainPro*

Step 1 − Process information representation

Information about the base process including the flowsheet, mass and energy balance information, material and stream classification, chemical environmental impact (WAR method) and economic data are specified. The flowsheet and material and energy balance information can be extracted from any process simulator (for example, HYSYS) in a XML format. This information is processed by ENVOP*Expert* to a suitable text file format decipherable by *SustainPro*.

Step 2 − Qualitative alternatives generation

Qualitative design modifications are identified by ENVOP*Expert* in two-step procedure: first a P-graph analysis is performed to identify the root cause of waste generation, next the hierarchical design method is used to generate heuristic alternatives. In parallel, a set of mass and energy indicators is calculated in *SustainPro* to identify the critical points in the process with respect to material accumulation and process economics. The most significant indicators are next identified using a sensitivity analysis method to propose design modifications (changes in design variables that improve targeted indicators, which in turn, improve a corresponding sub-set of sustainability metrics). The next step is to generate sustainable design alternatives (those that match the target indicators).Generally, the *SustainPro*'s alternatives are in-line with the suggestions of ENVOP*Expert*. However, the alternatives of *SustainPro* are more specific. For instance, using the material accumulation indicator, a high build-up of certain materials in a recycle stream can be diagnosed and hence suggestions for reducing this material

recycle can be proposed. Such specific suggestion cannot be inferred by ENVOP*Expert* as its analysis is purely based on qualitative reasoning.

Step 3 − New process design

The suggested new design alternatives are implemented to the base process. This is done by modifying the flowsheet of the simulator (HYSYS) followed by simulation of the new flowsheet to generate mass and energy balance data.

Step 4 − Alternative evaluation

The sustainability of the new alternatives are evaluated using *SustainPro*, where a set of sustainability and safety metrics, ([4] and [5]) are calculated and used for comparing the base and the new process. The indicators are again calculated for the new alternative to verify that the targets have been improved and the others have not been disturbed. This step is repeated for each design alternative.

Step 5 – Define the optimization problem

The new process alternative from step 4 is used as an initial estimate to further refine the search for the optimal sustainable process design. Thus, in this step, a list of process variables that can be manipulated is specified. *SustainPro* then screens these variables by prioritizing the ones that may significantly affect the process. The minimum and maximum ranges for optimizing these variables are also estimated based on reverse approach which uses sustainability metrics and other operational constraints.

Step 6 − Process variable optimization

Single or multi-objective optimization problem involving the selected process variables is solved. This is done through a simulation-optimization option within ENVOP*Expert*. The sustainability and safety metrics of the optimized process are calculated and compared with the base process using the *SustainPro* package.

3. Ammonia Production – Case Study

The developed framework has been applied on an ammonia production process based on synthesis gas route. The basic flowsheet of this case study is shown in Figure 2. A feed-gas stream containing nitrogen, hydrogen, and impurities (argon and methane) is passed through a multiple-bed catalytic reactor equipped with internal cooling. The products from the reactors are condensed using a coolant and then separated in a flash separator, where ammonia liquid is collected as product and the vapor mixture is recycled back to the reactors. To prevent the build-up of impurities within the process, a purge-gas stream is used, which becomes the only waste stream from this process. The step-by-step procedure employed to improve the sustainability of this process is described next.

Step 1 − Information of the process

Information about the process is obtained from a HYSYS simulation model. Here, a series of plug flow reactors is used as a representation model for the multi-bed reactor unit. The economic data, material and stream classifications and chemical environmental impacts are listed in Table 1.

Step 2 − Qualitative alternatives

Table 2 lists some of the qualitative suggestions identified by ENVOP*Expert* and *SustainPro*. As seen from the table, the two sets of alternatives complement each other.

Step 3 − New design process

Through the use of set of mass and energy *SustainPro* identified the critical paths in the process by performing a sensitivity analysis on the operational parameters related to these paths. This analysis pointed to the reduction of methane in the feed-gas stream as the most promising because of its effects on the targeted indicators. Following the

methodology described in [4], one possible separation process to reduce this impurity is to use a membrane. A specific membrane for this type of mixtures is reported in [5]. Hence, this alternative is implemented in the HYSYS model of the base process and a new simulation was carried out.

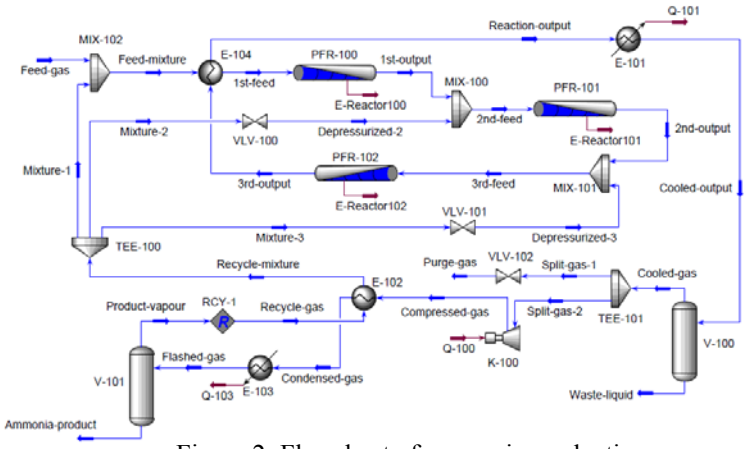


Figure 2: Flowsheet of ammonia production

Table 1: Ammonia process information

Stream	Status	Value/Cost	Material	Purpose	Env. Impact (mass basis)
Feed-gas	Feed	$ 0.13/kg	Methane	Impurity	0.0408
Purge-gas	Waste	$ 0.005/kg	Water	Utility	0.0
Ammonia-product	Product	$ 0.3/kg	Hydrogen	Raw material	0.0
Q-100	Energy	$ 0.07/kWh	Nitrogen	Raw material	0.0
Q-101	Energy	$ 0.006/kWh	Ammonia	Product	3.984
Q-103	Energy	$ 0.1/kWh	Argon	Impurity	0.0

Table 2: Qualitative alternatives

System	Qualitative Suggestions
ENVOP*Expert*	1) Prevent excessive feed of materials in the feed-gas stream, remove impurities (methane and argon) from feed-gas 2) Increase reactant conversion in the reactors 3) Recycle or recovery-recycle useful materials (nitrogen, hydrogen and ammonia) from purge-gas stream
SustainPro	1) Separate methane in the feed-gas stream before entering the process 2) Reduce the recycle of methane using a new separation process

Step 4 − Alternative evaluation

The new sustainability and safety metrics are calculated using the *SustainPro* and compared with the base case. Table 3 shows that the new alternative improves all the sustainability indicator metrics while maintaining the same safety index to their base values. Therefore, it is concluded that the new design is more sustainable.

Step 5 − -Define the optimization problem
Here, we consider feed and purge flowrates as the decision variables (i.e., two degree of freedoms). Using the reverse approach combined with sensitivity analysis, *SustainPro* verifies that manipulating the feed flowrate offers the best improvement to the process. The minimum and maximum range for the feed flowrate is found to be 1.22×10^5 < flowrate < 4×10^5 kg/h. This range is obtained by considering the maximum allowable profit loss (1%) and the constraints of the reactor performance.

Table 3: Selected sustainability metrics for the base and improved process

Metrics	Base	New	Improvement
Total Net Primary Energy Usage rate (GJ/y)	19336	18968	1.90%
% Total Net Primary Energy sourced from renewables	NA	NA	NA
Total Net Primary Energy Usage per Kg product (kJ/kg)	23.74	23.43	1.31%
Total Net Primary Energy Usage per unit value added (kJ/$)	0.012	0.011	8.33%
Total raw materials used per kg product (kg/kg)	1.23	1.21	1.63%
Total raw materials used per unit value added	6.14×10^{-4}	5.80×10^{-4}	5.54%
Fraction of raw materials recycled within company	NA	NA	NA
Net water consumed per unit mass of product (kg/kg)	7.92	7.76	2.02%
Net water consumed per unit value added	3.96×10^{-3}	3.71×10^{-3}	6.31%
WAR per kg product (impact/kg)	0.1541	0.1524	1.11%
Safety Index	26	26	0%
Profit ($/yr)	1.13×10^8	1.21×10^8	8.03%

Step 6 − Variables optimization
With the specified range, the feed flowrate is optimized using simulated annealing. A single objective problem involving profit maximization was considered. Table 4 lists the sustainability and safety metrics for two optimized process (local and global optima). As seen from the table, the profit reaches a maximum when operating near the maximum feed-gas flow. The other sustainability and safety metrics are also shown to be constant or slightly improved at this maximum flow. The only metric that becomes worse is the total energy used per year. However per kg of product, the energy consumption fraction is maintained constant, which means that to produce the same amount of product the energy spent is the same.

4. Conclusions

An integrated framework comprising of ENVOP*Expert* and *SustainPro* has been developed for sustainability analysis of chemical processes. The integrated analysis diagnoses the waste sources accurately, recommends design alternatives to improve the process and reduces the optimization space by screening the variables with most impact

on the process. The application of the framework has been illustrated on a continuous ammonia production process. Future work will include extension and application of the framework to cover batch process operations.

Table 4: Sustainability metrics for the optimized process

Metrics	Base	Local optima (Feed-gas = $3.55x10^5$ kg/h)	Global Optimal (Feed-gas = $3.97x10^5$ kg/h)
Total Net Primary Energy Usage rate (GJ/y)	19336	54792	61232
% Total Net Primary Energy sourced from renewables	NA	NA	NA
Total Net Primary Energy Usage per Kg product (kJ/kg)	23.74	23.44	23.42
Total Net Primary Energy Usage per unit value added (kJ/$)	0.012	0.011	0.011
Total raw materials used per kg product (kg/kg)	1.23	1.21	1.21
Total raw materials used per unit value added	$6.14x10^{-4}$	$5.81x10^{-4}$	$5.81x10^{-4}$
Fraction of raw materials recycled within company	NA	NA	NA
Net water consumed per unit mass of product (kg/kg)	7.92	7.76	7.76
Net water consumed per unit value added	$3.96x10^{-3}$	$3.70x10^{-3}$	$3.70x10^{-3}$
WAR per kg product (impact/kg)	0.1541	0.1524	0.1524
Safety Index	26	26	26
Profit ($/yr)	$1.13x10^8$	$3.51x10^8$	$3.93x10^8$

References

1. Bakshi, B.R. and Fiksel, J., 2003, The Quest for Sustainability: Challenges for Process Systems Engineering, AIChE Journal, 49: 1350-1358
2. Halim, I. and Srinivasan, R., 2002, An Integrated Decision Support System for Waste Minimization Analysis in Chemical Processes, Environmental Science and Technology, 36: 1640-1648
3. Carvalho, A., Matos, H. and Gani, R., 2008, Design of Sustainable Chemical Processes: Systematic Retrofit Analysis Generation and evaluation of alternatives, Process Safety and Environmental Protection, 86: 328-346
4. Azapagic, 2002, Sustainable Development Progress Metrics, *IChemE Sustainable Development Working Group*, IChemE Rugby, UK
5. Heikkilä, A.-M, 1999, Inherent Safety in Process Plant Design – An Index-based Approach, Ph.D-Thesis, VTT Automation, Espoo, Finland
6. Jaksland, C., Gani, R. and Lien, K., 1995, Separation Process Design and Synthesis Based on Thermodinamic Insights, Chemical Engineering Science, Vol. 50, No 3, 511-530
7. Wang, D., Teo, W. K, Li, K., 2002, Permeation of H2, N2, CH4, C2H6, and C3H8 Through Asymmetric Polyetherimide Hollow-Fiber Membranes, Journal of Applied Polymer Science, Vol. 86, 698–702

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Automated design of batch water-allocation network

Hong-Guang Dong,[a] Li-Juan Li,[a] Rui-Jie Zhou[b]

[a]*School of Chemical Engineering, Dalian University of Technology, Dalian 116012, PRC*
[b]*Department of Economics, Dalian University of Technology, Dalian 116024, PRC*

Abstract

The mathematical technique presented in this work deals with one step design of batch water-allocation network (WAN), where batch schedules, water-reuse subsystems, and wastewater-treatment subsystems are all taken into account. In the first place, a flexible schedule model is introduced to represent the precedence order of all operations. Then, two novel superstructures incorporating all basic elements (i.e. states, tasks, equipment and time) are adopted to capture all production schemes and batch WAN configurations. Specifically, by adding novel components in the original superstructure, a series of optimal network structures with multistage splitting and mixing options which have never been contained within previous superstructure can be easily generated. Finally, a reliable optimization strategy, where deterministic and stochastic searching techniques are combined, is suggested to deal with the resulting mixed integer non-linear programming (MINLP) model. Two illustrative examples are presented to demonstrate the effectiveness of the proposed approach.

Keywords: superstructure, batch schedules, batch water-allocation network.

1. Introduction

In the past, the tasks of optimizing batch schedules, water reuse and wastewater treatment subsystems were performed individually. Cheng and Chang (2007) first developed an effective procedure to incorporate these three components into a single comprehensive model. However, in their study, the discrete-time model embedded was inflexible. In addition, not all possible network structures were included in their superstructure and this superstructure may fail to wholly reflect the essential relationship between units and corresponding operations. Finally, interactions between process schedules and batch WAN were not clearly addressed mathematically, which might result in solving the proposed model on a sequential basis. To circumvent these problems, Zhou et al. (2008) integrated batch schedules and WAN into a simultaneous optimization model based on continuous time representation and the modified state space concept. Although the relationship between operations and units was addressed by the STN (Kondili, 1993) and SEN (Smith, 1996) extensions of the original state space framework (Bagajewicz and Manousiouthakis, 1992), yet assigning a fixed mixer and splitter to each unit during the whole time horizon inevitability leads to the preclusion of a class of optimal network structures, where the best cost-optimal scheme may actually lies. Finally, the continuous time formulation is unable to express the sequence of all operations and the optimization strategy has to be executed interactively between different components. Given these shortages, there is therefore a need to develop a more comprehensive design method for optimizing the batch WANs.

2. Superstructures

2.1. Superstructure for batch production

A novel superstructure has been introduced in this work to represent batch productions. (see Figure 1). In the first place, all material states embedded in distribution network (DN block) are classified into three groups-raw materials, intermediate products and final products, all of which are further divided into two parts: material states to be consumed/sold and produced/bought. Then, through the process operators (OP block), each water-using unit is illustrated as a block with many sub-blocks which correspond to operations performed at certain event points. Finally, all material states and units are connected through streams to reflect the material flows and a reference axis of time is set to identify the starting and ending points of each operation.

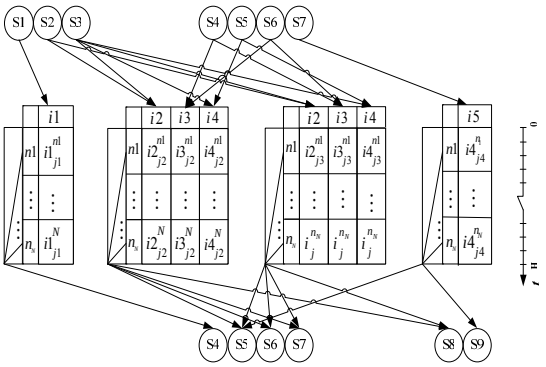


Figure.1. Superstructure for batch production.

2.2. Superstructure for WAN

In the previous superstructure (Zhou et al. 2008), the splitting and mixing can only take place in DN where a one-to-one correspondence between mixers/splitters and water users is ensured, and the possibilities that streams bypass water users without mass exchange were completely ruled out. As a result, streams can only mix once before entering the water user and in the same way they are only allowed to split once after leaving the water user. To overcome these deficiencies, we have developed a comprehensive STS superstructure for optimizing batch WAN.

The improved state-space superstructure for batch WAN is also viewed as two inter-connected blocks (see Figure 2). One is referred to as the distribution network (DN), in which all connections between units and junctions are embedded. The other is the so-called process operators (OP), which can be further divided into two sub-blocks, i.e., units and junctions. Detailed explanations of our superstructure are described in the sequel. In one time interval, while all input streams to DN are allowed to connect all exits leading to OP block or the environment, only inputs from junctions are allowed to split into several branches. Likewise, more than one input is allowed to mix before all junctions. Finally, in any time interval, splitting of inputs from units and mixing of streams before units are forbidden. As for OP, each unit is illustrated as a block with many sub-blocks which correspond to operations performed at certain event points. Furthermore, the optimal number and the corresponding assignment of units to junctions are not fixed a priori, thus being subjected to optimization process. It is worthy of note, that junctions can provide increased flexibility in formulating network

structures with multistage mixing/splitting options, which can never be captured by other frameworks. Specifically, all mixed inputs to DN block have further opportunities to match with other possible streams at all junctions and, if necessary, such looping procedure within junctions can be carried out for arbitrary times.

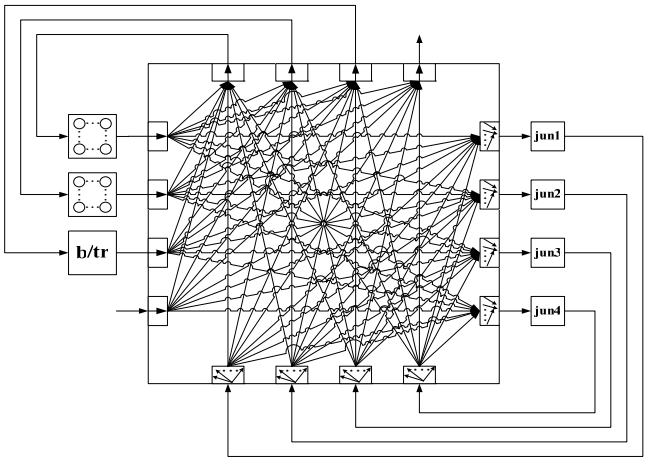


Figure.2. Superstructure for WAN.

3. Mathematical Model

Due to the length of this paper, only parts of the constraints are presented as follows.

3.1. Models for Batch schedules

3.1.1. Material balances

$$ST_{s,n} = ST_s^{in} - \sum_{i \in I_s} \rho_{s,i}^c \sum_{j \in J_i} B_{i,j,n} + r_{s,n} - d_{s,n} \qquad \forall s \in S, n \in N (n = n_1) \qquad (1)$$

$$ST_{s,n} = ST_{s,n-1} - \sum_{i \in I_s} \rho_{s,i}^c \sum_{j \in J_i} B_{i,j,n} + \sum_{i \in I_s} \rho_{s,i}^p \sum_{j \in J_i} B_{i,j,n} + r_{s,n} - d_{s,n} \qquad \forall s \in S, n \in N (n > n_1) \qquad (2)$$

$$ST_{s,N} = ST_s^{in} \qquad \forall s \in S \qquad (3)$$

Constraint 3 expresses that the amount of each state at the starting and ending times of any operation cycle should be kept identical.

3.1.2. Duration constraints

$$T_{i_j^n}^f = T_{i_j^n}^s + \alpha_{i,j} \cdot wv(i, n) + \beta_{i,j} \cdot b_{i_j^n} \qquad \forall i \in I,\ j \in J_i,\ n \in N \qquad (4)$$

$$B_{i_j^n} / B_{i,j} \le b_{i_j^n} < \left(B_{i_j^n} + B_{i,j} \right) / B_{i,j} \qquad \forall i \in I,\ j \in J_i,\ n \in N \qquad (5)$$

$$T_{i_j^n}^s = \chi_{i_j^n} \cdot \Delta t \qquad \forall i \in I,\ j \in J_i,\ n \in N \qquad (6)$$

$$T_{i_j^n}^f = \varphi_{i_j^n} \cdot \Delta t \qquad \forall i \in I,\ j \in J_i,\ n \in N \qquad (7)$$

where $\alpha_{i,j}$ and $\beta_{i,j}$ denote respectively the constant and each variable term of processing time; $B_{i,j}$ and $b_{i_j^n}$ are assumed, respectively, to be the amount processed at each variable term and the number of variable terms needed; Δt represents the length of each variable interval. Constraints 4-7 show that the processing time of each operation is made up of two parts: fixed term and variable terms.

3.2. Model for WAN

3.2.1. Mass and flow balances of each operation

$$f_{u,t}^{in} = \sum_{u' \in U} fs_{u',u,t} \qquad \forall u \in U, t \in T \qquad (8)$$

$$f_{u,t}^{in} \cdot c_{u,k,t}^{in} = \sum_{u' \in U} fs_{u',u,t} \cdot cs_{u',k,t} + \sum_{jun \in JUN} fs_{jun,u,t} \qquad \forall u \in U, k \in K, t \in T \qquad (9)$$

$$f_{u,t}^{out} = \sum_{u' \in U} fs_{u,u',t} + \sum_{jun \in JUN} fs_{u,jun,t} \qquad \forall u \in U, t \in T \qquad (10)$$

$$c_{u,k,t}^{out} = cs_{u,k,t} \qquad \forall u \in U, k \in K, t \in T \qquad (11)$$

3.2.2. Mass and flow balances of each unit

$$f_{j,t}^{in} = \sum_{i_j^n \in I_j^n} f_{i_j^n,t}^{in} \qquad \forall j \in J, t \in T \qquad (12)$$

$$f_{j,t}^{out} = \sum_{i_j^n \in I_j^n} f_{i_j^n,t}^{out} \qquad \forall j \in J, t \in T \qquad (13)$$

$$c_{j,k,t}^{in} = \sum_{i_j^n \in I_j^n} c_{i_j^n,k,t}^{in} \qquad \forall j \in J, k \in K, t \in T \qquad (14)$$

$$c_{j,k,t}^{out} = \sum_{i_j^n \in I_j^n} c_{i_j^n,k,t}^{out} \qquad \forall j \in J, k \in K, t \in T \qquad (15)$$

3.3. Objective function

The criterion of this study is the minimizing of cost, which can be expressed as follows:

$$Obj = (total\ income - purchasing\ cost) - (cost\ of\ fresh\ water + cost\ of\ water\ treatment + cost\ of\ buffer\ tanks + cost\ of\ junctions) \qquad (16)$$

4. Application examples

4.1. Example 1

The first example presented was originally solved simultaneously based on discrete-time representation by Cheng and Chang (2007). All the data used can be found in the original work. In a latter study, when network complexity was taken into account, Zhou et al. (2008) obtained a batch WAH design, of which the profit of batch production and cost of network were found to be 1366.7 and 388.31 units respectively. In the present study, On the basis of a time of horizon of 4 hr, the most appropriate production scheme and network configuration are presented in Figures 3. It can be found that while the schedule scheme is the same as the former study, the network configuration has subjected to a major transformation, as the multistage mixing/splitting stream from Jun1-Jun2 can be clearly identified. The corresponding cost of network in a production cycle can be reduced to 377.01 units, which represents a 3% improvement. This can be attributed to the fact that by allowing multistage mixing/splitting, the economic trade-off issues in WAH design can be carried out effectively. Specifically, the costs of freshwater supply, wastewater treatment, buffer tanks, junctions and pipelines were found to be 123.33, 201.68, 0 and 52 units respectively. As evident from the network structure, buffer tank is not embedded in the optimal network configuration and this can explain that buffer tanks are not necessary in some batch WAN if optimal production sequences and operating policies of water flows are adopted. Finally, it can be observed that the introduction of junctions allows a higher integration of the overall batch WAN, since only 2 junctions and 14 pipelines are needed.

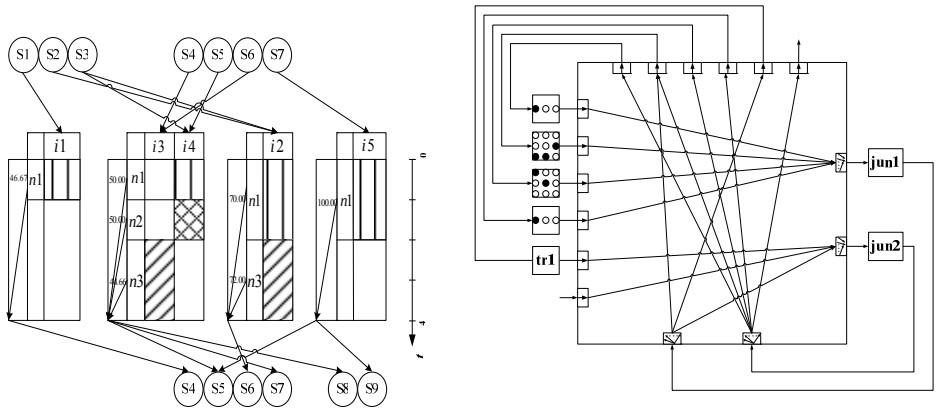


Figure.3. Optimal production scheme and network configuration for Case I

4.2. Example 2

In this example, multi-contaminant system design options are incorporated. Consider the process presented by Zhou et al. In this study, the removal ratios of tr1 and tr2 are chosen to be 0.9, 0.85, 0.9 and 0.95, 0.9, 0.95 respectively; the corresponding cost is set to 0.75 and 1 respectively. The resulting schedule results and its corresponding network are shown in Figure 4 and 5. Notice that the rewards of production and cost of network can be found to be 2108.57 and 333.95 units respectively. The freshwater and wastewater treatment costs in this case are 121.71 and 78.03 units respectively, while the costs of buffer tanks, junctions and pipelines are 67.21 and 67 units respectively. It can be concluded from this example that multistage mixing/splitting is indeed a desirable fashion and such methods enable cost-optimal designs of complex and large scale multi-contaminant systems.

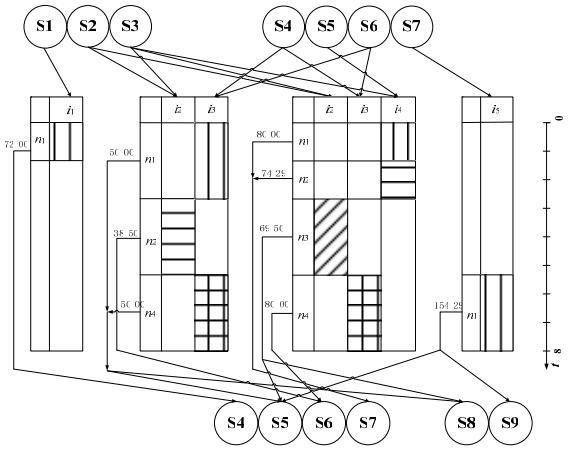


Figure.4. Optimal production scheme for Case II

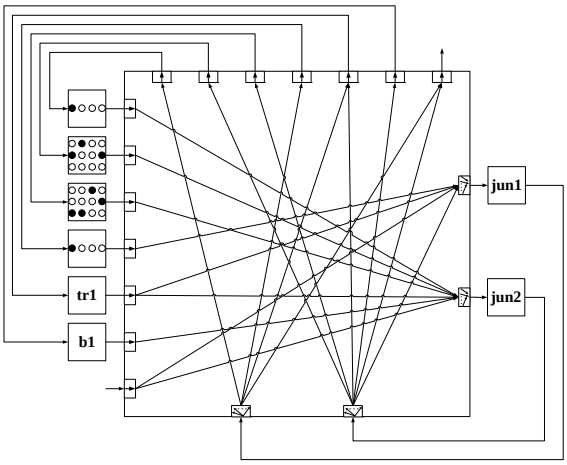


Figure.5. Optimal network configuration for Case II

5. Conclusions

Two superstructures have been presented in this study to formulate a MINLP model for one step design of batch WAN. The advantage of this approach is that not only all possible alternative network topologies can be captured, but the sequential order of all operations as well, can also be properly addressed. To ensure the quality and efficiency of the solutions, a hybrid optimization strategy integrating DICOPT and EA search techniques is developed. Finally, two examples dealt with single- and multi-contaminant WAN are presented to demonstrate the feasibility and effectiveness of the proposed automated design method.

Acknowledgments

This work is supported by the National Natural Science Foundation of China under Grant No.20876020.

References

Bagajewicz, M., & Manousiouthakis, V. (1992). On the mass/heat exchanger network representations of distillation networks. AIChE Journal, 38, 1769.

Cheng, K. F., & Chang, C. T. (2007). Integrated Water Network Designs for Batch Processes. Ind. Eng. Chem. Res., 46, 1241.

Kondili, E., Pantelides, C. C., & Sargent, R. W. H. (1993). A general algorithm for short-term scheduling of batch operations—I. MILP formulation. Comput. Chem. Eng., 17, 211.

Smith, E. M. (1996). On the optimal design of continuous processes, Ph.D. Dissertation, under supervision of C. Pantelides. Imperial College of Science, Technology and Medicine, London, UK.

Zhou, R.J., Li, L.J.,Xiao, W. & Dong H.G. (2008) Simultaneous optimization of batch process schedules and water-allocation network. Comput. Chem. Eng. (In press)

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
2009 Elsevier B.V. All rights reserved

Development of a Flowsheet Design Framework of Multi-step PSA Cycles for CO_2 Capture

Giovanna Fiandaca[a], Eric S Fraga[a], Stefano Brandani[b]
[a] *Centre for Process Systems Engineering,Department of Chemical Engineering, University College London (UCL), e.fraga@ucl.ac.uk*
[b] *Institute for Materials and Processes, School of Engineering and Electronics University of Edinburgh*

Abstract
A design framework for the evaluation of multi-bed/multi-step Pressure Swing Adsorption (PSA) cycles for CO_2 capture is introduced. The framework addresses the need to have a compact and effective representation of cycle configurations. This is necessary for the development and implementation of a flowsheet design framework for the synthesis of PSA cycles. Cycle configurations can be assessed through multi-criteria optimisation to take into account the trade-offs between conflicting aspect of the performance. The effect of cycle configuration has been investigated for 5 multi-step/multi-bed PSA cycles from the literature.

Keywords: Pressure Swing Adsorption, CO_2 Capture, Design, Multi-objective Optimisation

1. Design of Pressure Swing Adsorption Cycles for CO_2 Capture

The need to stabilise levels of greenhouse gases in the atmosphere motivates the development of more efficient CO_2 capture technologies. Pressure swing adsorption (PSA) is the most efficient separation option for middle scale operations [1]. PSA is a cyclic separation process whose main steps are adsorption, at high pressure (P_H), and regeneration of the adsorbent, at low pressure (P_L). During the adsorption step, the less strongly (or less rapidly) adsorbed species pass through the bed and are collected as a *light* product (LP). The *heavy* product (HP) consists of the adsorbed species removed during desorption step. Besides these basic steps, others have been introduced to improve the efficiency of separation. Originally, PSA was designed to provide a high purity stream of the light product (LP). However, CO_2 is the most preferably adsorbed species on the majority of adsorbents used. Advances in adsorbent materials and more complex cycle configurations have proven PSA to be efficient and economically viable for CO_2 capture [2]. However, the design of complex PSA cycles still remains mainly an experimental effort due to the complexity of the simulation. The need to develop automated tools for the synthesis of optimal PSA cycles has been noted [3, 4].

This paper describes the key elements for the development of a multi-criteria design framework for the generation and evaluation of multi-bed/multi-step PSA/VSA (vacuum swing adsorption) cycles for CO_2 capture. Specifically, we describe a compact representation of cycle configurations and the use of multi-objective optimisation to compare the performance of different cycles. A compact representation of the cycles enables the use of the cycle configuration as a design variable in an optimisation framework. The ability of the framework to distinguish between different cycle configurations is demonstrated for 5 multi-bed/multi-step vacuum swing adsorption (VSA) cycles, originally proposed in [5]. A VSA cycle is a PSA cycle where the low pressure achieved during the blowdown step is sub-atmospheric to allow a deeper regeneration of the adsorbent.

2. Compact Cycle Representation

A flowsheet design framework should be able to identify the best cycle configuration to perform the desired separation according to the criteria specified by the designer. In Fig.1, a flowchart of the framework is proposed. The first step is the generation of a cycle configuration, c, which satisfies the requirements of the designer. Then, the set of operating conditions, $\overline{x}$, which affect the performance for this configuration is identified. Potential operating variables are cycle time, dimension of the beds, feed rates. The actual set of design variables, and their bounds, will depend on the configuration generated. Next, the operating variables, $\overline{x}$, for cycle c are optimised using multiple objectives. The performance of the configuration is then fed back to the generator of cycles and taken into account when generating new configurations. The information the designer needs to provide are the problem of interest (e.g. mixture of gas to be separated, the desired products) and the physical properties needed to simulate the operation. Furthermore, the designer has to give indication about the design space to be investigated, regarding both the relevant cycle configurations and the operative conditions.

The performance of the cycle, for instance in terms of purity and recovery, is defined by the cycle configuration (sequence **S** of steps used, bed interconnections) and operating variables, but not by the number of beds (N_B) used to perform the cycle: all the beds will undergo the given sequence of steps in an identical manner and at cyclic steady state they all behave in an identical way. On the other hand, the number of beds is necessary to evaluate capital and operating costs of the cycle. Once the sequence of steps **S** which provides the desired product has been chosen, the minimum number of beds needed can be determined using the method proposed by Smith and Westerberg in [3]. The cost and the number of beds, should be taken into account in the evaluation of cycle c.

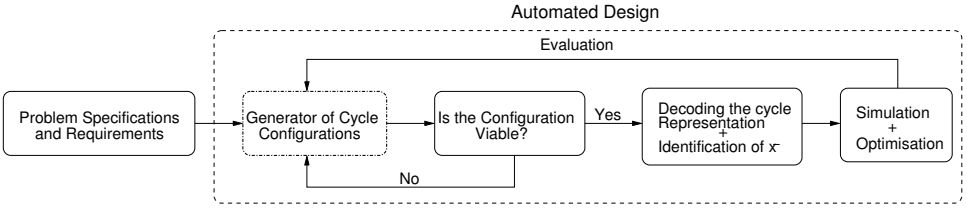


Figure 1: Diagram representing the elements of the ideal automated design framework for PSA cycles. According to the specification introduced by the designer, a viable cycle configuration, c, is generated. Then, the operating variables, $\overline{x}$, which determine the performance of c are identified. The optimal performance of c, in the design space defined by $\overline{x}$, is determined. The performance criteria are then fed back to the cycle generator so as to take it into account when generating new cycles.

The essential elements of a cycle configuration are the sequence **S** and the number (N_S) of steps, the number of products (N_P), the map of interconnections (between beds, feed tanks and product tanks), the pressure profile and the direction of the flow (co-current or counter-current with respect to the adsorption step) in each step.

One of the most critical decisions when developing automated design tools for a particular class of problems is the choice of the search space, or, equivalently, of a representation [6]. The representation needs to contain enough details for the implementation of the simulation and to cover the desired design space; it also has to be managed with ease by the generator of configurations. We propose, therefore, a compact mathematical representation of cycle configurations. The key idea motivating the representation we propose, is that, to synthesise new, more efficient, cycles, we need to think of cycles not as sequences of known steps (e.g. adsorption with feed, desorption etc.), but to allow for all the possible bed interconnections to be explored. Since all the beds undergo the same sequence, **S**, we can describe the whole cycle using only one bed. The connections implemented in a cycle can

be represented in a matrix, C, which has to specify, at each step, the step or steps from which the bed is receiving its inlet; it also has to indicate the steps from which products are withdrawn. Each row of C corresponds to the "receiving step", including product withdrawals, and each column to its "sources". The bed can act as a source at each step, as well as the feed tank. The number of receiving units (N_U) is therefore the summation of the number of steps (N_S) and the number of products (N_P). The number of sources (N_{in}) is given by the number of steps plus 1, the feed. If N_C is the maximum number of units feeding a single unit, the dimension of the matrix C will be $N_U \times N_C$. The representations of the cycles investigated in this paper are presented, in this compact form, in Table 1, together with the graphical visualisation of the cycles.

The matrix C describes in a compact way not only the map of interconnections in the cycle, but also information about pressure profiles and position of mixers and splitters. Once the output/input streams are specified, also the pressure profiles are: if both an input and an output streams are present, the pressure will remain constant (as in adsorption and desorption); if no input is fed to the bed, then the pressure will go down as the bed releases its content (as in the blowdown step); if no output stream leaves the bed, the pressure will increase during the step (as during pressurisation). The simulator can also deduce from the information in C when to include mixers (if a unit is receiving more than one stream, the incoming streams need to be mixed before been fed) and splitters (if a stream, or a mixture of streams, feeds more than one unit, splitters are needed). The information that the generator of configurations needs to produce are N_S, N_P, the matrix C and a vector $\overline{f}$ describing the direction of the flow for each step. $\overline{f}$ can be a vector of N_S elements, where $\overline{f}_i$ is 1 if the flow in step i is cocurrent with respect to the adsorption step, -1 otherwise.

The representation is general enough to represent all viable cycle configurations. However, it may also represent cycles that are not meaningful. It is therefore necessary for the designer to include some rules to guarantee that the configurations generated are viable. For example, one might want to specify the maximum number of steps that can be included, or that the first step must receive from the feed by imposing $C(1,1)$=-1, and the last step (N_S) is a pressurisation by specifying that $\forall(i,j) \in \{1,..,N_U\}\ C(i,j) \in \{-1,1,...N_S-1)\}$. Section 3.1 illustrates the rules used to generate the cycles investigated in this article. Besides imposing the viability of the cycles generated, the designer might want to indicate which interconnections need to be included in the cycles and which are to be avoided: for example, pressure equalization steps are usually not included when vacuum swing cycles are used, as no compression work has to be recovered. This can easily be imposed to the matrix C, requiring that a bed which is undergoing blowdown cannot feed a bed which is undergoing pressurisation. All of the above specification can be implemented by imposing some conditions on the generation of C. However, the rules should not be overspecifying the problem so that only known configurations are explored.

3. Case Study

We present the effect of cycle configuration on four 5-step cycles, originally proposed in [5], for the separation of CO_2 from flue gases. We also include the 4 bed/4-step cycle [2, 7]. All the cycles are illustrated in Table 1. The feed is a typical stack effluent at 575 K, containing 15% of CO_2, 75% N_2 and 10% of H_2O. The adsorbent used is a K-promoted hydrotalcite-like compounds (K-HTlcs). The performance of a PSA cycle needs to be evaluated when cyclic steady state has been reached. This can be evaluated via the dynamic simulation of the operation, which is described by a set of differential equations.

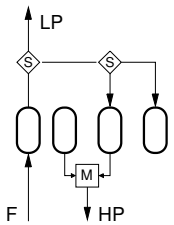
A simplified model, based on a series of CSTRs (continuous stirred tank reactors) for each bed, has been used to reduce computational costs involved with the simulation.

Table 1: Illustration of all the cycles investigated in this study. In the second column each cycle is represented in a compact way, in the second column the schematics of the cycle. The feed is indicated by "-1". The "-" entries in C corresponds to a 0 value in the actual computer representation.

Cycle	Representation	Scheme

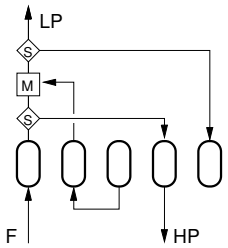
a

Destination	**Sources**	
I	-1	-
II	-	-
III	1	-
IV	1	-
LP	1	-
HP	2	3



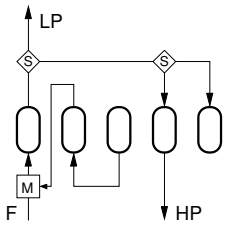
b

Destination	**Sources**	
I	-1	-
II	3	-
III	-	-
IV	1	-
V	1	2
LP	1	2
HP	4	-



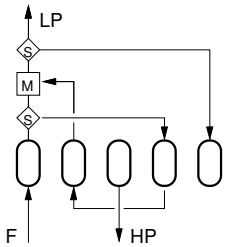
c

Destination	**Sources**	
I	-1	2
II	3	-
III	-	-
IV	1	-
V	1	-
LP	1	-
HP	4	-



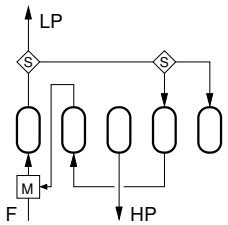
d

Destination	**Sources**	
I	-1	-
II	4	-
III	-	-
IV	1	-
V	1	2
LP	1	2
HP	3	-



e

Destination	**Sources**	
I	-1	2
II	4	-
III	-	-
IV	1	-
V	1	-
LP	1	-
HP	3	-



A good correspondence between our results and those of the more rigorous model used by Reynolds et al. [2] has been achieved using 6 CSTRs to simulate each bed [7]. The linear driving force (LDF) model has been used to describe the mass transfer in the solid phase [1]. All data used are from [5]. The details of the equations used have been provided in [7]. We achieved a faster convergence to cyclic steady state (CSS) by adopting the unibed approach [8].

3.1. Generation of configurations

In all cycles investigated here, the first step is adsorption with the feed, and that the last step is a pressurisation with the light product. The second step is *heavy reflux*, where the bed, at high pressure, receives a purge gas, rich in the heavy component, to the feed end. This heavy purge gas can come from step *III*, which is always a blowdown, or from step *IV*. Step *IV* is a *light reflux* step, during which the CO_2 is desorbed at low pressure using the output stream of step *I* as purge gas. The output of step *II* can either be recycled at the feed end of step *I*, or might be taken as a part of the light product and used to pressurise step *V*. The heavy product is withdrawn from the step, between *III* and *IV*, which is not feeding step *II*. All the steps are counter-current with respect to the first.

4. Results

The design problem is the simultaneous optimisation of CO_2 recovery (R) and purity (P) of the product in the design space defined by the cycle time (t_c), the purge-to-feed ratio (γ), the pressure ratio ($\pi_T = P_H/P_L$) and the throughput of the operation (θ). θ is the amount of feed fed to one reactor during the adsorption step (step *I*) divided by the cycle time and the mass of adsorbent in one column. The analysis of the cycles has been made with respect to the same parameters used in [2, 5]. The value of the high pressure has been kept constant, $P_H = 1.36\,\text{atm}$. The only constraint is the evaluation of R and P at cyclic steady state, which is identified by dynamic simulation.

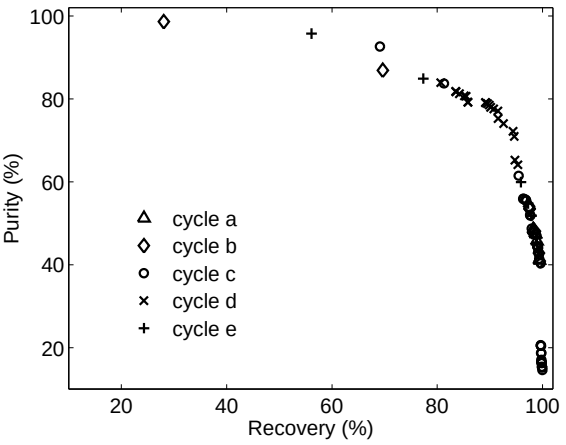


Figure 2: The Pareto set obtained by identifying the set of non-dominated points from the union of the individual Pareto sets obtained for each cycle configuration investigated. The aim is to show how the best performance of different cycles compare.

For the multi-objective optimisation, a multi-objective genetic algorithm (MOGA) has been used [7, 9]. The Pareto sets, consisting of equally optimal trade-offs between P and R, have been obtained using a crossover rate of 0.7, a mutation rate 0.01, a tournament size 2, a population size of 70, and 50 generations. Using a large population size, we obtain a reliable approximation to the Pareto front, as the statistical analysis carried out in [9] showed.

A Pareto set of each cycle configuration has been generated. The overall set of non-dominated points, from the set of points defined by the union of the individual Pareto sets, is shown in fig.2. The cycle configuration which corresponds to each point in this set is shown in the legend. Our results are qualitatively coherent with the conclusions by Reynolds et al. [5]. However, we did not replicate their

results quantitatively as our models differ. All 5-step configurations outperformed the 4-step cycle. The latter could only achieve a product purity of 60%. This result justifies the investigation of cycles where also the heavy recovery step is included (step *II* in all the 5-step cycles).

Comparing the results from cycle *b* and cycle *c*, the effect of the management of the output of the heavy recovery step (step *II*) can be deduced: when this is mixed with the output of step *I*, as in cycle *b*, lower recoveries are achieved as some CO_2 can be lost in the light product. The higher values of P have been achieved by cycle *b*. By comparing cycle *b* and cycle *c* on one side, and cycle *d* and cycle *e* on the other, we can understand the effect of the different arrangements for the input to step *II* and the heavy product. When the purge for the heavy recovery step comes from the blowdown (as in cycle *b* and cycle *c*), higher purity and higher recoveries can be achieved, as it becomes evident looking at both extremes of the Pareto set. However, the Pareto set suggests that the best trade-offs between P and R are achieved by cycle *d*, which dominates the region where the Pareto set bends. This analysis would encourage the cycle-generator in the design framework to propose new cycles which keep the characteristics of cycle *d* and cycle *b*, maintaining what they have in common (input to step *I* and *II*), and proposing intermediate arrangements of the input to step *II* and the heavy product, in which the two cycles differ.

5. Conclusion

A multi-criteria design framework for the generation and evaluation of multi-bed/multi-step PSA cycles for CO_2 capture is proposed. A compact representation of the cycles has been introduced, a necessary component for the development of the flowsheet design framework. The representation has been illustrated through the presentation of 5 multi-step/multi-bed cycles. These 5 cycles have been solved using a multi-objective optimisation tool. An overall Pareto set has been generated which allows us to understand the effect of different cycle configurations as well as to gain insight on how to improve the performance by manipulating the configuration.

References

1. Ruthven, D. M., Farooq, S., and Knaebel, K. S. (1993) *Pressure Swing Adsorption*. VCH.
2. Reynolds, S. P., Ebner, A. D., and Ritter, J. A. (2005) New pressure swing adsorption cycles for carbon dioxide sequestration. *Adsorption*, **11**, 531–536.
3. Smith, O. J. and Westerberg, A. W. (1990) Mixed-integer programming for pressure swing adsorption cycle scheduling. *Chem. Eng. Science*, **45**, 2833–2842.
4. Agarwal, A., Biegler, L. T., and Zitney, S. E. Optimal synthesis of a pressure swing adsorption process for CO_2 capture. *08AIChE Annual Meeting*.
5. Reynolds, S. P., Mehrotra, A., Ebner, A. D., and Ritter, J. A. (2008) Heavy reflux psa cycles for CO_2 recovery from flue gas: Part I. performance evaluation. *Adsorption*, **14**, 399–413.
6. Hamda, H., Roudenko, O., and Schoenauer, M. (2002) Multi-objective evolutionary topological optimum design. *Adaptive Computing in Design and Manufacture V*.
7. Fiandaca, G., Fraga, E. S., and Brandani, S. (2009) Multicriteria design framework for CO_2 capture via multi-step PSA cycles. *European Symposium on Computer-Aided Process Engineering*-19, in press.
8. Kumar, R., Fox, V. G., Hartzog, D. G., Larson, R. E., Chen, Y. C., Houghton, P. A., and Naheiri, T. (1994) A versatile process simulator for adsorptive separations. *Chem. Eng. Science*, **49**, 3115–3125.
9. Fiandaca, G., Fraga, E. S., and Brandani, S. (2008) A multi-objective genetic algorithm for the design of pressure swing adsorption. *In Proceedings of ACDM 2008,* 8^{th} *International Conference*, ISBN: 978-0-9552885-1-7.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Ethylene Separation by Feed-Splitting from Light Gases

Daniel Salerno[a], Harvey Arellano-Garcia[a], Günter Wozny[a]

[a] *Chair of Process Dynamics and Operations, Sekr, KWT-9, Berlin Institute of Technology, Strasse des 17 Juni135, D-10623, Berlin, Germany*

Abstract

Separation of ethylene from light gas mixture is one of the most energy intensive separations in petrochemical processes, which uses distillation columns up to 100 m tall and containing over 100 trays due the very small differences in the relative volatilities and very large reflux ratios and also due to the need for sub-ambient temperatures. In recognition of these costs, several attempts have been made in the past to develop processes with less energy and equipment costs. In distillation columns with condenser temperatures significantly below room temperature, such as in ethylene separation towers, it is essential to minimize the expensive energy requirements of the refrigeration cycle that produces the tower reflux. In this work, a solution has been found by expanding the gaseous distillate to decrease its temperature. Moreover, additional solutions applied to conventional ethylene fractionation columns have been implemented here in order to study this behavior. In this contribution, an outlet stream of an Oxidative Coupling of Methane (OCM) reactor, which has been previously stripped of its CO_2 content, is also introduced in a demethanizer tower to remove almost all of its CH_4 content before entering the ethylene fractionating column. Then, it is cooled exchanging its heat with the distillate stream of the ethylene tower, warming the distillate. The main goal is to reduce the condenser heat duty. This objective is achieved reducing a significant amount of heat required by the condenser, maintaining the mandatory product purity. Due to this improvement, it was also possible to reduce the reboiler heat in almost the same percentage amount that is achieved with the condenser. In addition, the reflux of ethylene column decreases. The sensitivity analysis and the corresponding simulations results will be discussed in order to show the efficiency of the presented approach. These results have also been used for the design of the pilot plant which is now being built at our department. The final design for the OCM process will also be presented.

Keywords: distillation, ethylene separation, energy saving, process simulation.

1. Introduction

Due to the large natural gas reserves worldwide, methane appears to be the cheapest and most available carbon source for the gas-based petrochemical industry. The development of a simple and commercially advantageous process for the direct conversion of natural gas to more easily transportable products was desired. Catalytic reaction for the direct conversion of natural gas to ethane and ethylene offered a new route for ethylene production. Among the numerous attempts for direct conversion, the oxidative coupling of methane (OCM) to C_2 hydrocarbons still remains one of the potential routes.[1-2] Regardless of the technology used in the production of ethylene, the process that follows its production is to separate the reaction products.

Distillation is by far the most common separation technique in the chemical process industry accounting for 90% to 95% of the separations.[3] Distillation columns are a major energy user in a petrochemical plants, and therefore, it is necessary to save as much as possible the energy demand for this towers, specially the heat expenses used in condensers and reboilers, because of the evaporation steps involved. Typically, more than half of the process heat distributed to plant operations ends up in the reboilers of distillation columns. By this, high-level energy is fed at the base of the column and about the same amount of energy is released at the top, unfortunately at a much lower temperature level. The difference between the two Gibbs energies can be seen as the necessary energy investment to reverse the mixing entropy and to separate the components of a given feed by a distillation process.[4] To find a solution to this problem some researchers have proposed alternatives focused on the feed-splitting concept.[5] In the conventional distillation operation for ethylene-ethane separation, as typified in Figure 1, low-temperature distillation column has been the preferred technology for several decades. However, this binary separation process consumes about 36 percent of the refrigeration energy required in, for example, an ethylene plant.[6]

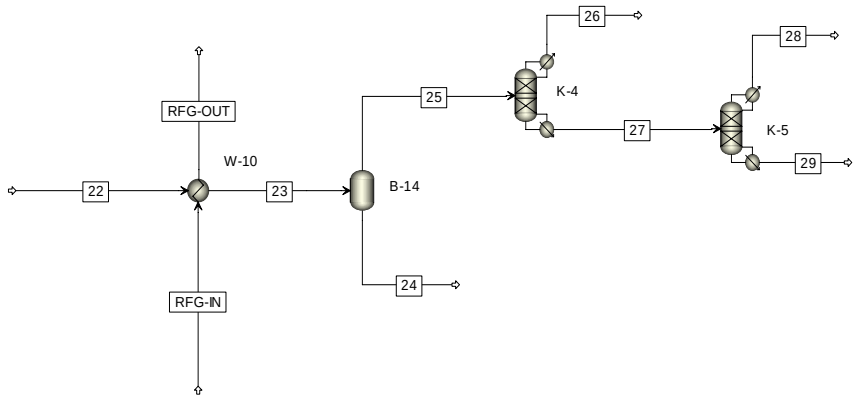


Figure 1 Conventional Ethylene Separation System

The C_2 splitter is commonly operated at high-pressure, utilizing closed-cycle propylene refrigeration, which is incorporated with the refrigeration systems serving other parts of the plant. The desired objectives for any ethylene separation process are to obtain a high-purity ethylene product combined with a high percentage recovery of the ethylene. The conventional distillation technology can accomplish both of these objectives. The main goal of this study is to reduce de amount of energy required by the reboiler in the ethylene-ethane column by partially precooling the feed(s) by the distillate, using a scheme symmetrical to that employed for warm columns.

2. Brief description of the OCM process

The complete process consists of three main sections: reaction, purification and separation section. Due to novel process design strategies, the downstream alternatives will also improve the reaction part and the catalyst design as well by means of for instance maximum carbon dioxide concentration for the reaction product. Figure 2 shows the flowsheet for this process.

2.1. Reaction section

The reactor is continuously fed with natural gas and oxygen. The feed gas is preheated to 700 °C and will catalytic partial oxidized at pressure of 0.5 MPa. The Reactor is jacked cooled to operating temperature of 750 ° C. The exothermic reaction heat has to be removed while steam is produced. The reaction products are compressed in a multi-compaction section to 3.5 MPa and cooled down to 40 °C later on. This stream is fed to the following purification section.

2.2. Purification section

In this section the carbon dioxide is removed by an amine scrubber and the following sodium hydroxide scrubber. The first scrubber is based on an absorption process with mono ethanolamine solutions (MEA) as detergents. The concentration of carbon dioxide is reduced from approximately 60 mol% to 10 mol% in this section. In the next following scrubber the carbon dioxide will be removed completely.

2.3. Separation section

This section consists of two cryogenic distillation columns. The first one is the demethanizer and the unconsumed methane is separated from the product stream. The product stream consists of ethane and ethylene that is separated in the deethanizer. The heat requirement Q of a single rectification column can be minimized best by effective heat exchange between the warm products and the cold feed, and also by efficient insulation of the column. Furthermore, the operational reflux ratio R should be kept as small as possible; optimum ratios are in the range of $R = 1.05 - 1.1\ R\text{min}$.[7] Soave has shown that in a cold tower, if the enthalpy of the feed is decreased by precooling with the distillate, the final result is a decrease in the condenser and reboiler duties; his study was applied to a demethanizer column of a "cold box" of an ethylene plant. In the present work the feed-splitting principle to an ethylene-ethane distillation tower is applied, using the steady-state Aspen Plus® simulation tool for the calculation and study of this ethylene column.

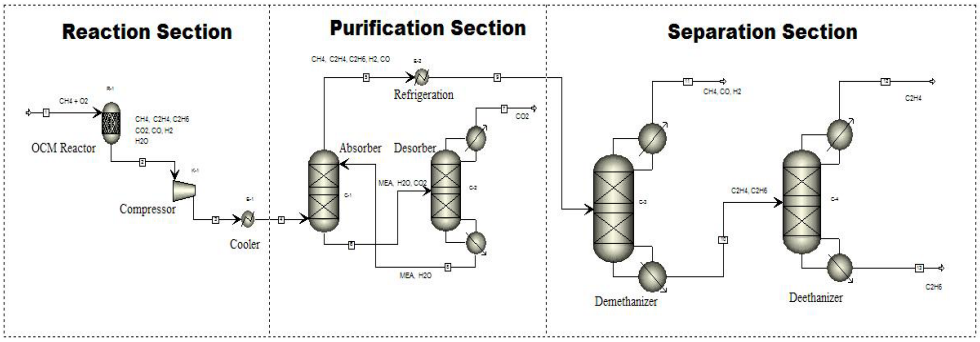


Figure 2 Flowsheet of the OCM Process

3. Simulation models

The first simulation model had the same schematic shown in figure 1. The data on the feed to the demethanizer column are reported in Table 1. Starting point is the outlet stream of an OCM reactor, in which all the contents of CO_2 are removed, introduced in a demethanizer tower to remove most of its CH_4 content before entering to the ethylene-ethane fractionating column. Both columns were modeled using rate-based calculations. With this approach, mass and heat transfer between the phases flowing each other is

assumed, based on the two-film theory (AspenONE Process Engineering webinar, 2008). Three alternative solutions of feed configurations are studied, and each one has to guarantee a distillate with at least 95% weight of ethylene.

Table 1 Demethanizer Column Feed

Property	Value	Flow rate (kg/hr)	
T (°C)	-23.000	H_2	9.870
P (bar)	32.000	CO	35.160
F (kg/hr)	442.319	CH_4	330.010
		C_2H_4	51.040
		C_2H_6	16.220
		H_2O	0.019

In order to reduce the condenser and reboiler duty different solutions of feed configuration to the ethylene-ethane column were proposed for improving its performance with the feed-splitting technique. In the first one (Figure 3) the feed is pre-cooled with the cool stream from the condenser, were a warmer distillate is obtained. The feed temperature to the ethylene-ethane column decreased, the minimum reflux ratio and the condenser duty also decreases.

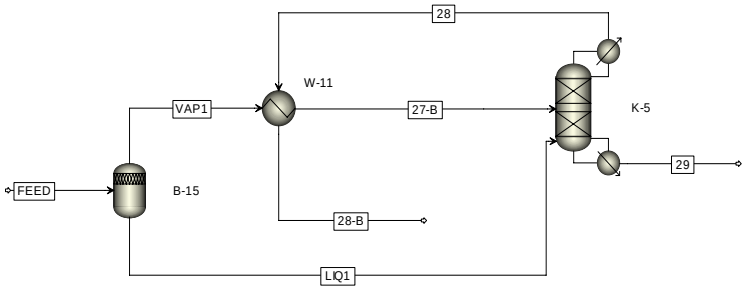


Figure 3 Alternative 1 for the Ethylene-Ethane Column

In the second one (Figure 4), part of the feed to the ethylene-ethane column is flashed, at constant pressure lowering its temperature, and consequently vaporizing certain amount of the stream. The vapor phase goes to the heat exchanger, decreasing its temperature a little more, and then is adiabatically flashed again to generate a second liquid stream which feeds the column K-5.

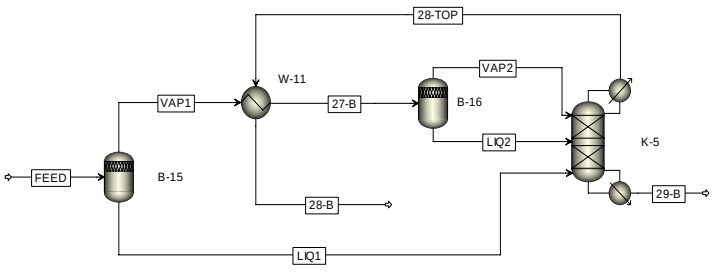


Figure 4 Alternative 2 for the Ethylene-Ethane Column

In the last alternative (Figure 5) only a fraction of the vapor phase is precooled. This fraction has to be determined by sensitivity analysis, making sure to maintain the quality specifications of the desired ethylene product at the top of this column.

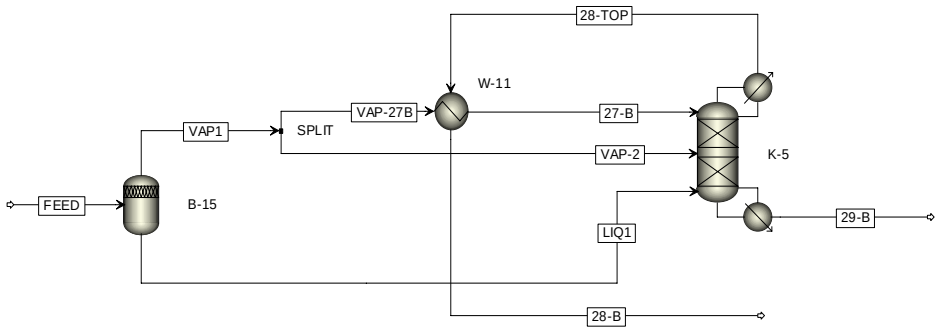


Figure 5 Alternative 3 for the Ethylene-Ethane Column

In all three cases the liquid phase obtained from the first flash is fed directly to the column, taking into account that this phase has a flow rate that is too low for good heat-exchange performance with the distillate. Also with these three studied configurations it was necessary to find the best feed trays. This can be done by changing the position of each feed stream one at a time. For all the simulations of the precooler the hot/cold outlet temperature approach was also kept constant.

4. Results and Discussions

To properly evaluate the simulation results the comparisons were made with the conventional separation ethylene-ethane column (K-5 from figure 1). Table 2 shows the duty values for condenser, reboiler, and flow rate of ethylene at the top of this column. Table 3 shows the ethylene recovery results for all the four cases studied.

Table 2 Ethylene-Ethane Column Results

Case Studied	Heat Duty (kW)		C_2H_4 flow (kg/hr)
	Condenser	Reboiler	
Conventional	-127.962	131.523	47.763
Alternative 1	-96.647	100.528	48.401
Alternative 2	-96.262	95.702	49.637
Alternative 3	-96.630	96.084	48.455

Table 3 Ethylene Recovery Results

Case Studied	Ethylene Flow (kg/h)		
	Total Column Feed	Top	Recovery
Conventional	50.266	47.763	0.9502
Alternative 1	50.464	48.401	0.9591
Alternative 2	50.463	49.637	0.9836
Alternative 3	50.464	48.455	0.9602

Comparing the conventional and the alternative 1 case, there is a substantial decrease in the condenser and boiler duties, around 30%, while at the same time the C_2H_4 product increases. Further improvements, alternative 2, lead to a greater ethylene recovery, small decrease in condenser duty and more substantial decrease in the reboiler duty (around 5%). When applying the feed-splitting concept to the ethylene-ethane column (K-5), both the condenser duty and the reboiler duty decrease under the same operating conditions (pressure, number of theoretical stages of the column, internals type, feed flow rate and composition). It can be observed that the condenser values decreases with each improved alternative, until a further splitting of the vapor phase can't be reached. In alternative 3, the 93.4% of the feed to the column was flashed into vapor phase being subjected to precooling. Comparing the results shown in Table 3 we can see the improvement in the flow rate of ethylene product at the top of the column K-5, therefore, alternative 2 shows the best performance in case of ethylene recovery.

5. Conclusions

The goal of significantly reduce the heat duty required by the condenser has been achieved by almost 25% and simultaneously has been reduced (27.2%) the amount of heat duty required by the reboiler. If the flow to the W-11 heat exchanger is not divided, there will be an increase of the ethylene recovery at the top of the K-5 tower. With this feed-splitting concept 98% of the initially ethylene can be recovered. Considering the data reported in Table 2, the advantages of applying the feed-splitting concept in a pilot plant is highly profitable, in order to save as much energy as possibly and reduce the expenses and optimizing resources. The sensitivity analysis and the corresponding simulations results show the efficiency of the presented approach. These results will be used also for the design of the pilot plant which is now being built at our department using the results of these simulation specifications.

Acknowledgment

The authors acknowledge support from the Cluster of Excellency "Unifying Concepts in Catalysis" coordinated by the Berlin Institute of Technology and funded by the German Research Foundation – Deutsche Forschungsgemeinschaft.

References

1. S. Bhatia; C.Y. Thien; A.R. Mohamed, 2009, Oxidative coupling of methane (OCM) in a catalytic membrane reactor and comparison of its performance with other catalytic reactors, Chemical Engineering Journal, doi:10.1016/j.cej.2009.01.008.
2. N.A.S. Amin; S.E. Pheng, 2006, Influence of process variables and optimization of ethylene yield in oxidative coupling of methane over Li/MgO catalyst, Chemical Engineering Journal, 116, 187-195.
3. P.C. Wankat, 2007, Separation Process Engineering.
4. G. Soave; J.A. Feliu, 2002, Saving energy in distillation towers by feed-splitting. Appl. Therm. Eng., 22, 889-896.
5. G. Soave; S. Gamba; L. A. Pellegrini; S. Bonomi, 2006, Feed-Splitting Technique in Cryogenic Distillation. Ind. Eng. Chem. Res., 45, 5761-5765.
6. J.C. Bruno; A. Vidal; A. Coronas, 2006, Improvement of the raw gas drying process in olefin plants using an absorption cooling system driven by quench oil waste heat, Energy Conversion and Managemet, 47, 97-113.
7. J. Stichlmair, 2002, Distillation and Rectification, Ullmann´s Encyclopedia of Industrial Chemistry.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Cost minimization in noncatalytic biodiesel production plants

Susana Espinosa[b], M. Soledad Diaz[a] and Esteban A. Brignole[a]

[a] *Planta Piloto de Ingeniería Química (PLAPIQUI), Universidad Nacional del Sur-CONICET, Camino La Carrindanga Km 7, Bahía Blanca 8000, Argentina*
[b]*Universidad Nacional del Comahue,Buenos Aires 1400, 8000 Neuquén, Argentina*

Abstract

In the present study, we propose an optimization model for the minimization of capital and operating costs in a supercritical methanol biodiesel plant, considering alternative process schemes together with the convenience or not to use a cosolvent in the chemical reaction. A rigorous process simulation model, based on a group contribution equation of state is integrated with a successive quadratic programming algorithm to solve the nonlinear problem. Optimization results show that noncatalytic biodiesel production with supercritical methanol can be economically competitive with conventional catalyzed production.

Keywords: biodiesel, supercritical, transesterification, optimization.

1. Introduction

Biodiesel is comprised of esters of short chain alcohols made from renewable biological sources, such as vegetable oils, and it can be used as an alternative diesel fuel. Biodiesel can considerably reduce emissions, it is biodegradable and nontoxic. However, raw material and production costs are the main concern for commercialization of this product. Most of the conventional methods for biodiesel production use a basic or acid catalyst (Freedman et al., 1984, Fukuda et al., 2001, Lotero et al., 2005) and more recently, enzimes (Watanabe et al., 2000; Varma and Madras, 2007). While conventional catalytic transesterification of oils gives several byproducts, the single phase supercritical transesterification takes place at high pressures and temperature, not requiring the use of a catalyst. This feature makes the process not sensitive to water and free fatty acid content in the feed, avoiding the need for a pre-treatment step (Kusdiana and Saka, 2004; Cao et al., 2005) and additional separation steps downstream the reactor. Furthermore, free fatty acids in the feed oil are esterified simultaneously. Reaction time takes seconds to minutes in supercritical processes while catalyzed transesterification takes minutes to hours. The absence of pretreatment, soap removal and catalyst removal can significantly reduce the capital cost of a biodiesel plant. However, the need for high pressures and temperatures can be associated to high operating costs for the supercritical process. Therefore, development of reliable simulation and optimization models is required for the design of economically feasible supercritical biodiesel production processes.

In this work, we develop an optimization model for the minimization of capital and operating costs in a supercritical methanol biodiesel plant, based on rigorous process models, nonlinear programming techniques and thermodynamic predictions with a Group Contribution Equation of State with Association (Gross., 1996; Ferreira et al., 2004). Thermodynamic predictions have in turn been checked with available

experimental equilibrium data from the literature and data obtained in our group laboratory (Hegel et al., 2007). We have formulated a nonlinear programming (NLP) problem for an 80,000 ton/y biodiesel plant, using soybean oil and methanol as raw materials. The model allows exploring different plant capacities and alternative productions schemes by solving different NLPs, together with the convenience of including or not a co-solvent in the chemical reaction. Numerical results confirm that non-catalytic biodiesel production with supercritical methanol is economically competitive with conventional catalyzed production.

2. Thermodynamic modeling of supercritical soybean oil transesterification

Thermodynamic modeling of phase equilibrium of the reactants and reaction products in the two phase region and the determination of the homogeneous region are required for a reliable reactor design. Furthermore, the design of the downstream separation train requires the accurate prediction of the distribution of components between phases in the heterogeneous region (Espinosa et al., 2007, Diaz et al., 2008).

Binary and ternary systems including reactants and reaction products and water have been studied at different operating conditions. Binary interaction parameters have been taken from the literature (Espinosa et al. 2000; Ferreira et al., 2004; Andreatta et al., 2008). Figure 1 shows a good agreement between the Group Contribution Equation of State (GCA-EoS) estimations when compared to experimental data on LLE of the ternary methanol + methyl oleate + glycerol by Negi et al. (2006) at 333 K and atmospheric pressure. Reliable phase equilibria predictions of this mixture are especially important to identify operating conditions in the first separator that assure high recovery and purity of methanol in the vapor phase. Our study concludes that three phases in equilibrium (two liquid phases and one vapor phase) are necessary in order to obtain methanol rich vapor phase while glycerol and methyl esters remain in the liquid phases. The existence of two liquid phases in the first separator has been impose as a restriction in the mathematical model.

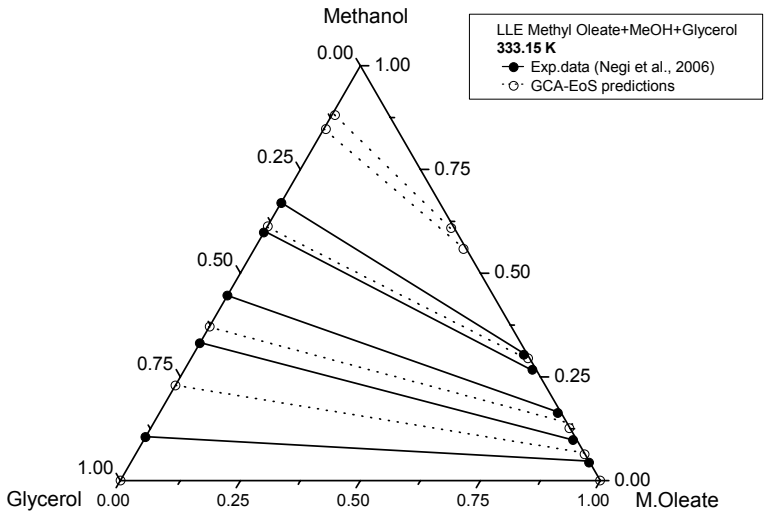


Figure 1. Ternary phase equilibria for the reaction products (methyl oleate and glycerol) and excess methanol at 333.15 K and one bar. Experimental data (Negi et al., 2006) and GCA-EoS predictions

3. Optimization Model

The optimization model for the noncatalytic transesterification of soybean oil with methanol at high pressure includes reliable thermodynamic predictions with the GCA-EoS equation of state and alternative schemes indicated as discontinuous lines in Fig. 2.

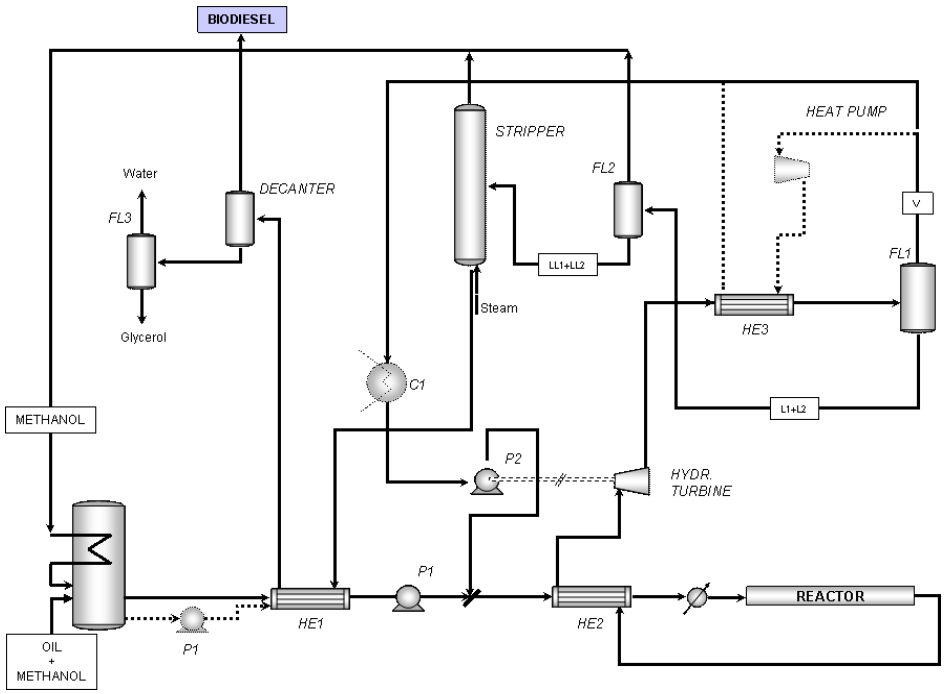


Figure 2. Superstructure of supercritical biodiesel production plant

We have formulated a nonlinear programming (NLP) problem for each analyzed process scheme, for an 80,000 ton/y biodiesel plant. The objective is the minimization of capital and operating costs for a supercritical methanol plant considering a project life of ten years. Capital cost correlations have been implemented from Ulrich (1984), Peters and Timmerhaus (1991) and Institut Français du Pétrole (1981).

Operating costs correspond to steam and cooling water requirement in heat exchangers, and electrical consumption for pumps and compressors. Optimization variables are the operating pressure in the supercritical reactor and in the first methanol recovery unit (FL1), pressure change in the heat pump, inlet steam flowrate in the stripper and distillate to feed ratio in this unit. The process model includes mass and energy balances in equipment, design equations, thermodynamic equilibrium equations (GCA-EoS) and recycle convergence equations as nonlinear equalities. Inequality constraints impose the existence of two liquid phases in the 1st separator, aqueous and organic phases, to provide high methanol recovery in the recycle stream with low glycerol content. A minimum methanol recovery of 90% and maximum glycerol mass concentration of 1 % in this stream are required. Additionally, constraints on maximum water content in the recycled methanol from the top of the stripper, products purity specification of 99 wt% for methyl oleate (biodiesel) and product purity specification above 95wt% for glycerol are imposed. Product specifications, as required by European norms (EN 14214), have been also included as additional constraints in the optimization model. The resulting

nonlinear programming problem has been solved with a Successive Quadratic Programming (SQP) algorithm (Biegler et al., 1985).

4. Discussion of Results

Based on experimental results (Hegel et al., 2007) operating pressure in the reactor has been taken from 98 to 110 bar to ensure around 100% oil conversion to biodiesel, at 570 K. Following Kusdiana and Saka (2001) experimental results, the reaction has been modeled as first order and rate equations have been included in the optimization model. The supercritical transesterification reaction is carried out in a tubular reactor with 24 to 1 methanol/oil molar ratio. The possibility to use propane as a co-solvent (also to obtain 100% conversion at lower pressures), as suggested by Cao et al. (2005) has been considered, taking a molar propane/methanol ratio of 0.05 in the reactor. As compared to these authors' results, we have obtained a slightly higher residence time in the reactor (13.59 against 10 min) and slightly lower residence time with the process without co-solvent (13.59 against 15.37). Process variables, equipment energy consumption and biodiesel composition in the final product are shown in Table 1 for supercritical processes with and without co-solvent.

Table 1. Main variables of the supercritical biodiesel process

Variable	C3/Meth=0.00	C3/Meth=0.05
Separator Pressure, FL1 (bar)	6	6.5
Pres. change Heat Pump (bar)	3	9.5
Reactor pressure (bar)	100	106
Methanol recovery in recycle (%)	93	95
Glycerol in recycle (wt%)	0.3	0.6
Biodiesel in recycle (wt%)	0.2	0.3
Biodiesel purity (wt%)	99.9	99.8
Methanol in Biodiesel (wt%)	0.07	0.1
Glycerol in Biodiesel (ppm)	76	97
Water in Biodiesel (ppm)	473	395
Tubular Reactor Volume (m^3)	25	31
Reactor Tubes Length (m)	28	30
Reactor Residence Time (min)	15.4	13.6
Feed Pump Power, P1 (kW)	99.8	99.9
Compressor Power, HP* (kW)	116.6	337.0
Heat in Heat Pump (kW)	2362.3	2956.6
Hidraulic Turbine Power (kW)	69.9	75.2
Recycle Pump Power, HP[a] (kW)	61.7	69.8

[a] HP: Heat Pump

It can be noted that optimal operating pressure for the first separator is 6 bar, which ensures LLV equilibrium at the conditions determined in the heat pump (150ºC to

160ºC) and a pressure change of 3 bar has been determined in the heat pump. However, when propane is used as a co-solvent, this pressure change is not enough to completely condense the propane + methanol mixture. Therefore, one alternative is the addition of external cooling, as this mixture has its bubble point around 363 K, which is too low to heat the feed to the first separator. Another alternative is to increase the upper bound on pressure change in the heat pump, which reduces the latent heat of vaporization of the condensing methanol + propane mixture, avoiding the requirement of additional cooling. The propane/methanol mixture must be subcooled to avoid cavitations in the pump. Biodiesel purity is above 99%wt in both cases, as required. The feed pump and the compressor have the main energy consumption. The integration between the hydraulic turbine and the recycle pump has been possible in the analyzed cases.
Table 2 shows optimal cost values with and without propane used as co-solvent in a 0.05 molar ratio to methanol. In cost calculations, investment cost has been annualized considering a project life of ten years and soybean oil price as 60 US$/t. When operating with a co-solvent both operating and capital costs are higher due to the need for compression at higher pressure in the heat pump compressor of the methanol-propane recycle. However, as mentioned earlier, the reaction kinetics is increased when propane is used as cosolvent, reducing the residence time in the reactor. The last column in Table 2 shows results for a process scheme with no heat pump. As compared to the one with heat pump, it can be seen that operating costs have increased in 470 % due to higher steam requirement for heating the feed to the first separator. Capital costs are a slightly lower because there is no need for a compressor and the biodiesel production cost has increased in almost 4 %.

Table 2. Cost comparison for the plant with and without the use of a co-solvent

Costs	24-to-1, HPump C3/Meth=0.	24-to-1, HPump C3/Meth=0.05	40-to-1, HPump C3/Meth=0.	40-to-1, no HPump C3/Meth=0.
Operating (US$/h)	34.32	55.81	50.08	284.89
Operating (US$/t biodiesel)	3.41	5.56	4.98	28.23
Capital (US$/t biodiesel)	4.66	5.88	5.15	4.47
Plant (US$)	2,287,826	2,876,631	2,526,000	2,195,050
Production Cost (US$/t biodiesel)	598.98	608.55	626.32	650.72

5. Conclusions

Different alternatives to supercritical biodiesel production have been studied in the present work by formulating nonlinear programming problems based on rigorous process models and thermodynamic predictions for each process scheme. Numerical results indicate that the use of propane as co-solvent slightly reduces the residence time in the reactor but increases both production and capital costs. Operating costs are significantly reduced considering the use of a heat pump scheme, making the supercritical biodiesel production economically competitive with conventional catalytic production and providing additional advantages that include no need for feed treatment, catalyst and soap removal.

References

Andreatta, A.E., Casas, L.M., Hegel, P., Bottini, S.B., Brignole, E.A., 2008. Phase Equilibria in Ternary Mixtures of Methyl Oleate, Glycerol and Methanol. Ind. Eng. Chem. Res. 47, 5157-5164.

Biegler, L. T., Cuthrell, J. E.., 1985. Improved Infeasible Path Optimization for Sequential Modular Simulators II: The Optimization Algorithm. Computers and Chemical. Enginering 9, 257-265.

Cao, W., Han, H., Zhang, J., 2005. Preparation of Biodiesel from Soybean Oil using Supercritical Methanol and co-solvents. Fuel 84, 347-351.

Diaz, M. S., Espinosa,S., Brignole, E. A. Model-based cost minimization in noncatalytic biodiesel production plants. Submitted to Bioresour. Technol., November 2008.

Espinosa, S., Foco, G., Bermúdez, A., Fornari, T., 2000. Revision and Extension of The Group Contribution Equation of State to New Solvent Groups and Higher Molecular Weight Alkanes. Fluid Phase Equilibria 172, 129-143.

Espinosa, S., Diaz, M.S., Brignole, E. A., 2007. Process synthesis and optimization of a supercritical methanol biodiesel production plant. ECCE-6, 6th European Congress of Chemical Engineering, Copenhagen, 16-20 September 2007, 423-424.

Ferreira, O., Brignole, E.A., Macedo, E.A., 2004. Modelling of phase equilibria for associating mixtures using an equation of state. J. Chem. Thermodynamics 36, 1105-1117.

Freedman, B., Pryde E.H., Mounts T.L., 1984. Variables affecting the yields of fatty esters from transesterified vegetable oils. *J.Am. Oil Chem. Soc.* 61, 1638

Fukuda, H., Kondo, A., Noda, H., 2001. Biodiesel Fuel Production by Transesterification of Oils. *Journal of Bioscience and Bioengineering* 92, 405-416.Gros, H.P., Bottini, S.B., Brignole, E.A., 1996. A group contribution equation of state for associating mixtures, Fluid Phase Equilibria 116, 537-544.

Hegel, P., Mabe, G., Pereda, S., Brignole, E.A., 2007. Phase Transitions in Biodiesel Reactor using Supercritical Methanol, Ind. Eng. Chem. Res. 46, 6360-6365.

Institut Français du Pétrole, 1981. Manual of Economic Analysis of Chemical Processes.. McGraw-Hill Book Company, France.

Kusdiana, D., Saka, S., 2001. Kinetics of transesterification in rapeseed oil to biodiesel fuel as treated in supercritical methanol. Fuel 80, 693-698.

Kusdiana, D., Saka, S., 2004. Effects of water on biodiesel fuel production by supercritical methanol treatment. Bioresour. Technol. 91, 289-295.

Lotero, E., Liu, Y., Lopez, D.E., Suwannakarn, K., Bruce, D.A., Goodwin, Jr., J. G., 2005. Synthesis of Biodiesel via Acid Catalysis. *Ind. Eng. Chem. Res.* 44, 5353-5363

Negi, D.S., Sobotka, F., Kimmel, T., Wozny, G., Schomäcker, R., 2006. Liquid-Liquid Phase Equilibrium in Glycerol- Methanol- Methyl Oleate and Glycerol – Monoolein –Methyl Oleate Ternary Systems. Ind. Eng. Chem. Res. 45, 3693-3696.

Peters, M.S., Timmerhaus, K.D., 1991. Plant Design and Economics for Chemical Engineers. McGraw-Hill, NY.

Ulrich G.D. A Guide to Chemical Engineering Process Design and Economics, John Wiley & Sons, NY, 1984.

Varma, M.N., Madras, G., 2007. Synthesis of Biodiesel from Castor Oil and Linseed Oil in Supercritical Fluids. *Ind. Eng. Chem. Res*. 46, 1-6.

Watanabe, Y., Shimada, Y., Sugihara, A., Noda, H., Fukuda, H., Tominga, Y., 2000. Continuous production of biodiesel fuel from vegetable oil using immobilized Candida antarctica lipase. J. Am. Oil Chem. Soc. 77, 355-360

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Nested Heuristic and Gradient-based Method for Generalized Disjunctive Programming

Daqing Tian[1], Lingyu Zhu[2], Xi Chen[1*], Zhijiang shao[1], Jixin Qian[1]

[1]*State Key Lab of Industrial Control Technology, Department of Control Science and Engineering, Zhejiang University, Hangzhou 310027, China.*
[2]*College of Chemical Engineering and Materials Science, Zhejiang University of Technology, Hangzhou 310014, China*

Abstract

The Generalized Disjunctive Programming (GDP) has been proposed as an alternative to the mixed integer nonlinear programming (MINLP) for applications in process synthesis. It is straightforward in conditional modelling. Besides, it can reduce the complexity of the NLP sub-problems. In this paper, a nested method combining heuristic algorithm and gradient-based optimizer is proposed to solve the GDPs. It is generally a two-layer method, where the heuristic algorithm such as Tabu Search (TS) or Genetic Algorithm (GA) performs master iterations in the outer loop dealing with the logical variables of GDPs, and a gradient-based NLP solver such as Sequential Quadratic Programming (SQP) is applied in the inner layer dealing with the sub-NLP problems. In each call by the outer-layer heuristic algorithm, a set of logic values is specified and passed to the inner-layer. The general GDP model can thus be transformed into a reduced NLP sub-problem by automatically eliminating a number of equations based on the logic disjunction, nonlinear ones in some cases. After the execution of the gradient-based NLP solver, the result of each NLP sub-problem is returned to the outer-layer for further iterations. The heuristic method is responsible to find the global optimum of logic variables. Good performance has been demonstrated by applying the combined method into process synthesis such as heat exchanger networks (HEN) and the integrated water systems.

Keywords: process synthesis, generalized disjunctive programming, nested method.

1. Introduction

The synthesis of process network problems involving discrete variables are usually formulated as Mixed Integer Nonlinear Programming (MINLP) models. In the past decade however, researchers have presented Generalized Disjunctive Programming (GDP) as an alternative modelling method for MINLP.[1] Compared with MINLP, the GDP model is more intuitive and straightforward. Besides, it reduces the complexity of the generated sub-NLP problems, when the Boolean variables are specified.

A number of deterministic methods which used to solve the MINLP have been adapted to solve the GDP problems.[2] Although these algorithms have the advantage of fast convergence, it is well known that global optimality cannot be guaranteed for non-convex models; or in other words, the global optimization relies on the mathematical form of the handled problem.[3] In contrast to the inherent flaw of deterministic methods on global optimization, the heuristic algorithms are born to have advantage on this

* : Corresponding author. E-mail: xichen@iipc.zju.edu.cn

issue. Methods that can be categorized into heuristic algorithms include simulated annealing,[4] tabu search[5,6] and genetic algorithm,[7] etc. For example, Tabu search has already been applied to solve chemical process optimization problems[8].

Combining deterministic and heuristic algorithms, the idea of hybrid method has also been proposed for global optimization with fast convergence.[9] In this paper we will propose a nested framework combining heuristic and gradient-based methods to solve the GDP problems. Numerical results of several typical process synthesis problems which are formulated as GDP are also given.

2. Nested heuristic and gradient-based method

2.1. The decomposition strategy of nested method

The general form of the GDP model is as follows:

$$
\begin{aligned}
&Min \quad Z=\sum_k c_k + f(x) \\
&s.t. \quad g(x) \le 0 \\
&\quad \begin{bmatrix} Y_k \\ h_{1k}(x) \le 0 \\ c_k = \gamma_k \end{bmatrix} \vee \begin{bmatrix} \neg Y_k \\ h_{2k}(x) \le 0 \\ c_k = 0 \end{bmatrix} \quad k \in D \qquad \text{(GDP)} \\
&\quad \Omega(Y)=\text{True} \\
&\quad x \in R^n, c \in R^m, Y \in \{true, false\}^m
\end{aligned}
$$

The GDP model allows conditional modelling of equations by using Boolean variables. When Y_k, the Boolean variable denote the existence of a unit and is false, the continuous variables in the second block are usually set to be zero. As soon as a set of Boolean variables' value is specified, the GDP model will be transformed into a NLP sub-problem. Equations belonging to the inactivated terms will be eliminated, and variables corresponding to the non-existing units usually will be set to zero and maybe also removed from the NLP sub-problem. With the reduction in dimension and avoidance of singularities, the NLP sub-problem can always be solved efficiently.

Here we propose a nested method, involving two layers of sub-solver: in the outer loop a heuristic algorithm such as TS or GA performs the master iterations, dealing with the Boolean variables only; in the inner layer, a gradient-based NLP solver such as Sequential Quadratic Programming (SQP) is applied, to solve the reduced NLP sub-problems when the Boolean variables are specified. This method decomposes the GDP model into two layers, as is shown in Fig 1. The heuristic algorithm solves a combinatorial optimization problem, which has no explicit mathematical formulations. And it's responsible for iteratively locating the value of Boolean variables to escape from being trapped in a local optimum and therefore reach global optimality of the GDP problem. In each call by the outer loop, a set of Boolean values is passed to the inner layer, producing a reduced NLP problem to be solved by the inner gradient-based NLP solver. The sets is defined as: $S1=\{m|Y_m=true\}$ and $S2=\{n|Y_n=false\}$.

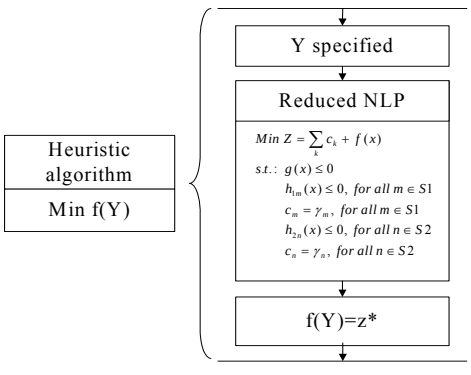


Fig 1 Structure of the nested method

Next we look at the GDP model and the nested method from a somewhat different viewpoint. After a set of Boolean values is

specified, a sub-network is produced from the postulated superstructure, with the "false-value" units disappeared. Reflected in the mathematical formulation, the GDP problem is transformed into a NLP problem with a reduction in both the equation and variable dimensions. Therefore, each Boolean value, when denoting the existence of a unit, possesses two sorts of members: variables and equations (relations of variables). When this unit is selected in the reduced NLP, the variables belonging to this unit will be present, so do the relations of these variables; otherwise, the variables corresponding to the non-existing unit will be set to zero and removed from the produced NLP.

This idea is quite similar to the Objected-Oriented Programming (OOP) idea in the computer programming language, C++. When an object is created, it means all the class members belonging to this object are created in the computer memory. The GDP model then can be transformed as follows:

$$
\begin{aligned}
Min \quad & Z = \sum_k \mathrm{Unit}_k.c_k + f(\mathrm{Unit}.x) \\
s.t. \quad & \mathrm{g}(\mathrm{Unit}.x) \le 0 \\
& \begin{bmatrix} \mathrm{unit}_k \\ h_{1k}(\mathrm{Unit}.x) \le 0 \\ \mathrm{Unit}_k.c_k = \gamma_k \end{bmatrix} \quad \mathrm{k} \in \mathrm{D} \\
& \Omega(\mathrm{unit}_k) = \mathrm{True} \\
& x \in \mathrm{R}^n, c \in R^m, \mathrm{unit}_k \in \{true, false\}^m
\end{aligned}
$$

Here we use the "." operator to indicate the "belong-to" relationship. Generally every continuous variable belongs to a unit or another. In the disjunctive part, instead of non-existing continuous variables being set to zero, the second block of equations usually does not exist at all. The g(.) functions which will hold irrespective of the disjunctive values, are usually used for mass and enthalpy balance. If certain unit doesn't exist, relevant variables in the g(.) should also be removed from the sub-NLP. The conditional modelling of equations has been discussed by the original GDP model. The conditional selection and removal of variables can be accomplished through logical operators in the equation itself. For example, in the HENS case, the heat balance at each stage can be implemented like this: $(T_{i,k} - T_{i,k+1})Fcp_i = \sum_{unit_{i,j}=true} Q_{i,j,k}$. The variable $Q_{i,j,k}$ will be activated only when the $unit_{i,j,k}$ is selected.

This variant will produce a better sub-NLP than the original GDP, in that the sub-NLP will have fewer variables and avoid singularities effectively. However, both models can be solved by the nested method proposed in this paper.

2.2. The heuristic algorithm in the outer layer

Both Genetic Algorithm (GA) and Tabu Search (TS) have been implemented to perform the outer loop of the GDP solver. Due to space limitation, only the results and discussion of The Tabu Search method will be presented.

Tabu Search (TS) is a meta-heuristic method which employs neighbourhood search to explore the solution space and prevent it from being trapped in a local optimum. Define a set neighbourhood $\mathrm{NY_k}$ as follows:

$$
NY_k = \{ y' \left| \sum_{m=1}^{M} \left| y'_m - y_m \right| = k \right\}, k \in N^+
$$

where M is the number of disjunctions, y is the current solution, and y' is one of its neighbours. The definition of the NY_k means that its element y' is exactly k bits different from the current solution. From the viewpoint of configuration, a neighbour is different from the current one in k unit's presence or absence.

As for the tabu and aspiration criteria, all the past current solutions and the local optimums (which are defined by no improvement for L iterations) will be kept for tabu use. There is also a probability p=2% to aspire a tabu point as a candidate for the selection of the next current solution. Besides, almost all the parameters used are initialized depending on the number of disjunctions, such as the maximum iteration "MaxIterCount", the maximum iteration defining a global solution to be found "MaxIterNoImprove", the length of the tabu lists, etc.

2.3. The gradient-based optimizer in the inner layer

With Boolean variables specified in the outer loop, the GDP model is branched into a typical NLP sub-problem through conditional modelling of equations. This is a typical optimization problem which is solved for the best operating levels for a given structure. Numerous solution techniques have been developed so far to solve the constrained NLP problems. For example, Sequential Quadratic Programming (SQP) accomplishes this by optimizing a sequence of quadratic programming (QP) sub-problems iteratively. The global solution of the GDP problems relies on the global solution of the sub-NLP directly. For the nested method, since the GDP's conditional modelling has greatly reduced the complexity of the NLP problem, and the same set of Boolean variables may be revisited by the outer loop, global optimization of the generated NLP becomes attainable for most problems. And for the NLP problems of examples in chapter III, produced when the global structure are given, the successful ratio of global optimization are all above 95%.

3. Test problems and numerical results

In this section we will apply the proposed nested method to two typical process synthesis problems: a HENS problem and a water treatment network problem. The program of the nested method is implemented with Microsoft Visual Studio 6.0 for the heuristic part and General Algebraic Modelling System (GAMS) for the inner NLP part. The execution of the program is run on an Intel 2.8 G Hz, 3.0G memory, PC, with Windows Server 2003. The nested method adopts the Tabu search in the outer layer and the SNOPT in the inner layer.

Case 1: HENS example

The first case to be studied is a Heat Exchanger Network Synthesis (HENS) problem, cited from Example 5 in Zamora and Grossmann's paper.[10] It involves 3 hot and 2 cold streams, steam and cooling water, and no stream splits. We reformulate it in a GDP model and then solve it by the nested method. The starting configuration is constructed assuming that the coolers and heaters all exist and no exchangers are present, which is an intuitive and feasible structure as listed in the second row of Table 1. As for the TS's parameters, the "MaxIterNoImprove" is same as the number of disjunctions, ND. The maximum number of different bits defining a neighbour K is half of ND, which is relatively large because the logical constraints are quite strict. The iterative process of the proposed method is given in Table 1. As can be seen, after 8 iterations, the nested method finds the global optimal solution and after ND iterations of no improvement for this solution, the method exits with the objective value of 82042.91.

The results of the nested method are listed in Table 2. The second column indicates the successful times to reach the global optimality among 30 runs. The third column gives an average count of the actually used NLP runs, which is highly relevant with the computation time. The forth column shows the ratio r of the actually used NLP count to the possible NLP count on average, which is 2^{23} here. The fifth column shows the major iteration count to reach the solved best solution versus the actually used iterations on

average. These three columns highly reveal the performance of the heuristic method looking for the optimal structure, and therefore the global solution.

Table 1: Iterative process for Case 1

Iteration No.	Boolean variables (0/1)	Objective value	Feasibility
0	000000000000000000011111	95415.23	Yes
1	000000000000010000011101	84189.18	Yes
2	000000000000001000011101	84189.18	Yes
3	000100000000001000011101	87639.28	Yes
4	000010000100011000011101	92573.90	Yes
5	000010000000001000011101	95039.28	Yes
6	000010000000001010011101	92830.18	Yes
7	000010000000001010011001	85431.14	Yes
8	000000000000001010011001	82042.91	Yes
……	……		

Table 2: Performance data for Case 1

No. of Disjunction	Global runs	Average NLP Count	Average Ratio r	Average Major iteration	Average time (s)
23	26/30	1210	1/6933	8/31	106

Case 2: WUTN example

Case 2 is a water network problem consisting of three effluent streams with three contaminants and three treatment units, which is adapted from published examples of Wang and Smith[11] and Gunaratnam.[12] It is complicated in that it involves highly nonlinear terms in the objective function and bilinear terms in the constraints.

The objective function consists of (1) the capital cost of the treatment units; (2) the operating cost of the treatment units; (3) the piping cost. The cost of a pipe $Cost_{i,j}^{pipe}$ is calculated as follows:

$$F_{i,j} = A_{i,j} v_{i,j} / 3600$$

$$Cost_{i,j}^{pipe} = (a A_{i,j} + b Y_{i,j}) d_{i,j}$$

where the flow rate F is in t/h, the cross-sectional area A in m^2, the distance d between the units is in m, and the velocity v in m/s, a=3603.4, b=124.46. The objective function is as follows:

$$Total\,cost = \sum_{j \in TU} Cap\,cost_j + \sum_{j \in TU} Oper\,cost_j + \sum_{k \in U} cost_{k,w}^{pipe} + \sum_{i \in PU, j \in TU} cost_{i,j}^{pipe} + \sum_{j \in TU} cost_{j,j}^{pipe}$$

The economic and other process data can be found in Gunaratnam's paper.[12] Besides, the velocity of the flows in each pipe is constrained to be 1m/s and the flow rate of each stream, if it exists, must be above 0.5 t/h.

As can been seen above, this example has 21 Boolean variables: 6 denote the existence of discharge pipes for each unit, 9 for pipes from 3 process units to 3 treatment units and 6 for treatment units interconnections.

Existing methods for water treatment networks always rely on the insights into the water networks, such as at least one of the contaminants must reach a maximum outlet concentration level, and then decompose the original complex models. Here however we apply the nested method to solve this problem without employing any engineering insights, just to examine its performance, and the starting Boolean variables are randomly generated. As for the TS's parameters, the "MaxIterNoImprove" is 1/3 of disjunctions, ND, and maximum number of different bits defining a neighbour K is also 1/3 of ND. The performance data for case 2 is listed in Table 3.

Table 3: Results of Case 2

No. of	Global	Average NLP	Average	Average Major	Average

Disjunction	runs	Count	Ratio r	iteration	time (s)
21	43/50	1227	1/1709	6/13	272

The objective of the solution is 482688.97 and the corresponding configuration is illustrated in Fig 2. To check the global optimality, 10 runs have been repeated by strengthening the TS search ability, which cost more than 7000 NLP counts and 1900 seconds in each run. Starting from 10 sets of random Boolean values all tests reached this same solution.

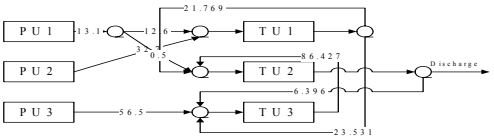


Fig 2: Global optimum solution structure for Case 2

4. Conclusion

In this paper, a nested framework combining general-purpose heuristic algorithm and gradient-based NLP solvers is proposed to solve the GDP model. It avoids the complexity and incapability of deterministic methods for global optimization, while performs quite well to solve different process synthesis problems. An Objected-Oriented viewpoint is also introduced to offer better understanding of not only the GDP model but also the proposed method.

Acknowledgements

We gratefully acknowledge the financial support of the National Basic Research Program of China (No. 2009CB320603) and National Natural Science Foundation of China (No. 60704029).

References

M. T**ü**rkay, I. E. Grossmann. Logic-based MINLP algorithm for the optimal synthesis of process networks. Computers and Chemical Engineering 1996; 20: 959-978.

I. E. Grossmann. Review of nonlinear mixed-integer and disjunctive programming techniques. Optimization and Engineering2002; 3: 227-252.

C. A. Floudas, I. G. Akrotirianakis, S. Caratzoulas, C. A. Meyer, J. Kallrath. Global optimization in the 21st century: advances and challenges. Computers and Chemical Engineering 2005; 29: 1185-1202.

S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. Science 1983; 220: 671-680.

F. Glover. Tabu search- Part I. ORSA Journal on Computing 1989; 1:190-206.

F. Glover. Tabu search- Part II. ORSA Journal on Computing 1990; 2:4-32.

D. E. Goldberg. Genetic algorithm in search, optimization, and machine learning. Addison-Wesley, Reading, MA, 1989.

B. Lin, D. C. Miller. Solving heat exchanger network synthesis problems with Tabu Search. Computers and Chemical Engineering 2004; 28: 1451-1464.

X. Chen, et al. Nested Tabu Search (TS) and Sequential Quadratic Programming (SQP) method, combined with adaptive model reformulation for heat exchanger network synthesis (HENS) Industrial and Engineering Chemistry Research 2008; 47: 2320-2330.

J. M. Zamora, I. E. Grossmann. A global MINLP optimization algorithm for the synthesis of heat exchanger networks with no stream splits. Computers and Chemical Engineering 1998; 22(3): 367-384.

Y. P. Wang, R. Smith. Design of distributed effluent treatment systems. Chemical Engineering Science 1994; 49(18): 3127-3145.

M. Gunaratnam, A. Alva-Argez, A. Kobossis, J.-K. Kim, R. Smith. Automated design of total water systems. Industrial and Engineering Chemistry Research 2005; 44(3): 588-599.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A sequential variable decoupling method for rigorous calculation of molecular weight distribution of batch free radical polymerization

Xi Chen[a*], Feijun Jiang[a], and Zhen Yao[b*]

a State Key Laboratory of Industrial Control Technology, Department of Control Science & Engineering., Zhejiang University, Hangzhou 310027, China.
b State Key Laboratory of Chemical Engineering,(Polymn. Division), Department of Chemical Engineering, Zhejiang University, Hangzhou 310027, China

Abstract

This paper introduces a new method on calculating the Molecular Weight Distribution (MWD) of free radical polymerization, termed the Sequential Variable Decoupling (SVD) method. The original model of MWD calculation in free radical polymerization is based on fundamental reaction mechanism and diffusion controlled reaction rate coefficients, resulting in a group of large scale differential algebraic equations (DAEs). However the traditional DAEs solvers always require huge computer memory to solve this kind of model simultaneously, making it impossible to implement on a normal PC. The SVD method decomposes the large DAEs model into a small scale DAEs model and a series of small scale ordinary differential equations (ODEs) models. This decomposition reduces computer memory greatly and makes it executable on a normal PC. As an example, the SVD method has been used to calculate the MWD of methyl methacrylate (MMA) homopolymerization with good precision. To further reduce computation time cost, techniques with variable step size in the numerical integration along time and chain length of the polymers are also proposed. By using the methods, the computational efficiency is improved by 8.2 times at a cost of increasing the relative error only by 0.56%.
Keywords: free radical polymerization, molecular weight distribution, sequential variable decoupling method.

1. Introduction

The Molecular Weight Distribution (MWD) of free radical polymers is the major factor that determines the end-use properties of polymers. A general and precise method that predicts MWD based on the operation parameters is very helpful. For a typical polymerization process, the range of molecular weight is from 10^3 to 10^5 or even more. Therefore, the models of polymerization are always large in scale, which presents difficulties while solving these models.

In the previous research, a great diversity of methods have been introduced, such as lumping method[1], passage to continuous variable method, the method of moments[2], the discrete Galerkin h-p-method[3], the Monte-Carlo method[4][5][8]. All these methods

* : Corresponding author. E-mail: xichen@iipc.zju.edu.cn
* : Corresponding author. E-mail: yaozhen@iipc.zju.edu.cn

include simplifying assumptions which help reduce equations, but at the same time reduce its precise and limit its application area[6][7].
In the first part of our paper, we introduce a general model of MWD calculation. Our model of free radical polymerization is based on the fundamental reactions mechanisms, including material balance equations, the volume relations, and chemical kinetic expressions thermodynamic properties. In the second part, we will introduce the Variable Decoupling Sequential Method, to decompose the original large scale DAEs model into small scale DAEs with a series of small scale ODEs sequential model. Finally, we will discuss the solver for this specific model, and we will optimize the compute process, which will improve the algorithm efficiency greatly.

2. Model decomposition of MWD calculation

2.1. Model I, Original model
The kinetics of typical free-radical homo-polymerization initiated by mono-functional initiators could be described by the following reactions: initiation, propagation, transfer and termination by combination and disproportion. A modified Stickler-Panke-Hamielec (SPH)[9] Model based on free-volume theory is used in this work. Here we only take the case of mono-functional initiators and bulk polymerization into account for simplicity. Equations, describing the whole model of free radical polymerizations, contain both ODEs and algebraic ones, the so-called large scale DAEs. The dimension of the ODEs group depends on the maximum value of the chain length. However, theoretically the chain length will grow to infinite, which make it impossible to solve this large scale DAEs. Therefore it is necessary to implement the moment method to convert the infinite item into finite item.

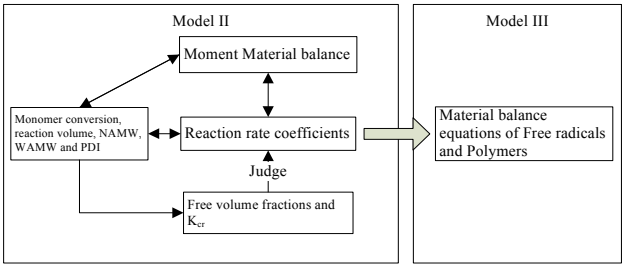


Fig.1 Structure of Model II and its relationship with Model III

2.2. Model II, Moment model
By using the moment method, we replace the ODEs of free radicals and polymers, the infinite items, with the ODEs of j^{th} moment of free radicals and polymers, whose dimension is limited.
The j^{th} moment of free radicals and polymers are defined as follows:

$$\lambda_j = \sum_{m=1}^{\infty} n^j R_n^{\bullet} \tag{1}$$

$$\mu_j = \sum_{m=1}^{\infty} n^j P_n \tag{2}$$

The moment model also contains the SPH model which is the same with model I.
And it also contains ordinary differential equations and algebraic equations, so this model is a DAEs model. However, it is a small scale DAEs group. It only contains 13 ordinary differential equations and 11 algebraic equations.
However, by solving this small DAEs group, we can only obtain the time-evolution of moment. In this case, we still can not simply derive the MWD since the concentration of polymers and free radicals with different chain lengths are not available. But with the solution of the small DAEs, we know the time-evolution of the moments, the reaction

coefficients k, the reaction volume V and other algebraic variables. Then, we can treat these variables as known variables, and solve the ODEs of free radicals and polymers to calculate the MWD, and this is the model III.

2.3. Model III, MWD calculation

Since the sums of free radicals and polymers of all chain length, are 0^{th} moments of free radicals and polymers, so the equations of free radicals and polymers can be converted into the following:

$$\frac{1}{V}\frac{d(R_1^\bullet \cdot V)}{dt} = k_i I^\bullet M - k_p M R_1^\bullet + k_{tr,M} M\lambda_0 - (k_{tc} + k_{td}) R_1^\bullet \lambda_0 + k_{tr,P}(P_1\lambda_0 - R_1^\bullet \mu_0) \tag{3}$$

$$\frac{1}{V}\frac{d(R_n^\bullet \cdot V)}{dt} = k_p R_{n-1}^\bullet M - k_p R_n^\bullet M - k_{tr,M} R_n^\bullet M - k_{tr,I} R_n^\bullet I_2 + k_{tr,P}(P_n\lambda_0 - R_n^\bullet \mu_0) - (k_{tc} + k_{td}) R_n^\bullet \lambda_0 \quad (n=2, \ldots, CL_{max}) \tag{4}$$

$$\frac{1}{V}\frac{d(P_n V)}{dt} = k_{tr,M} R_n^\bullet M + k_{tr,I} R_n^\bullet I_2 + k_{tr,P}(R_n^\bullet \mu_0 - P_n\lambda_0) + k_{td} R_n^\bullet \lambda_0 + \frac{1}{2} k_{tc} \sum_{s=1}^{n-1} R_s^\bullet R_{n-s}^\bullet \quad (n=2, \ldots, CL_{max}) \tag{5}$$

In the above equations, we can get the time-evolution of k_i, k_p, $k_{tr,m}$, $k_{tr,I}$, $k_{tr,P}$, k_{tc}, k_{td}. I_2, $I^\bullet$, M, λ_0, μ_0, V from the solution of the small DAEs group in Model II. The concentrations of polymers and free-radicals with different chain lengths, which represent MWD, can be obtained by solving the ordinary equations (3) (4) (5). This relationship is represented in Fig.1. From this figure, it is easy to conclude that by solving model II and III sequentially, we can finish calculating MWD. And also the maximum of chain lengths that are taken into consideration can be very large with our solver represent in the next chapter.

3. Sequential Variable Decoupling method with control on truncation error

3.1. Decomposition of Model III with Sequential Variable Decoupling method

In the equations (3) (4) (5), the model III is still quite large in scale. And with the present ODEs solver, it will take quite large computer memory to solve this model simultaneously, which makes it not applicable on a personal computer. And we have tried to solve the ODEs directly, and a computer with 4GB memory can only solve the maximum chain length of 2000, which is far from sufficient for MWD predication.

In the equations (3) (4) (5), we can find that, for a specific chain length, the equations of polymers and free radicals do not contain any information from bigger chain lengths' components, which means we can solve the whole model sequentially, from small chain length

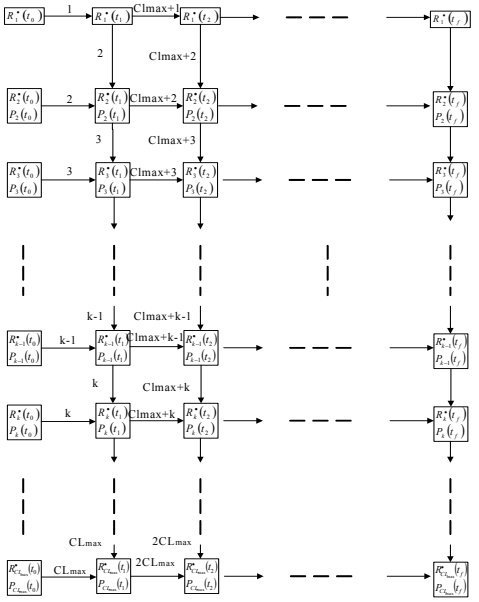


Fig.2 Structure of Decoupled Sequential Model

to big chain length components, thus the large scale ODEs group is decoupled into lots of such units sequential models. Fig.2 shows this structure, the structure of decoupled sequential model. In the Fig.2, the simulation is from the initial moment t_0 to the final moment t_f, and the whole period is sampled into several discrete points.

Since that we can solve $R_k^\bullet$ and $P_k^\bullet$, from k=1 to k=CL_{max}, sequentially, we can save lots of computer memory and make this model solvable on a personal computer. This is the key idea of the sequential variable decoupling method. The CL_{max} represents the maximum of the chain lengths that need to be considered.

However, problem rises while solving this model. Since the model III has become a large scale sequential model, if we implement traditional ODE solver on it, the computation time is quite long. And the sequential model needs synchronization on the steps among different chain lengths which make this model inflexible. So later in this chapter will discuss the optimization of this computation process.

3.2. Mechanism of variable step size

The Fig.2, the structure of decoupled sequential model, also shows the discretization of the whole reaction period, thus the whole reaction period is divided into many steps between discrete points. How to divide this period is an important issue for the computation. Here, we use first order backward differentiation formula (BDF) to solve these ODEs. Since the step size decides the efficiency and the accuracy of the algorithm, a mechanism determining the appropriate step size is very essential.

The Gear's BDF algorithm introduced a method to estimate the error of numerical integration. Here we use the first order method and build an error estimator. The step size of the numerical integration is determined by the error and error tolerance. Each model has a specific requirement on the accuracy. For this model, 10% relative error is acceptable. For each iteration, the mechanism is described as follows:

(1) Use the old step size *h* to calculate a new step.

(2) Run the error estimator to estimate the error for this step

(3) If the error is smaller than the tolerance, then increase the step size and go to next step (1); if not, reduce the step size and calculate this step again, also go to (1), and this step is the feedback and the control on the step size to maintain the error within the tolerance.

3.3. Variable step size on time scale

Fig.2 shows the structure of the sequential ODEs group. It denotes that, at each step of numerical integration, the method will calculate the polymers and free radicals fractions for all chain lengths. To ensure this synchronization, the step sizes for a certain step, of all variables, must be the same, which means, it needs to choose the minimal step size to satisfy the accuracy requirement of all variables. If one variable's accuracy is violated, the whole step needs to be calculated again.

3.4. Variable step size with chain length

In the error estimator mentioned above, there is a variable α calculated to represent the sensitivity towards the step size. For this specific model, we find that the α increases with the chain length, which means, for a certain step, the variables that represent larger chain length polymers and free radicals are more inert than the variables that represent smaller chain length polymers and free radicals.

Because the large chain length polymers and free radicals are always inert, and according to their accuracy requirement, their step size can be further amplified. At each

even step we set a transition point. The calculation of the variables for the chain length larger than this point is skipped on this step. The step size of the variables of the chain length larger than this point is twice the size of the variables of the chain length smaller than this point. For the variables of small chain length, their step size is still determined by the feedback control mechanism. We decide the transition point in this way: if the α of the polymer variable for a certain chain length is very large, we can conclude that from this chain length on, the polymer and free radicals are inert, so this point is the transition point.

4. Results and Discussion

4.1. MWD results from model III with constant step size

Here we implement an example of MMA homopolymerization, the 0.3% (weight fraction) AIBN at 90□ for 30 minutes. First we solve the Model II by DAEs solver, and generate the time-evolution of the moments, the reaction rates *k*, the reaction volume *V*, which are the inputs of model III.

By applying the Sequential Variable Decoupling method, we can calculate the very wide chain length range. In our example, we calculate the MWD with maximum chain length of 5×10^4, which is sufficient for polymer design.

Here the step size is constant, and in order to decide an appropriate step size, a comparison between different step sizes is needed. The comparison among step size of 0.5 min, 0.1 min and 0.01min is show in Tab.1. From the comparison, we know that the step size of 0.1min is optimal for this model, because the accuracy improvement for the step size from 0.5min to 0.1min is quite significant, but the accuracy improvement for the step size from 0.1min to 0.01min is not so significant and the maximum relative error between 0.1min and 0.01min is within the accuracy requirement 10%. So here for the constant step size the 0.1 min step size is an optimal one. And for the later discuss, we take the results of the calculation with step size 0.01min as the standard data for comparison.

Tab.1 Comparison of MWD calculations

Step size	Compute time (s)	Maximum relative error with step size=0.01min(%)
0.5min constant	292.38	50.2%
0.1 min constant	769.26	9.01%
Feedback control on time	138.17	9.86%
Step Size variable with Chain length	94.13	9.57%
0.01min consant	5324.78	

4.2. Variable step size on time scale

The result of the calculation is quite positive, by applying the feedback control, the efficiency improved by 5.6 times, which means it only takes 18.0% computation time of the constant step size method with a step size of 0.1min, and the maximum relative error is 9.86% comparing to 9.01% for the result of 0.1min. The comparison is also shown in Tab.1. Since the MWD is constituted of tens of thousands

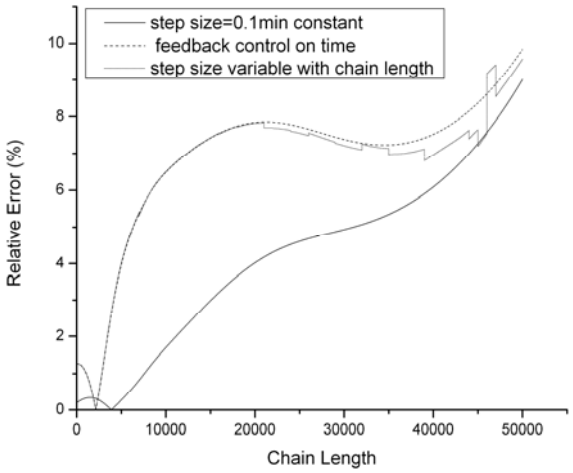


Fig.4 Relative error comparison of different methods

variables, and the accuracy of the calculation should be represent as a distribution of the relative error, it is show in Fig.4.

4.3. Variable step size on chain length

The result of calculation is positive too. The comparison is also contained in Tab.1. And the relative error distribution and its comparison with the relative error of feedback control on time scale, is in Fig.4. The time cost further decreased by 31.9% and the relative error do not increase noticeably.

5. Conclusion

This paper introduced a novel method on MWD calculation. We decompose the large scale DAEs model into a group of small scale sequential ODEs, which make the MWD solvable on a normal PC. We further discuss the improvement of computational process by varying step size on both time scale and chain length, which greatly reduces the time at a cost of little deterioration in accuracy.

Acknowledgements

We gratefully acknowledge the financial support of the National Basic Research Program of China (No. 2009CB320603) and National Natural Science Foundation of China (No. 60704029).

References

[1] A. M. Basedow, K. H. Ebert, H. J. Ederer, 1978, Kinetic Studies on the Acid Hydrolysis of Dextran, Macromolecules, 11, 774

[2] S. Katz, G. M. Seidel, 1967, Moments of the size distribution in radical polymerization, AIChE J., 13, 319

[3]M. Wulkow, 1996, The simulation of molecular weight distributions in polyreaction kinetics by discrete Galerkin methods, Macromol. Theory Simul., 5, 393-416

[4]Gurutze Arzamendi and Jose Ramon Leiza, 2008, Molecular Weight Distribution (soluble and Insoluble Fraction) in Emulsion Polymerization of Acrylate Monomers by Monte Carlo Simulations, Ind. Eng. Chem. Res., 47, 5934-5947

[5]Hidetaka Tobita, 2006, Molecular Weight Distribution of Living Radical Polymers, Macromol. Theory Simul. 15, 23-31

[6]Timothy J.Crowley and Kyu Yong Choi, 1997, Calculation of Molecular Weight Distribution from Molecular Weight Moments in Free Radical Polymerization, Ind. Eng. Chem. Res., 36, 1419-1423

[7]Alessandro Butte, Alessandro Ghielmi, Giuseppe Storti, 1999, Massimo Morbidelli Calculation of molecular weight distributions in free-radical polymerization with chain branching. Macromol. Theory Simul. 8, 498-512

[8]H. Tobita, N. Hamashima, 2000, Monte Carlo simulation of size exclusion chromatography for branched polymers formed through free-radical polymerization with chain transfer to polymer, Macromol. Theory Simul., 9(8), 453-462

[9]Stickler, M., Panke, D., and Hamielec, A.E., 1984, Polymerization of Methyl Methacrylate up to High Degrees of Conversion: Experimental Investigation of Diffusion-controlled Polymerization, J. Polymer Sci. Poly. Chem. 22, 2243-2253

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Global Optimization of Nonconvex Generalized Disjunctive Programs

Juan P. Ruiz, Ignacio E. Grossmann
Carnegie Mellon University, 5000 Forbes Ave. , Pittsburgh 15213, United States

Abstract

This paper is concerned with the global optimization of Bilinear and Concave Generalized Disjunctive Programs. The efficiency of methods to solve these problems relies heavily on their capability for predicting strong lower bounds that in turn depend on the strength of their relaxations. The major objective of this work is to propose a procedure to find such relaxations. We first present a framework for obtaining a hierarchy of linear relaxations for nonconvex Generalized Disjunctive Programs (GDP), that combines the linear relaxations proposed in literature for nonconvex MINLPs with the results in the work of Sawaya & Grossmann [1], to obtain relaxations for the case of Linear GDPs. We further exploit the theory behind Disjunctive Programming to guide the generation of better relaxations more efficiently. We show through a set of examples how these new relaxations can strengthen substantially the lower bounds for the global optimum leading to an improvement in the performance when used within a spatial branch and bound framework.

Keywords: Global Optimization, Generalized Disjunctive Programming, Convex Relaxations, Bilinear GDP, Concave GDP

1. Introduction

Generalized Disjunctive Programming (GDP), developed by Raman and Grossmann [2], has been proposed as a framework that facilitates the modeling of discrete-continuous optimization problems by allowing the use of algebraic and logical equations that are expressed in terms of Boolean and continuous variables. In order to take advantage of the existing MILP/MINLP solvers, GDPs are often reformulated as MILP/MINLP problems by using either the Big-M (BM), or the Convex Hull (CH) [3] reformulation. However, these reformulations are not unique and may have associated relaxations that are not very tight, consequently having an adverse effect on the efficiency of the algorithm that is used. In nonconvex GDP problems the direct application of traditional algorithms to solve the reformulated MINLPs may correspond

to a local optimum. Therefore, specialized algorithms should be used in order to find the global optimum [4]. Nonconvex GDP problems with bilinear and concave constraints are of particular interest since these arise in many applications [5,6]. To tackle these problems, Lee and Grossmann [7] proposed a global optimization method that first relaxes the bilinear terms by using the convex envelopes of McCormick [8] and the concave terms by using linear under estimators. The convex hull [9] is then applied to each disjunction and used within a spatial branch and bound technique. A major question is whether one might be able to obtain stronger lower bounds to enhance the efficiency for globally optimizing GDP problems. Sawaya and Grossmann [1] have recently established new connections between Linear GDP and the Disjunctive Programming theory by Balas [10]. As a result, a family of tighter reformulations has been identified. These are obtained by performing a sequence of basic steps on the original disjunctive set (i.e. each basic step is characterized by generating a new set of disjunctions by intersecting the former), bringing it to a form closer to the Disjunctive Normal Form (DNF), and hence tightening its discrete relaxation [9]. In this work we build on the work by Sawaya and Grossmann [1] exploiting the new hierarchy of relaxations in order to solve more efficiently nonconvex GDP problems.

2. Background

The general structure of a nonconvex GDP can be represented as follows [2,3,6]

$$
\begin{aligned}
&Min\ Z = f(x) + \sum_{k \in K} c_k \\
&s.t. \quad g^l(x) \le 0 \qquad\qquad l \in L \\
&\bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ r_{ik}^j(x) \le 0 \quad j \in J_{ik} \\ c_k = \gamma_{ik}(x) \end{bmatrix} \qquad k \in K \qquad\qquad (GDP_{NC}) \\
&\Omega(Y) = True \\
&x^{lo} \le x \le x^{up} \\
&x \in R^n, c_k \in R^1, Y_{ik} \in \{True, False\}
\end{aligned}
$$

where $f : R^n \to R^1$ is a function of the continuous variables x in the objective function, $g^l : R^n \to R^1$, $l \in L$, belongs to the set of global constraints. The disjunctions $k \in K$, are composed of a number of terms $i \in D_k$, that are connected by the OR operator . In each term there is a Boolean variable Y_{ik}, a set of inequalities $r_{ik}^j(x) \le 0$, $r_{ik}^j : R^n \to R^1$, $j \in J_{ik}$ and a cost variable c_k. If Y_{ik} is true then $r_{ik}^j(x) \le 0$ and $c_k = \gamma_{ik}(x)$ are enforced, otherwise they are ignored. Also, $\Omega(Y) = True$ are logic propositions for the Boolean

variables. As indicated in Sawaya and Grossmann [1], we assume that the logic constraints $\underset{i\in D_k}{\vee} Y_{ik}$ are contained in $\Omega(Y) = True$. In a nonconvex GDP, f, r_{ik}, γ_{ik} and/or g^l are nonconvex functions. In this work we aim at solving two particular nonconvex GDP cases, namely, bilinear GDPs (GDP_B) and concave GDPs (GDP_{Co}). A GDP_B is a nonconvex GDP where some of the functions f, r_{ik}, γ_{ik} and/or g^l are bilinear, whereas a GDP_{Co} is a nonconvex GDP where some of the functions f, r_{ik}, γ_{ik} and g^l are concave.

3. Methodology to obtain strong relaxations

In this section we present a general framework to obtain a hierarchy of continuous linear relaxations for the nonconvex GDP problem (GDP_{NC}) that can serve as a basis for predicting tight lower bounds of the global optimum. This framework will be presented for general nonconvex GDP problems, although it will later be only applied to the bilinear and concave case (i.e. GDP_B and GDP_{Co}). Without loss of generality we can consider in (GDP_{NC}) f to be a linear function of x. Hence, we can represent it as $d^T x$.

Let us define the following sets:

$L_c := \{ l \in L \mid g^l(x) \text{ is convex} \}$, $L_{nc} := \{ l \in L \mid g^l(x) \text{ is nonconvex} \}$, $J_{cik} := \{ j \in J_{ik}, i \in D_k, k \in K \mid r^j_{ik}(x) \text{ is convex} \}$, $J_{ncik} := \{ j \in J_{ik}, i \in D_k, k \in K \mid r^j_{ik}(x) \text{ is nonconvex} \}$, $D_{cik} := \{ i \in D_k, k \in K \mid \gamma_{ik}(x) \text{ is convex} \}$, $D_{ncik} := \{ i \in D_k, k \in K \mid \gamma_{ik}(x) \text{ is nonconvex} \}$

In the first step of this approach we replace each nonconvex function with a valid linear over/under estimator. Generalizing,

$g^l(x) \le 0$ is replaced by $A^l_g x \le b^l_g$, $l \in L_{nc}$, $r^j{}_{ik}(x) \le 0$ is replaced by $A^j_{rik} x \le b^j_{rik}$, $j \in J_{ncik}$, $i \in D_k$, $k \in K$ and $c_k = \gamma_{ik}(x)$ is replaced by $A_{\gamma ik}(x, c_k) \le b_{\gamma ik}$, $i \in D_{ncik}$, $k \in K$

Note that the dimensions of the matrices, A^l_g, A^l_{rik}, $A_{\gamma ik}$, and the right hand side vectors, b^l_g, b^j_{rik}, $b_{\gamma ik}$, depend on the particular under/over estimators that are chosen.

Similarly, we replace each convex inequality with a valid linear outer-approximation,

$g^l(x) \le 0$ is replaced by $A^l_g x \le b^l_g$, $l \in L_c$, $r^j{}_{ik}(x) \le 0$ is replaced by $A^j_{rik} x \le b^j_{rik}$, $j \in J_{cik}$, $i \in D_k$, $k \in K$ and $c_k = \gamma_{ik}(x)$, is replaced by $A_{\gamma ik}(x, c_k) \le b_{\gamma ik}$, $i \in D_{cik}$, $k \in K$

Note that the dimensions of the matrices, A^l_g, A^l_{rik}, $A_{\gamma ik}$ and the vectors b^l_{rik}, b^l_g, $b_{\gamma ik}$ depend on the polyhedral outer-approximation technique that is chosen [4].

Replacing the nonconvex functions in (GDP_{NC}) by the corresponding under/over estimators and outer-approximations, leads to the following Linear GDP:

$$
\begin{aligned}
&\underset{\cdot}{Min}\ Z^L = d^T x + \sum_{k \in K} c_k \\
&s.t.\quad A_g^l x \le b_g^l \qquad\qquad l \in L \\
&\bigvee_{i \in D_k} \begin{bmatrix} Y_{ik} \\ A_{rik}^j x \le b_{rik}^j \quad j \in J_{ik} \\ A_{\gamma ik}(x, c_k) \le b_{\gamma ik} \end{bmatrix} \qquad k \in K \qquad\qquad (GDP_{RLP}) \\
&\Omega(Y) = True \\
&x^{lo} \le x \le x^{up} \\
&x \in R^n, c_k \in R^1, Y_{ik} \in \{True, False\}
\end{aligned}
$$

In order to predict strong lower bounds for the global optimum of (GDP_{NC}), as the second step in this approach, we consider the hierarchy of relaxations for (GDP_{RLP}) from the work of Sawaya & Grossmann [1]. These authors proved that any Linear Generalized Disjunctive Program (LGDP) that involves Boolean and continuous variables can be equivalently formulated as a Disjunctive Program that only involves continuous variables. Hence, Disjunctive Programming Theory can be used to solve Linear GDPs. One important result in Disjunctive Programming Theory, as presented in the work of Balas [9], is that we can systematically generate a set of equivalent DP formulations going from the CNF to the DNF by using an operation called "basic step" which preserves regularity (See Theorem 2.1 [9]). Although the formulations obtained after the application of basic steps on the disjunctive sets are equivalent, their *continuous relaxations* are not. As described in Theorem 4.3 [9], the application of a basic step on a disjunctive set leads to a new disjunctive set whose relaxation is at least as tight, if not tighter, as the former. Consider now the linear relaxation of (GDP_{NC}), namely (GDP_{RLP0}), which is equivalent to (GDP_{RLP}). We introduce the subscript i to indicate the number of basic steps that has been applied to the initial Linear GDP relaxation (GDP_{RLP0}) that is obtained by the under/over estimation of the nonconvex terms from (GDP_{NC}). Then we can show that:

$$h - rel\, GDP_{RLP0} \supseteq h - rel\, GDP_{RLP1} \supseteq ..h - rel\, GDP_{RLPi} .. \supseteq h - rel\, GDP_{RLPt} \supseteq GDP_{NC}$$

where *h-rel* GDP_{RLPi} is the convex hull relaxation of GDP_{RLPi} (Lee and Grossmann [2])
We can then establish the following proposition.

PROPOSITION: The lower bounds for the global optimum obey the following relationship: $Z_{RLP0} \le Z_{RLP1} \le .. Z_{RLPi} \ldots \le Z_{RLPt} \le Z_{NC}$ where Z_{RLP0} , Z_{RLP1} ,.. Z_{RLPi} , Z_{RLPt} are the optimal solutions of the hull-relaxations of problems GDP_{RLP0} , GDP_{RLP1},.. GDP_{RLPi} .., GDP_{RLPt} respectively and Z_{NC} is the optimal solution of GDP_{NC}.

In order to implement the basic steps more efficiently we developed a set of rules based on Theorem 4.5 [9] which exploits the particular structure of the problems.

4. Numerical Results and Conclusions

In order to test the performance of the framework to generate strong relaxations we used 5 examples. Example 1 deals with the optimal selection of two reactors with bilinear constraints in the demand. Example 2 deals with the optimal selection of two reactors whose investment costs are given by a concave univariate function. Example 3 deals with the optimization of a Heat Exchanger Network with discontinuous investment costs for the exchangers and can be represented by a nonconvex GDP with bilinear and concave constraints [8]. Example 4 deals with the optimization of a Wastewater Treatment Network whose associated nonconvex GDP formulation is a bilinear GDP . Finally, example 5 is a Pooling Design problem that can be also represented as a bilinear GDP [7]. Table 3 summarizes the characteristics and size of the examples, and Table 4 shows the computational performance of the Lee and Grossmann [7] relaxation and the one proposed in this work.

Table 3 Characteristics and size at example problems

	Bilinear Terms	Concave functions	Discrete Variables	Continuous Variables
Example 1	1	0	2	3
Example 2	0	2	2	5
Example 3	4	9	9	8
Example 2	36	0	9	114
Example 3	24	0	9	76

Table 4 Performance of proposed framework

	Global Optimum	Lower Bound (Lee & Grossmann Relaxation)	Lower Bound (Proposed Relaxation)	Best Lower Bound
Example 1	-1.01	-1.28	-1.10	-1.10
Example 2	6.31	5.65	6.08	6.08
Example 3	114384.78	91671.18	94925.77	97858.86
Example 4	1214.87	400.66	431.9	431.9
Example 5	-4640	-5515	-5468	-5241

All the examples show an improvement in the lower bound which is a direct indication of the reduction of the relaxed feasible region. The column “Best Lower Bound” in the Table 4 represents the MIP solution of the Linear GDP after the relaxation of the nonconvex terms/functions (i.e. bilinear terms and concave terms). This ultimately represents the best lower bound attainable by the application of the basic steps (i.e. the solution of the hull relaxation of the DNF of GDP_{RLP0}). Clearly, this can be used as an indicator of the performance of our set of rules to apply basic steps (e.g. note that Examples 1, 2 and 4 reach this limit). A further indication of tightening is shown in Table 5 where numerical results of the branch and bound algorithm proposed are

presented. As it can be seen the number of nodes that the spatial branch and bound algorithm needs to visit before finding the global solution is significantly reduced in Examples 1, 3 and 4. In Examples 2 and 5 the performance is not altered significantly.

Table 5 Performance of proposed methodology

		Global Optimization Technique using Lee & Grossmann Relaxation			Global Optimization Technique using Proposed Relaxation		
	Global Optimum	Nodes	Bound contract. (% Avg)	CPU Time (sec)	Nodes	Bound contract. (% Avg)	CPU Time (sec)
Example 1	-1.01	5	35	2.1	1	38	1.4
Example 2	6.31	1	33	1.0	1	33	1.0
Example 3	114384.78	13	85	11.0	1	99	6.0
Example 4	1214.87	450	8	217	227	16	139
Example 5	-4640	502	1	268	497	1	285

References

[1] Sawaya & Grossmann (2008): Reformulations, Relaxations and Cutting Planes for Linear Generalized Disjunctive Programs. Submitted for publication

[2] Raman, R., & Grossmann, I. E. (1994). Modeling and Computational Techniques for logic based integer programming. Comp. and Chem. Eng. 18 (7), 563-578.

[3] Lee, S. & Grossmann I. E. (2000) New algorithms for nonlinear generalized disjunctive programming. Computers and Chemical Engineering 24, 2125-2141

[4] Tawarmalani, M., Sahinidis, N. (2002) Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming. , Kluwer Academic Publishers

[5] Quesada, I. & Grossmann I. E. (1995b). Global optimization of bilinear process networks with multicomponent flows. Comp. and Chem. Eng. 19 (12), 1219-1242.

[6] Turkay M. & Grossmann I.E. (1996) Disjunctive Programming Techniques for the Optimization of Process Systems with Discontinuous Investment Costs-Multiple Size Regions. Ind. Eng. Chem. Res. 1996, 35, 2611-2623

[7] Lee, S. & Grossmann I. E. (2003) Global optimization of nonlinear generalized disjunctive programming with bilinear inequality constraints: application to process networks. Computers and Chemical Engineering 27 , 1557-1575

[8] McCormick, G. P. (1976) Computability of global solutions to factorable nonconvex programs. Part I . Math. Prog. 10, 146-175.

[9] Balas, E. (1985) Disjunctive Programming and a hierarchy of relaxations for discrete optimization problems. SIAM Journal on Alg. and Disc. Methods, 6, 466-486.

[10] Balas, E. (1979) Disjunctive Programming. Annals of Discrete Mathematics 5 (1979) 3-51

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Solid-liquid Equilibrium Modelling and Stability Tests for Triacylglycerols Mixtures

Moises T. dos Santos,[a,b] Galo A.C.Le Roux, [a] Xavier Joulia,[b] Vincent Gerbaud [b*]

[a]*Universidade de São Paulo, Escola Politécnica, Laboratório de Simulação e Controle de Processos, Av. Prof. Lineu Prestes, São Paulo,5088-900 BRASIL*
[b] *Université de Toulouse, Laboratoire de Génie Chimique (LGC) UMR CNRS INP/UPS, Allée Emile Monso, Toulouse, 31000 France. *Vincent.Gerbaud@ensiacet.fr*

Abstract

Computer-Aided Mixture Design for product development can take advantage from equilibrium modelling. Systems composed by triacylglycerols (TAG) mixtures are widely used for many applications (foods, cosmetics, pharmaceutical and lubricants) and their end-use properties are very close related to phase behaviour (melting and crystallization). Such molecules can have different polymorphisms in solid state, leading to a lack of intersolubility and consequently formation of multiple solid phases. This work has implemented the solid-liquid equilibrium for TAG mixtures in a two step approach: stability tests and equilibrium compositions computations for two phase mixtures. The Michelsen's method for stability analysis was adapted to cope with polymorphisms and was successful for phase-split detection. Melting curves for mixtures composed by 9 TAGs in different compositions and molecular structures were simulated revealing good agreement with physical background for such systems. Further implementations of other initialization independents stability tests and robust optimization techniques show, therefore, a great potential for use as auxiliary computational framework for match improved mixtures in structured lipids research.

Keywords: triacylglycerols, polymorphisms, solid-liquid equilibrium, CAMD.

1. Introduction

Knowledge of phase equilibrium calculations has been widely applied in chemical process design, simulation and optimization. However, equilibrium calculations have also a great potential to be applied in product design, as many products have their properties directly related to phase behaviour in multi-component mixtures.

In chemical industry, phase behaviour plays an important role in the product functionality of many products: paints, rubber, plastic composites, agglomerated powders, extruded products, foams and foods (Bruin; Jongen, 2003). Phase equilibrium calculations are very useful in food industry for product design and fat-based foods are particular examples where phase behavior directly affects product requirements. The solid fat content and melting profile of such products influences the texture; the mouthfeel related to this texture and to its destruction during mastication is the key factor to final product quality and appreciation by the consumer (Bruin, 1999).

Vegetable oils and fats (natural or modified) are the main raw materials for these fat-based foods. Control of melting and solidification behaviour of such systems is therefore essential for manufacturing these products (Bruin; Jongen, 2003). Despite the importance for the design of fat-based foods with improved chemical, physical and nutritional properties, the use of phase equilibrium calculations remains still unsolved in

such systems, due to their complex nature. In fact, fats and oils are lipids systems composed mainly by triacylglycerols (TAG) molecules. For computing solid fat content, up to 1990, mainly empirical calculation methods were available (Wesdorp, 1990). Such empirical approaches limit the application range and provide no fundamental understanding of phase behaviour of TAG systems. After 1990, experimental data on fatty systems (Liang *et al*., 2003; Inoue *et al*., 2004; Zhou and Hartel, 2006; Zhang *et al*., 2007) were published along with some discussions about general TAG phase equilibrium modelling (Wesdorp, 1990; Won, 1993; Himawan *et al*., 2006).

However, a robust and general methodology to cope with the complexity of the problem is still unsolved. One reason is that fats and oils are multi-component mixtures of different triacylglycerols that can crystallize in 3 different polymorphisms (a, ß' and ß). At these solid states, such molecules have limited miscibility, leading to the formation of several solid phases. Similar problems can be found in oil industry as precipitation of waxes in pipelines due to solid phase behavior of n-alkanes (Lucia *et al*., 2005).

2. Molecular Structures and Polymorphisms

2.1. Triacylglycerol structure and fatty acids

Triacylglycerol molecules (TAGs) are responsible for more than 95% of vegetable oils composition. They are made by three fatty acids sterified to a glycerol backbone, as illustrate in Figure 1.

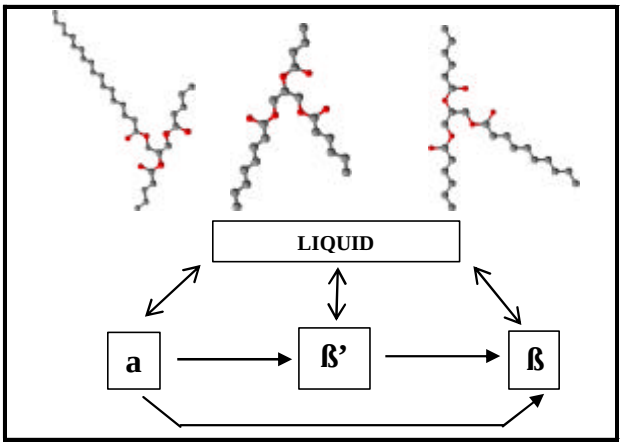


Figure 1: Triacylglycerol structures and state transitions.

Figure 1 shows a large variety of molecules according to the type of the 3 fatty acids attached to the glycerol. The carbon chain length, the number of double bounds and even the position of fatty acid in the three possible positions directly affect melting point, melting enthalpy, viscosity and solid miscibility in mixtures. As consequence, equilibrium between different species in liquid multi-solid phases is directly affected by these molecular issues.

2.2. Polymorphisms

TAG molecules present basically three different polymorphisms (Figure 1): unstable a, metastable ß' and stable ß, each one corresponding to a solid crystalline arrangement (Sato, 2001). Differences in molecular packing affect important properties related to solidification and melting behaviour, such as melting points and enthalpy of fusion

(molecule level), Gibbs energy, melting ranges and solid content (mixture level) as well as subjective properties related to fat crystal networks in edible products, as texture and spreadability.

3. Multi-phase Multi-Component Solid Liquid Equilibrium

3.1. General Modelling

Computing phase equilibrium is the solution of a nonlinear programming (NLP) problem searching for the global minimization of the total Gibbs energy subject to material balance constraints that can be also represented as an equivalent set of nonlinear equations. Experimental and theoretical backgrounds allow considering liquid and alpha solid phases as ideal (Wesdorp, 1990; Bruin, 1999; Himawan *et al*., 2006). Thus, only ß' and ß solid phases need description with an excess Gibbs free-energy model. The Margules equations are used in this work. Interaction parameters for the model were obtained by using the isomorphism based correlations, as illustrated by Wesdorp (90).

3.2. Stability Analysis

As the number of phases coexisting in equilibrium is not known *a priori*, the flash calculation is performed in two steps: stability analysis and calculation of phase fraction and equilibrium compositions. McDonald and Floudas (1997) demonstrated that while one can use global optimization to solve the NLP problem, it is computationally more efficient to use a two-stage approach, since the dimensionality of the global optimization problem in phase stability is less than that of the full equilibrium problem.

Stability tests concern the evaluation if a phase can split in two other phases, given T, P and overall composition. In this work, we use the tangent plane criterion for stability problem, presented by Michelsen (1982a,b): a necessary and sufficient condition for stability of a mixture at fixed temperature, pressure and overall composition (**z**) is that the Gibbs energy surface be at no point below the plane tangent to the surface at **z**. This condition for n components mixture is expressed as follows:

$$D(x) = m(x) - m_0 - \sum_{i=1}^{n} \left(\frac{\partial m}{\partial x_i} \right)_0 (x_i - z_i) \geq 0 \quad \forall x_i \tag{1}$$

Where $m(x) = \sum_{i=1}^{n} x_i \ln x_i + g^E(x)$ and $g^E(x) = G^E / RT$ is the reduced molar excess Gibbs energy. The subscript zero indicates evaluation at $x = z$. D(x) is the tangent plane distance function, which represents that if there is any possible trial phase with composition x to be added with negative value of D(x), it is possible to have a negative variation of Gibbs free energy (initial phase is unstable and splits). As all of the minima in D(x) are located in the interior of the permissible region ($x_i = 0$, ? $x_i = 1$) (Michelsen, 1982a), the value of D(x) will be nonnegative if it is nonnegative at all stationary points.

For achieving stationary points, Michelsen proposed an iteration procedure from a several initial guesses. The adaptation of this method to cope with polymorphisms is done by increasing the number of initial guesses: one pure phase of each component in each polymorphism and mixtures with compositions evaluated as an average of the mixtures already present (Wesdorp, 1990), leading to a 3n+4 initial guesses.

4. Results

Table 1 shows some stability tests performed in a mixture composed by 4 mono-TAGs formed by the fatty acids Palmitic, Stearic, Myristic and Oleic. The overall composition

(Z) and their physical state, temperature, stationary points and tangent plane distance are shown together with the physical state of the new phase that will be formed in case of instability (most negative value of tangent plane distance).

Table 1: Stability tests in a four component TAG systems.

Feed (Z1,Z2,Z3,Z4)	Temperature (°C)	Stationary Points (x1,x2,x3,x4)	D(x)/RT	Phase Split
ß' (0.25,0.25,0.25,0.25)	80	(0.1135, 0.1713, 0.5438, 0.1713)	-0.9610	Liquid
		(0.1598, 0.0441, 0.7520, 0.0441)	-0.3044	beta'
		(0.1369, 0.4103, 0.0424, 0.4103)	-0.0417	beta'
		(0.0546, 0.0058, 0.9338, 0.0058)	-0.4127	beta'
		(0.9170, 0.0227, 0.0376, 0.0227)	1.1466	Beta
		(0.0035, 0.4980, 0.0005, 0.4980)	0.1060	Beta
		(0.0011, 0.0001, 0.9986, 0.0001)	-0.3533	Beta
ß' (0.1, 0.1, 0.1, 0.7)	35	(0.0555, 0.0330, 0.6802, 0.2313)	-1.2008	Liqud
		(0.0223, 0.0010, 0.9701, 0.0067)	-0.8398	beta'
		(0.0880, 0.1049, 0.0728, 0.7343)	0.0012	beta'
		(0.0196, 0.0008, 0.9743, 0.0053)	-0.8402	beta'
		(0.6140, 0.0110, 0.2983, 0.0768)	1.3439	beta
		(0.0022, 0.1246, 0.0009, 0.8723)	0.1268	beta
		(0.0004, 0.0000, 0.9994, 0.0001)	-0.8160	beta

Systems composed by 9 triacylglycerols were used to simulate melting curves. They contain saturated as well as unsaturated molecules with different chain lengths. Mixture 1 is composed by the followings TAGs: MMM, PPP, SSS, OOO, PPS, PSP, SSM, MMP and MMS. Figure 2 shows the melting curves for mixture 1 at three solid states. It can be noted that as molecules in a state have the lowest melting points, the a curve lies below the two other curves. For ß' and ß curves, there is a region (between 20°C and 55°C) where the behavior is inverted (ß curves below ß' curve), highlighting that solid liquid equilibrium can clarify different behaviors in TAG systems. For ß' and ß equilibrium, initial compositions given by stability analysis were used.

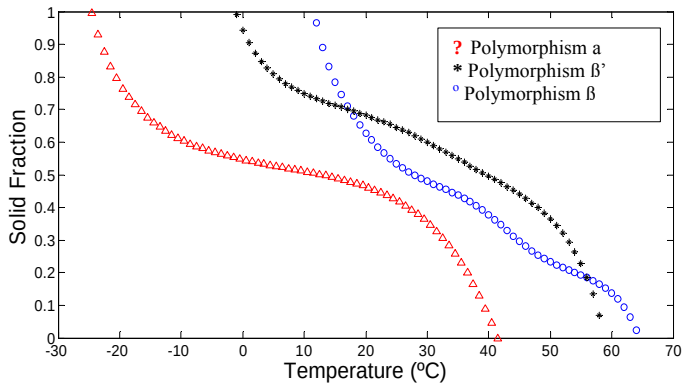


Figure 2: Melting curves for 3 polymorphic states (Mixture 1).

In Figure 3, mixture 1 was simulated using two different compositions, both in ß form: Composition 1 is [0.05;0.05;0.2;0.3;0.06;0.1;0.05;0.05;0.14] while composition 2 is [0.1;0.1;0.1;0.5;0.05;0.05;0.03;0.03;0.03]; therefore, in composition 2 half of the system is made by TAG OOO (formed by three oleic acids), which is highly unsaturated (low melting points). Thus, it can be observed the large influence of these unsaturations in the melting profile of the mixture (lower solid contents for all temperature range).
Resolution of solid-liquid equilibrium can also give useful information about the influence of TAG structures in the melting profile of the mixtures. In order to achieve this, mixture 1 was modified: TAG MMM (14:0) was substituted for LLL (12:0)-mixture 2 and for CCC (10:0)-mixture 3. As the carbon number decreases, the melting points are lower (for same degree of unsaturations); thus, a lower solid fraction for a given temperature is expected, what can be observed in Figure 4. This is true especially between 0°C and 20°C: in the final regions of solidification (more than 80 % of solid) and final region of melting (more than 70 % of liquid), the curves are very similar. Even in the middle regions, the difference is not large as the structure of these mixtures is quite similar. However, the curves reveal the sensibility of the model even for small changes in molecular structures.

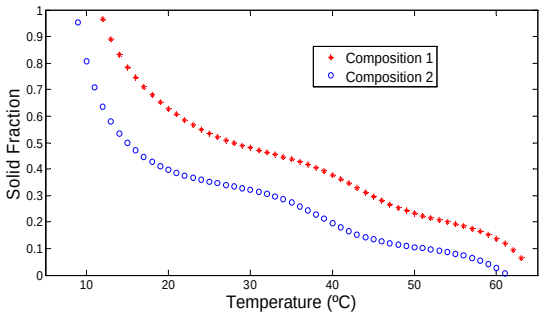


Figure 3: Influence of composition in melting curves (mixture 1; ß state).

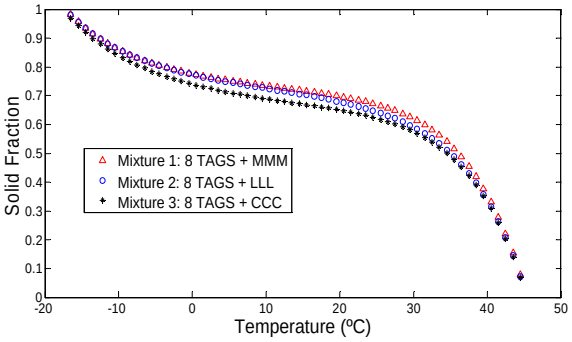


Figure 4: Melting curves for 3 mixtures differing of one TAG.

5. Conclusions

Melting and solidification behavior of triacylglycerol mixtures has a great importance for quality assessment of many products. The resolution of solid-liquid equilibrium with polymorphisms for triacylglycerol systems reveals useful information about melting and solidification profile of these complex mixtures. The effect of changes in mixture

composition or small changes in molecular structures can be observed by solid fractions in different temperatures. The technique has showed potential use for theoretical predictive capacity for new mixtures. However, the very large numbers of possible molecules that can be formed by fatty acids in a glycerol backbone requires the use of simulated melting curves coupled with optimization techniques in a Computer Aided Mixture Design environment for triacylglycerols. Besides this, other stability methods (as interval methods) must be tested in order to compare its robustness and the independence of phase split detection in relation to initialization, for using in multi-solid phase problems. Further DSC curves simulations can also be assessed by SLE, and experimental curves can be used as model validation.

The move towards the molecular level for product design is a major trend and challenge for Process System Engineering (PSE) in the future (Grossmann and Westerberg, 2000) and lipid-based products can take advantage of thermodynamic modelling of their phase behavior, allowing to achieve new products with improved functional properties and sustainability principles, as natural resources of triacylglycerols and fatty acids are renewable and provide alternative products using glycerol structure.

References

Bruin S., 1999, Phase equilibrium for food product and process design, Fluid Phase Equilibrium, 158–160, p. 657–671.

Bruin S., Th.R.G. Jongen, 2003, Food Process Engineering: The Last 25 Years and Challenges Ahead, Comprehensive Reviews in Food Science and Food Safety, 2, 42 – 81.

Grossmann I.E., A.W. Westerberg, 2000, Research Challenges in Process Systems Engineering, AIChE Journal, 46, 9, 1700–1703.

Himawan C., V.M. Starov, A.G.F. Stapley, 2006, Thermodynamic and kinetic aspects of fat crystallization, Advances in Colloid and Interface Science, 122, 3 – 33.

Inoue T., Y. Hisatsugu, R. Yamamoto, M. Suzuki, 2004, Solid–liquid phase behavior of binary fatty acid mixtures 1. Oleic acid/stearic acid and oleic acid/behenic acid mixtures, Chemistry and Physics of Lipids, 127, 143–152.

Liang B., Y. Shi, R.W. Hartel, 2003, Phase Equilibrium and Crystallization Behavior of Mixed Lipid Systems, JAOCS, 80, 4, 301–306.

Lucia, A., DiMaggio, P.A., Bellows, M.L., Octavio, L.M., 2005, The phase behavior of *n*-alkane systems, Computers and Chemical Engineering, 29, 2363–2379.

McDonald C.M., C.A. Floudas, 1997, GLOPEQ: a new computational tool for the phase and chemical equilibrium problem, Computers & Chemical Engineering, 21, 1–23.

Michelsen M.L., 1982a, THE ISOTHERMAL FLASH PROBLEM. PART I. STABILITY, Fluid Phase Equilibrium, 9, 1– 19.

Michelsen M.L., 1982b, THE ISOTHERMAL FLASH PROBLEM. PART II. PHASE-SPLIT CALCULATION, Fluid Phase Equilibrium, 9, 21– 40.

Sato K., 2001, Crystallization behaviour of fats and lipids - a review, Chemical Engineering Science, 56, 2255 – 2265.

Wasylkiewicz S.K., L.N. Sridhar, M.F. Doherty, M.F. Malone, 1996, Global Stability Analysis and Calculation of Liquid-Liquid Equilibrium in Multicomponent Mixtures, Ind. Eng. Chem. Res., 35, 1395–1408.

Wesdorp L.H., 1990, Liquid – multiple solid phase equilibrium in fats, Thesis. Delft University of Technology, Netherlands.

Won K.W., 1993, Thermodynamic Model of Liquid-Solid Equilibrium for Natural Fats and Oils, Fluid Phase Equilibrium, 82, 261-273.

Zhang L., S. Ueno, S. Miura, K. Sato, 2007, Binary Phase Behavior of 1,3-Dipalmitoyl-2-oleoyl-sn-glycerol and 1,2-Dioleoyl-3-palmitoyl-rac-glycerol, JAOCS, 84, 219 – 227.

Zhou Y., R.W. Hartel, 2006, Phase Behavior of Model Lipid Systems: Solubility of High-Melting Fats in Low-Melting Fats, JAOCS, 83, 6, 505–511.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Design of T-shaped microreactors by reduced-order approach

Osamu Tonomura, Masato Kubota, Manabu Kano, Shinji Hasebe

Dept. of Chemical Engineering, Kyoto University, Nishikyo-ku, Kyoto 615-8510, Japan
E-mail: tonomura@cheme.kyoto-u.ac.jp (O. Tonomura)

Abstract

The final goal of this research is to establish a systematic design method of T-shaped microreactors (T-MRs) with secondary flow using a reduced-order model (ROM). In this paper, a ROM which combines computational fluid dynamics (CFD) model and multi-resolution lamellar model is developed. Multi-resolution lamellar model is constructed from the *N* kinds of lamellar models with different width of layers. In each lamellar model, thin multi-layers of two fluids are arrayed alternately at the entrance of the mixing channel to express the increased material interface, and the concentration profiles in the flow direction are estimated by a reaction diffusion simulation between thin layers under the assumption of plug flow. The developed ROM requires the minimum number of CFD pre-simulations to extract model parameters such as weighting coefficients for superimposing different lamellar models. The main findings are: 1) the complicated mass transport processes of T-MRs obtained by 3-D CFD model is well captured by the ROM of 2-D, which is constructed from multi-resolution lamellas, and 2) the computational time based on the ROM is shortened to one-third of that based on just CFD model.

Keywords: microreactors, design, model order reduction, computational fluid dynamics.

1. Introduction

Microreactors can achieve the higher yield and selectivity of a product owing to their advantages of rapid mixing, high heat transfer rate, and precise residence time control. Although various microreactors have been proposed so far, their design depends on trial and error. Unless this situation is improved, the design period cannot be shortened. In the previous work by the authors, the microreactor design method based on computational fluid dynamics (CFD) model was proposed. CFD simulation provides the highly accurate detailed information on temperature and concentration distributions in microreactors without conducting experiments. However, it is not very practical to repeat CFD simulation for designing microreactors because of its heavy computational load. From this viewpoint, model reduction is required to obtain computationally tractable and physically meaningful models.

The secondary flows appear at the junction of T-shaped microreactors (T-MRs) under high throughput operation, increase the contact area between the mixing species, and accelerate the mixing rate (Bothe et al., 2006). The efficient mixing of T-MRs is confirmed through CFD simulation, whose main conditions and results are shown in Fig. 1. In order to efficiently design T-MRs, it is not practical to directly apply CFD

simulation to the T-MR design problems. The development of a reduced-order model (ROM) of T-MRs with secondary flow is effective in reducing the computational time for designing the T-MRs. This paper shows results of study toward the final goal of establishing a ROM-based design approach of T-MRs.

2. T-shaped microreactor (T-MR)

The fluid dynamics around the junction of T-MRs is complicated due to the generation of secondary flow, and it is difficult to analyze it without CFD simulation. As shown in Fig. 1, the concentration distribution greatly changes from the cross section ① to ③. On the other hand, there are few differences between concentration distributions on cross sections ③ and ④. This simulation result suggests that the complicated flow develops and becomes laminar along the mixing channel of T-MRs and that the convective mixing changes to the diffusion mixing.

In this research, the CFD model and the lamellar model are used as ROM. Diffusion and reactions in the mixing channel are modeled as reaction-diffusion simulations between thin layers. The two-dimensional concentration distributions on the cross sections are modeled as the one-dimensional distributions with two kinds of thin layers arrayed alternately at the entrance of the mixing channel as shown in Fig. 2. Thinner layers cause more rapid mixing. The concentration profiles under laminar flow condition in the flow direction are estimated by the lamellar model based on the reaction-diffusion simulations. As the first step of building the lamellar model, CFD simulation of mixing and reaction around the junction of the T-MR with secondary flow is conducted. Then, the lamellar model, which can estimate concentration profiles along the mixing channel, is constructed. Equation (1) is the governing equation of reaction-diffusion simulation under the assumptions of uniform flow and no axial diffusion.

$$u\frac{\partial C_i}{\partial x} = D\frac{\partial^2 C_i}{\partial y^2} + r_i \tag{1}$$

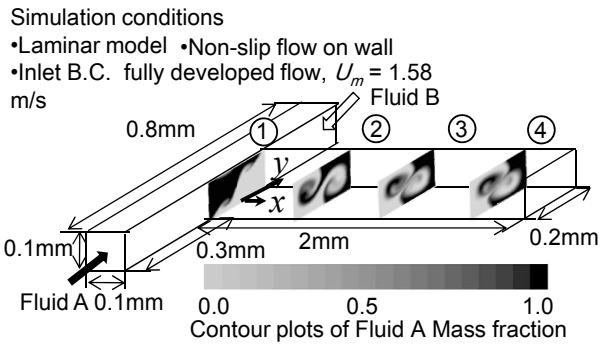


Fig. 1 Simulation conditions and results.

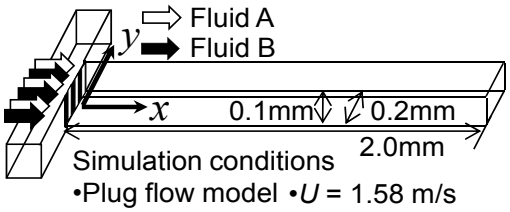


Fig. 2 Simulation of diffusion and reactions between the thin layers.

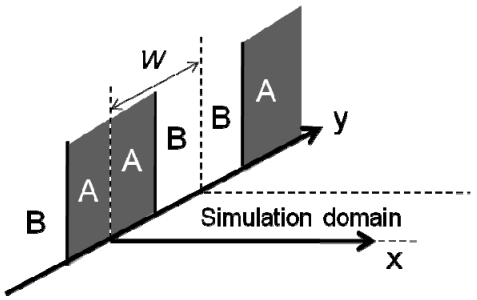


Fig. 3 Reaction-diffusion simulation domain.

where C_i denotes concentration of material i, x the position in the flow direction, y the position in the direction perpendicular to the flow direction, u flow velocity, D diffusion coefficient, and r_i reaction rate of material i. Equation (2) shows the boundary conditions of the reaction-diffusion simulation when width of layers is equal to W as shown in Fig. 3.

$$\text{B. C.} \quad \begin{aligned} &\frac{\partial C_i}{\partial y} = 0 && \text{at } y = 0, W \\ &C_A = C_{A0}, C_B = 0 && \text{at } x = 0, 0 < y \le W/2 \\ &C_A = 0, C_B = C_{B0} && \text{at } x = 0, W/2 < y < W \end{aligned} \tag{2}$$

where C_{i0} denotes concentration of material i at the entrance of mixing channel. The diffusional mixing and reactions in one-dimensional concentration distribution with single width of layers are analyzed by solving Eq. (1) under the boundary conditions shown in Eq. (2). In this research, Eq. (1) is solved with gPROMS®. However, the concentration distributions on the cross sections of T-MRs cannot be modeled with the arrayed layers having the same width because they consist of various scales of segregation. In order to express the concentration distributions, the lamellar model is constructed from the N kinds of reaction-diffusion simulations with single width of layers. In each simulation, the width of layers is different from those in the other simulations. The profiles of mass fractions of the lamellar model are given by Eq. (3):

$$F_{\mathrm{LM}}(i,x) = \sum_{n=1}^{N} a_n f_n(i,x) \tag{3}$$

$F_{LM}(i, x)$ denotes mass fraction of material i at position x in the flow direction. $f_n(i, x)$ denotes mass fraction of material i at position x in the flow direction in n-th reaction-diffusion simulation, where the width of layers is equal to W_n. a_n denotes the weighting coefficient of the result of n-th reaction-diffusion simulation. The values of weighting coefficients are determined as the solution of the minimization problem defined by Eq. (4):

$$\min_{a_n} \quad J = \frac{\sum_{i,k} \sqrt{\left(F_{CFD}(i, x_k) - F_{LM}(i, x_k)\right)^2}}{I\ K} \qquad \text{s. t.} \quad 0 \le a_n \le 1, \quad \sum_{n=1}^{N} a_n = 1 \tag{4}$$

J is the mean absolute error of mass fractions between the result of CFD simulation and the result estimated by the lamellar model. $F_{CFD}(i, x_k)$ denotes the CFD simulation result of mass fraction of material i at position x_k in the flow direction. I denotes the number of chemical species. K denotes the number of points in the flow direction for calculating the errors of mass fractions. The positions in the flow direction for calculating the errors of mass fractions have to be chosen. It is difficult to express the convective mixing and reaction with reaction-diffusion simulations between thin layers. Therefore, the positions for calculating the errors of mass fractions have to be where there is little effect of secondary flow on mixing and reactions.

3. Case study

The validity of the lamellar model is confirmed through a case study. At first, the lamellar model of the T-MR is constructed from the result of CFD simulation analyzing T-MR with 2 mm long mixing channel as shown Fig. 1. Then, the profiles of mass fractions between 2 mm and 10 mm in the mixing channel length are estimated by the constructed lamellar model. The estimated profiles are compared with the results of CFD simulation of mixing and reaction in the TMR with 10 mm long mixing channel. As the reaction system, the following competitive-consecutive reaction scheme is considered.

$$\begin{aligned} &\mathrm{A} + \mathrm{B} \rightarrow \mathrm{R} \quad && r_1 = k_1 C_A C_B \quad && k_1 = 18\ \mathrm{m^3/(kgmol \cdot s)} \\ &\mathrm{B} + \mathrm{R} \rightarrow \mathrm{S} \quad && r_2 = k_2 C_B C_R \quad && k_2 = 15\ \mathrm{m^3/(kgmol \cdot s)} \end{aligned} \tag{5}$$

The physical properties of Fluids A, B, R, and S are same with those of water at 300 K. Molecular weights of A, B, R, and S equal to 18, 18, 36, and 54 respectively. The positions x_k selected to calculate J are given at intervals of 0.1 mm between x_k = 1.3 mm and x_k = 2.0 mm. The solution of the minimization problem defined by Eq. (4) is found through comparing the values of J among all combinations of values of weighting coefficients, which are assumed to range from 0 to 1 at the 0.05 interval.

The lamellar model based on five kinds of reaction-diffusion simulations, i.e. N = 5, W_1 = 1 μm, W_2 = 6 μm, W_3 = 15 μm, W_4 = 50 μm, and W_5 = 100 μm, is abbreviated as LM-5. In addition, the lamellar model based on single width layers, i.e. N = 1 and W_1 = 6 μm, is abbreviated as LM-1. The width of layers of the LM-1 is determined so that the value of J is the smallest of all lamellar models when N = 1. The profiles of mass

fraction R of LM-5 (F_R LM-5), LM-1 (F_R LM-1), and CFD simulation result (F_R CFD) are plotted in Fig. 4. In addition, the values of J and J^*, the mean absolute error of mass fractions between 2 mm and 10 mm in the mixing channel length, are listed in Table 1. J^* is calculated by Eq. (6):

$$J^* = \frac{\sum_{i,k^*} \sqrt{\left(F_{\mathrm{CFD}}(i, x_{k^*}) - F_{\mathrm{LM}}(i, x_{k^*})\right)^2}}{I\ K^*} \tag{6}$$

The positions x_k^* selected to calculate J^* are given at intervals of 1 mm between $x_k^* = 2$ mm and $x_k^* = 10$ mm. K^* denotes the number of the selected positions to calculate J^*. The weighting coefficients of LM-5 are listed in Table 2.

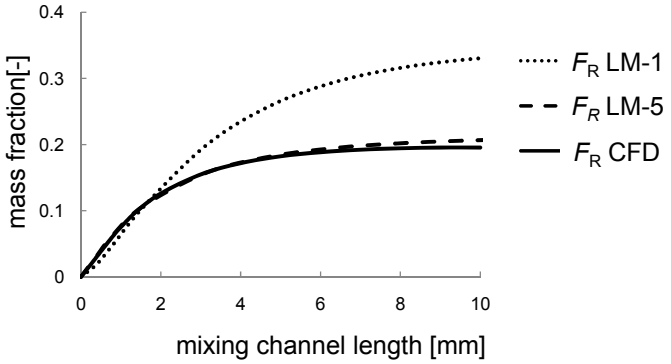


Fig. 4 Estimation results.

Table 1 Values of J and J^*

	LM-1	LM-5
$J \times 10^3$	4.53	2.04
$J^* \times 10^2$	8.76	0.95

Table 2 Weighting coefficients

n	W_n	a_n
1	1 μm	0.30
2	6 μm	0.20
3	15 μm	0.05
4	50 μm	0.20
5	100 μm	0.25

In Fig. 4, the profiles of F_R LM-5 (dashed line) and F_R CFD (solid line) are very similar between 2 mm and 10 mm in the mixing channel length, though the profile of F_R LM-1 (dotted line) is quite different from the others. Figure 4 clarifies that the model accuracy of LM-5 is much better that of LM-1. In addition, Table 1 shows that values of J and J^* of LM-5 are smaller than those of LM-1. That is, the model accuracy of the lamellar model is improved by using various kinds of widths of layers. Table 2 shows that various widths of layers are used in LM-5 to express the concentration distributions on the cross sections.

4. Conclusions

In this research, the ROM of T-MRs, which is based on the CFD model and the multi-resolution lamellar model, is proposed. The lamellar model is constructed from the results of the reaction-diffusion simulations to minimize the mean absolute error of mass fractions, J. Through the case study, it is confirmed that the lamellar model can estimate the profiles of mass fractions in the mixing channel of T-MRs accurately.

Acknowledgement

This research was partially supported by the New Energy and Industrial Technology Development Organization (NEDO), Project of Development of Microspace and Nanospace Reaction Environment Technology for Functional Materials.

References

D. Bothe, C. Stemich, and H. Warnecke, 2006, Fluid Mixing in a T-shaped Micro-Mixer, Chemical Engineering Science, 61, 2950–2958.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Novel Molecular Design Technique using Property Operators based on Signature Descriptors

Nishanth G. Chemmangattuvalappil, Charles C. Solvason,
Susilpa Bommareddy, Mario R. Eden
Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA

Abstract

In this work, we are introducing an algorithm that uses molecular signature descriptors for reverse problem formulations. This algorithm utilizes molecular property operators based on signatures for solving the reverse problem of obtaining the molecular structures corresponding to the property targets estimated during the process design step. The algorithm allows property models based on different Topological indices (TI) and group contribution models to be represented on the same property platform.

Keywords: Molecular design, Signature description, Reverse problem.

1. Introduction

1.1. Reverse problem formulation

Reverse problem formulation (RPF) is a technique used to reduce the complexity of integrated process and product design problems (Eden et al., 2004). The complexity of such problems is due to the non-linearity of property models. To overcome this issue, the single forward problem is broken down into two reverse problems. The first reverse problem provides the property targets to achieve optimum process performance and the second reverse problem is used to generate the molecular structures that meet the property targets. To track properties, property operators are used which are functions of the original properties tailored to obey linear mixing rules. The normalized property operator, Ω_{js} is obtained by dividing it by a reference value (Shelley and El-Halwagi, 2000). If $\psi_j(P_{js})$ is the property operator of the j^{th} property P_{js} of stream s, x_s is the fractional contribution, N_s is the number of streams:

$$\Omega_{js} = \frac{\sum_{s=1}^{N_S} x_s \psi_j(P_{js})}{\psi_j^{ref}} \qquad (1)$$

1.2. Molecular signature descriptors

Molecular signature is a molecular descriptor used for representing the atoms in a molecule using the extended valancies to a pre-defined height (Visco *et al.,* 2002, Faulon *et al.,* 2003a). If G is a molecular graph and x is an atom of G, the atomic signature of height h of x is a canonical representation of the sub graph of G containing all atoms that are at a distance h from x. The detailed procedure for generating signature can be found in the work published by Visco *et al.* (2002) and Faulon *et al.* (2003a). Since atomic signatures up to a given height are of finite size, any molecule can be uniquely represented with its atomic signatures. The signature of a molecule is defined

as the linear combination of its atomic signatures (Visco *et al.,* 2002, Faulon *et al.,* 2003a). The signatures produce meaningful QSPR/QSARs and accuracy of these descriptors is comparable to the existing TIs (Visco *et al.,* 2002, Faulon *et al.,* 2003a). The reason for this correlation is that many of the TIs can be derived from the molecular signature. Faulon *et al.,* (2003a) have provided the relationship between TIs and signatures:

$$TI(G) = k \cdot {}^{h}\alpha_G \cdot TI(root({}^{h}\Sigma)) \tag{2}$$

Here, k is a constant, ${}^{h}\alpha_G$ is the vector of occurrences of atomic signature of height h and TI(root (${}^{h}\Sigma$)) is the vector of TI values calculated for each root of atomic signature.

2. Reverse problem formulation using molecular signatures

Since many existing QSPR expressions can be reformulated in terms of molecular signatures, algorithms based on signatures will be able to track a wide variety of property targets. In addition, algorithms are available to obtain the molecular structures once the solution to RPF is available as signatures (Faulon *et al.,* 2003b).

2.1. General problem statement

Identify the non-cyclic molecules with the best dominant property, which also satisfy a set of property constraints. The first step is to identify the QSAR/QSPR expressions corresponding to all the given properties/activities. In the next step, identify the heights of molecular signature corresponding to the TIs used in QSPR. If θ is the property function of property *P,* the property operator corresponding to *P* is estimated as follows:

$$\theta = f(TI) \tag{3}$$

$$TI = \sum_{i=1}^{N} {}^{h}\alpha_i \cdot TI(root({}^{h}\Sigma)) = \sum_{i=1}^{N} {}^{h}\alpha_i L_i \tag{4}$$

$$\psi(P) = \sum_{i=1}^{N} x_i L_i \tag{5}$$

The dominant property, which is expressed in terms of the occurrences of atomic signatures, can be maximized or minimized subject to the property constraints. If Ω_j is the property operator corresponding to the dominant property and Ω_{ij} is the normalized property operator of molecule *i*, an optimization problem can be formulated as follows:

$$Max / Min \ \ \Omega_j \tag{6}$$

$$\Omega_j^{\min} \le \Omega_j \le \Omega_j^{\max} \tag{7}$$

To ensure that the molecule will be formed without any free bonds, the *handshaking lemma* from graph theory is used (Trinajstic, 1992). For a graph without any circuits, if *D* is the number of degrees of each vertex *i* and *V* is the number of vertices:

$$\sum_{i=1}^{N} D(i) = 2(V-1) \tag{8}$$

In order to differentiate between different types of atoms, vertex coloring has been used. The coloring function has to be selected based on the types of atoms and the nature of the final molecule. For instance, in the design of an alkane molecule, the coloring function is the degree of carbon atoms at all levels. The coloring should start from the root atom to all atoms up to level *h-1*. An example is shown below:

C(C(CC(CCC))CC(C(C)C(CC)C))→ C3(C3(C1C4(CCC))C1C4(C2(C)C3(CC)C1))

In a complete molecule, since each edge is shared by two vertices, the colors of the edge that joins the two vertices must be the same for both vertices. However, the order of colors will be different for both the vertices since the reading of color has to start from the root. For instance, consider the signatures, C1(C2) and C2(C1). Here, the reading of colors is 1→2 in the first signature where as it is 2→1 for the second signature. The presence of both edges ensured that there is a linking between the vertices. While writing the atomic signature, only one atom is being described by relating it to its neighboring atoms at different levels of neighborhood. Therefore, every color sequence of edges between two heights must be complemented by another signature having one edge with the same colors in the reverse order. If $(l_i \rightarrow l_j)_h$ is one coloring sequence $l_i \rightarrow l_j$ at a level *h,* then:

$$\sum (l_i \rightarrow l_j)_h = \sum (l_j \rightarrow l_i)_h \tag{9}$$

If the color sequence in one signature is $l_i \rightarrow l_j$ with $i=j$, then, there must be another signature present in the set of signatures with the same color sequence to complement the previous one. If η is the number of color sequences $l_i \rightarrow l_j$ on one signature with $i=j$ and x is the number of such color sequences and K is an integer, then:

$$\sum_{i=j} \eta_i x_i = 2K \tag{10}$$

In order to form a connected tree, it must be ensured that the total number of signatures, where the degree of the vertex at a higher level is more than the degree at a lower level should be less than the total number of vertices with the higher degree. If n_i is the number of child vertices with a higher degree than the parent vertex, i and j represents the child and parent colors:

$$\sum x_i n_i \leq \sum x_j \tag{11}$$

The dominant property function can be maximized or minimized subjected to the constraints. The signatures obtained will form the best molecule. In order to form other feasible molecules, integer cuts can be used. The algorithm developed by Faulon *et al.* (2003b) can be used to enumerate the molecules from the identified signatures.

2.2. Expression of group contribution (GC) models with signatures

In GCM, the property function of a compound is estimated as the sum of property contributions of all the molecular groups present in the molecular structure. The molecular signatures of sufficient height can be used to re-write GC expressions. This transformation will allow us to solve the property models based on TIs and GC models simultaneously. The height and number of the signature used to write the molecular groups in the GC model depends on the number of atoms and the nature of final molecule. Consider the design of a molecule with amine and alkyl groups. The maximum number of amine groups on each alkyl group and attached carbon atom on

each amine group is limited to one. The potential groups available in GCM are (Marrero and Gani, 2001):

CH_2 NH_2 CH_2NH_2 $CHNH_2$ CNH_2 CH_3 CH_2 CH C

The root atom on every signature can be colored with two numbers - the first color is the number of neighboring C atoms and the second color is the number of neighboring N atoms. Signature of height 2 is required for complete coloring. All signatures with root N can be used to re-write different amine groups. If the neighboring C atom of an N root has three neighbors, then, the root N atom will be equivalent to a CNH_2 group. Similarly, all N root signatures can be assigned the property contributions of CH_2NH_2 and $CHNH_2$ groups based on the colors of their nearest neighbors. The signatures with root C atoms will form the alkyl groups. Here, the signatures with root C atom having N in any of its nearest neighbor should not be considered as a group since that signature has been taken care of in the amine groups. For the rest of the signatures with root C, property contributions can be assigned based on the numbers of neighboring C atoms.

2.3. Property models with different signature heights

In a molecular design problem with multiple property constraints, different TIs may be describing different properties of interest. It is also possible to have one QSAR/QSPR containing different TIs. If the signature heights of the TIs are different, the signatures corresponding to the largest height have to be enumerated first and the number of signatures of smaller height can be obtained from the higher signatures. Consider the situation where QSPR is based on both zero and first order connectivity indices. The corresponding signatures are of heights one and two respectively. The property operator for property *Y* can be written in terms of signatures as follows:

$$Y = a_0{}^0\chi + a_1{}^1\chi \tag{12}$$

$$f(Y) = \sum_i L_i{}^1\alpha_i + \sum_j L_j{}^2\alpha_j \tag{13}$$

If the height of the largest signature of interest is *h* and height of the lower signature is *h-m,* then the total number of each *h-m* level signatures can be obtained by adding the *h* level signatures under same color at level *h-m*. For example, if the signatures of interest are N1(C), N2(CC) and N3(CCC) which are signatures with height one. Now, signatures of height two with root vertex N can be divided into three sets S_1, S_2 and S_3. The elements in these sets are signatures of height two with N vertex and vertex color 1, 2 and 3 respectively. The occurrences of signature N1(C) can be obtained as follows:

$$N(C):\ {}^1\alpha_{N(C)} = \sum_{S_1} {}^2\alpha_i \tag{14}$$

The number of occurrences of other signatures of height one is similarly estimated.

3. Case Study: Design of alkyl substituent for a fungicide

The optimal substituent selection for dialkyldithiolanylidenemalonate (DD), which was originally solved by Raman and Maranas (1998) has been reworked here. The target properties are affinity (*log* (V_E)), mobility (*log* (μ)) and retention (*log* [*R/(1-R)*]). These properties are correlated to hydrophobic factor (Uchida, 1980), which is correlated to a first order molecular connectivity index (Murray *et al.,* 1975). In this work, an additional property constraint of toxicity (LC_{50}) is included. The GCM can be used to

predict LC_{50} (Martin and Young, 2001). The property models are shown in Table 1. The objective of this case study is to identify the alkyl substituents of DD that give maximum affinity subject to the constraints based on mobility, retention and toxicity. The first order connectivity index can be re-written in terms of molecular signatures of height two (Visco *et al.,* 2002) according to Eq. 15:

$$ {}^{1}\chi = \sum_{i=1}^{K_G}\left(\frac{1}{2} \sum_{u \in {}^{l}V_2({}^{h}\chi_i)} \left[\deg(u) \deg({}^{-1}\nu_\sigma(u)) \right]^{-\frac{1}{2}} \right) x_i = \sum_{i=1}^{K_G} L_i x_i \qquad (15) $$

It is possible to form 65 signatures of height 2 with three signatures are corresponding to the CH_3- group, nine signatures corresponding to the CH_2- group, nineteen corresponding to the CH- group and thirty-four corresponding to the C- group. The property targets and the normalized property operators are shown in Table 1. The molecular design problem can be written in terms of signatures using Eqs. 16-19.

Table 1. Property operators and targets

Property	*Property model*	*Property operator*	*Ref.*	*L.B.*	*U.B.*
Mobility	$-0.6984({}^{1}\chi) + 2.0143$	$1.432(2.0143 - \log(\mu))$	1	-0.3	0.3
Retention	$0.787({}^{1}\chi) - 2$	$1.271\left(\log\left(\frac{R}{1-R} \right) + 2 \right)$	1	-0.3	1.0
LC50	$\exp\left(-\sum_{i=1}^{N} n_i \alpha_i \right)$	$-\log(LC_{50})$	1	-	0.04
Affinity	$0.5751({}^{1}\chi) - 0.2942$	$1.739(\log(V_E) + 0.2942)$	1	-	-

$$ Max\, \Omega_{V_E} \qquad (16) $$

$$ 2.455 \le \sum_{i=1}^{65} h_i x_i \le 3.314 \quad \text{(Includes the constraints on mobility and retention)} \qquad (17) $$

$$ 3.129 \le \sum_{i=1}^{3} c_1 x_i + \sum_{i=4}^{12} c_2 x_i + \sum_{i=13}^{31} c_3 x_i + \sum_{i=32}^{65} c_4 x_i \quad \text{(Constraint on LC50)} \qquad (18) $$

$$ \sum_{i=1}^{3} x_i + 2\sum_{i=4}^{12} x_i + 3\sum_{i=13}^{31} x_i + 4\sum_{i=32}^{65} x_i = 2\left[\left(\sum_{i=1}^{65} x_i \right) - 1 \right] \quad \text{(Handshaking Lemma)} \qquad (19) $$

To differentiate among different types of carbon atoms, the vertices are colored with the degree of each carbon atom. Equation 9 is applied to ensure that the signatures will complement each other. For illustration, one of the equations from Eq. 9 is shown:

$$ x_1 = x_4 + x_5 + x_6 \qquad (20) $$

Here, x_1, x_4, x_5 and x_6 are the occurrences of the signatures C1(C2(C)), C2(C2(C)C1), C2(C3(CC)C1) and C2(C4(CCC)C1). Equation 10 is used to ensure the consistency of signatures with same color sequence on same edges. An example is shown in Eq. 21:

$$x_4 + 2x_7 + x_8 + x_9 = 2K \quad (21)$$

Equation 11 ensures that the number of any color in the child level would not exceed the total number of that color when it is in parent level. One example is shown in Eq.22:

$$2x_{11} < \sum_{3\rightarrow 2} x_i \quad (22)$$

Here, x_{11} is the number of occurrences of the signature C2(C3(CC)C3(CC)). Since this signature contains two 2→3 color sequences, the total number of signatures with 3→2 sequence should be more than two times x_{11} to ensure the connected graph. Equations 16-22 are solved and the molecular structures are enumerated from the signatures using the algorithm provided by Faulon *et al.* (2003b). The solution and the estimated properties are shown in Table 2. This is the same solution as found by Raman and Maranas (1998). There are only four feasible solutions while the fifth solution in the original work is no longer valid because of the additional constraint on toxicity.

Table 2. Final solution

Affinity	Mobility	Retention	Toxicity	R_1	R_2
				methyl	3-methyl-butyl
1.608	-0.296	0.603	0.035	methyl	2-pentyl
				ethyl	*sec*-butyl
				methyl	*iso*-pentyl
1.586	-0.269	0.574	0.035	ethyl	*iso*-butyl
				n-propyl	*iso*-propyl
1.535	-0.207	0.503	0.040	methyl	2-methyl-2-butyl
1.504	-0.169	0.460	0.040	*iso*-propyl	*iso*-propyl

4. Conclusions

In this work, the molecular signature descriptors have been included in the RPF framework. The developed algorithm can include different QSAR/QSPR expressions based on multiple TIs thereby increasing the applicability of RPF. GC methods can be coupled with TI based expressions in the new algorithm. Since all property models are expressed as a function of occurrences of signatures, the problem can be solved for a single descriptor even if the TIs are described with signatures of different heights.

References

M.R. Eden, S.B. Jorgensen, R. Gani and M.M. El-Halwagi, (2004), *Chemical Engineering & Processing* 43, 595-608

J.L. Faulon, D. P. Visco Jr., R. S. Pophale, *J. Chem. Inf. Compt. Sci.* 2003a, 43, 707-720

J.L. Faulon, C.J.Churchwell, D. P. Visco Jr., *J.Chem.Inf.Compt. Sci*. 2003b, 43, 721-734.

J. Marrero and R. Gani, (2001), *Fluid Phase Equilibria*, 183/184, 183-208.

T.M. Martin and D.M.Young,(2001), *Chem.Res.Toxicol.* 2001,14,1378-1385.

W.J. Murray, L.H. Hall and L.B. Kier (1975), *J. Pharmaceutical Sci.* 64, 1978-81.

V. S. Raman, C. D. Maranas, *Computers & Chemical Engineering* 1998, *22* (6), 747-763.

M.D. Shelley and M.M. El-Halwagi, (2000), *Computers & Chemical Engineering*, 24.

N. Trinajstic, Chemical Graph Theory, 2nd Edition, CRC press, Boca Raton, FL, 1992

M. Uchida, *Pesticide Biochemistry and Physiology* 1980, 14, 249.

D. P. Visco Jr., R. S. Pophale, M. D. Rintoul, J.L. Faulon, *Journal of Molecular Graphics and Modeling* 20 (2002) 429-438

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Process integration aspects of the design of a gas separation system for the upgrade of crude Synthetic Natural Gas (SNG) to grid quality in a wood to methane process

Martin Gassner, François Maréchal

Laboratory for Industrial Energy Systems, Ecole Polytechnique Fédérale de Lausanne, CH – 1015 Lausanne, Switzerland

Abstract

The present paper investigates the prospects of process integration of the biogenic production of crude synthetic natural gas (SNG) and its upgrade to grid quality. At the example of a separation by means of a membrane cascade, a holistic design approach targeting the overall process performances is presented. Compared to a design obtained from an isolated approach, it is shown that a considerable reduction of the size and cost of the separation system is possible if a tight process integration is considered in the system design.

Keywords: SNG, process integration, CO2 separation, membrane system design

1. Introduction

Among the possible routes for producing fuels from biomass, thermochemical processing is one of the most promising options since a complete conversion of lignocellulosic biomass is achieved. In these processes, the reactive steps are performed at high temperature and a viable process integration is crucial to obtain an efficient conversion. Several recent studies have investigated suitable technology and processes for SNG production (Mozaffarian and Zwart, 2003; Duret et al. 2005; Gassner and Maréchal, 2005; Heyne et al., 2008; Luterbacher et al., 2008), the necessary gas upgrading step has however not received much attention. The present paper investigates this issue in more detail and aims to identify the advantages of a tight integration of the reactive steps with a gas separation system by membranes.

2. Problem statement

2.1. Crude SNG production

Representing the biomass as a chemical molecule with the carbon atom as reference, the conceptual design of its conversion into methane is based on Equation (1):

$$CH_{1.35}O_{0.63} + 0.3475\,H_2O \xrightarrow{\Delta\tilde{h}_r^0 = -10.5\,kJ/mol_{wood}} 0.51125\,CH_4 + 0.48875\,CO_2 \qquad (1)$$

Technically, the most envisaged route is to gasify the biomass and convert the producer gas into methane. Prior to gasification, the raw material is dried to below 20-25%wt humidity in order to prevent excessive losses due to water evaporation. From the current state of research and process development (Mozaffarian and Zwart, 2003; Stucki, 2005), it is expected that the first installations will be based on indirectly heated fluidised bed gasification of FICFB-type that has been developed and commercialised by (Hofbauer

et al., 2002). Due to the presence of dust, tars and catalyst poisons like sulphur compounds, the producer gas is cleaned before entering a catalytic fluidised bed reactor, where it is converted at 300-400°C to methane.
The stoechiometric equation (Eq. 1) shows that the overall conversion of wood to methane is exothermic and releases about 450 kJ/kg_{wood} of heat. This net release is distributed over the gasification and methanation reactions. The first one is endothermic and requires heat at high temperature, whereas the second one is exothermic and releases heat at lower temperature. The process thus requires additional energy and the quality of the process integration will define the overall process efficiency. The process is pinched at the gasification temperature and about 200 - 250 kW of high temperature heat at 850 – 900°C are typically required to convert 1 MW of wood to crude SNG.
The expected gas compositions of the producer gas and the crude methanation product computed with a previously developed process model (Duret et al., 2005; Gassner and Maréchal, 2005) are reported for methanation at atmospheric pressure in Table 1.

	C_2H_4	CH_4	H_2	CO	CO_2	N_2	H_2O
Gasification	2.2	10.9	38.8	25.5	18.6	0.5	3.5
Methanation	-	44.8	5.9	0.1	45.1	1.0	3.1
- without H_2O, CO_2	-	86.4	11.3	0.3	-	2.0	-
- without H_2O, CO_2, H_2	-	97.4	-	0.3	-	2.3	-
Grid specifications		> 96	< 4	< 0.2		< 6	-8°C

Table 1: Cold gas compositions and upgrading requirements (in %vol).

2.2. Design problem definition
According to the new Swiss directive for the supply of biogas to the natural gas grid, unlimited amounts of gas can be fed to the grid if its methane content is higher than 96%vol and the CO_2, H_2 and CO content less than 6, 4 and 0.2%vol respectively, while having a dew point at grid pressure (50 bar) lower than -8°C (SVGW, 2007). Assuming that nitrogen can not be easily separated from methane, Table 1 also shows the maximum obtainable purities after complete removal of water, carbon dioxide and hydrogen. Provided that an ideal separation process for removing CO_2 from the SNG is used, the hydrogen concentration is expected to rise to around 11%. The CH_4 purity will be at best around 86%, and at least some of the unconverted hydrogen must be removed.
The identification of the minimum energy requirements of the process furthermore adds a supplementary aspect to the design problem. Gasification has a relatively important heat demand at high temperature, which is usually supplied by burning producer gas and therefore reduces the methanation flow. In the combined heat and power application for which the gasification technology has originally been developed, few other fuel alternatives are present. However, in an SNG plant that necessarily suffers from a non-ideal separation, depleted gas streams from the SNG upgrading section could be used instead of the intermediate product. The condition for this is that its flame temperature is sufficiently higher than the gasification temperature. If the depleted gas is too diluted, it must in any case be treated in a catalytic combustion to eliminate the residual methane due to its high environmental impact as a greenhouse gas.
From these considerations, it is obvious that the design of the separation system is not a trivial problem. It requires the development of an appropriate multicomponent separation system, while also considering possible cost and energy savings originating from the trade-off between using the producer gas or a depleted gas as hot utility. In the framework of the actions aimed to climate change mitigation, rather pure carbon

dioxide (95% minimum) may also represent a valuable by-product. If sequestration instead of venting the separated carbon dioxide is targeted, care must be taken not to dilute this by-product with nitrogen from combustion with air. A possible option is thereby to use oxygen produced by electrolysis in the catalytic combustion, whose by-produced hydrogen could be supplied to the methane synthesis and would increase the SNG output (Gassner and Maréchal, 2008b). A general block flow diagram of these different upgrading options is given on Figure 1.

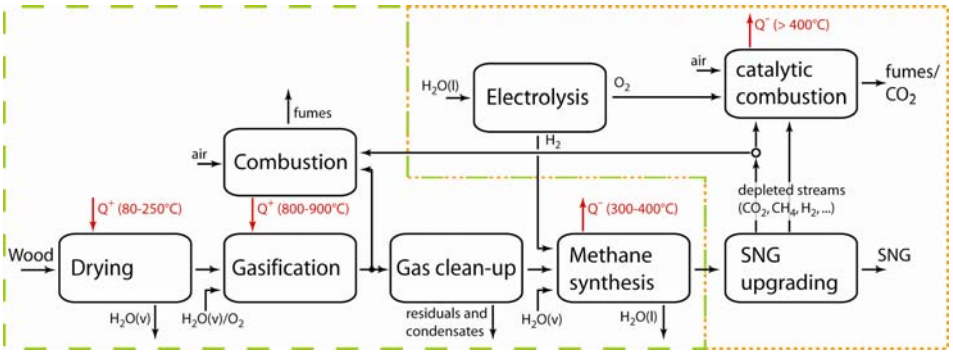


Figure 1: Block flow diagram of the crude SNG production (dashed) and upgrading (dotted) units.

3. Process integration study

3.1. Modelling

According to the applied design methodology (Gassner and Maréchal, 2008a), its thermodynamic part is divided into an energy-flow and an energy-integration model. In the energy-flow model, the thermodynamic conversions occurring in the different process units are modelled with a commercial flowsheeting tool (Belsim SA, 2007). The heat requirement of the process units is transferred to the energy-integration model, where the energy conversion and heat transfer system is determined by a mixed integer linear programming model targeting minimum operating costs with respect to the minimum approach temperature constraints. The thermodynamic conditions are then considered as design targets for the process equipment, which is then rated and costed by design heuristics and data from existing experimental and pilot plant facilities.

In order to investigate the impact of process integration on the optimal design of the separation system and the reactive sections of the plant, multi-stage membrane separation has been chosen among the identified candidate technologies for a detailed modelling. According to the theory of gas separation by membranes, a practical description of the permeation process through a membrane is to consider the difference of partial pressure in the bulk as driving force (Hwang and Kammermeyer, 1975). Using a phenomenological constant termed permeability P_i, the permeation of species i is described as:

$$\frac{d\dot{n}_{i,p}}{dA} = \frac{P_i}{\delta} \cdot \left(p_{i,r} - p_{i,p}\right) \qquad (2)$$

where $\dot{n}_{i,p}$ is the partial molar flow of i that permeates the membrane, A the membrane area, δ its thickness, and $p_{i,r}$ and $p_{i,p}$ the partial pressure on the retentate and permeate side, respectively. The selectivity $\alpha_{i,j}$ of the membrane between two different species i and j is defined by the ratio of their respective permeabilities P_i/P_j.

As the permeation is described by a set of *i* differential equations of the same type as Eq. 2, its resolution depends on the flow pattern of the membrane. Except for the simple case of perfectly mixed permeate and retentate chambers, no analytic solutions have been reported for multicomponent systems, but some simplified analytic or numeric procedures have been suggested. For flowsheet calculations, a simplified algebraic design model for hollow-fiber modules in counter-current operation developed and validated by Pettersen and Lien (1994) proves to be appropriate and has been implemented in the process model.
With the properties of commercially available membranes (Table 2), a single membrane stage is generally not sufficient to reach the required purity and an acceptable recovery of the crude SNG. Several membrane stages must therefore be cascaded, and a multitude of possibilities for the feed location, membrane connections and intermediate product withdrawals to be used as hot utility exist. From a first screening of all the possible cascades schemes, a superstructure of the most promising flowsheets has been defined (Figure 2) and is subjected to the process optimisation.

P_{CO2}	9 barrer	$\alpha_{CO2,H2}$	3.42 -
δ	1000 Å	$\alpha_{CO2,CH4}$, $\alpha_{CO2,CO}$, $\alpha_{CO2,N2}$	21.1 -

Table 2: Properties of commercial cellulose acetate membranes (Bhide and Stern, 1993; Phair and Badwal, 2006). 1 barrer = 7.5e-18$m^2/s^{-1}Pa^{-1}$.

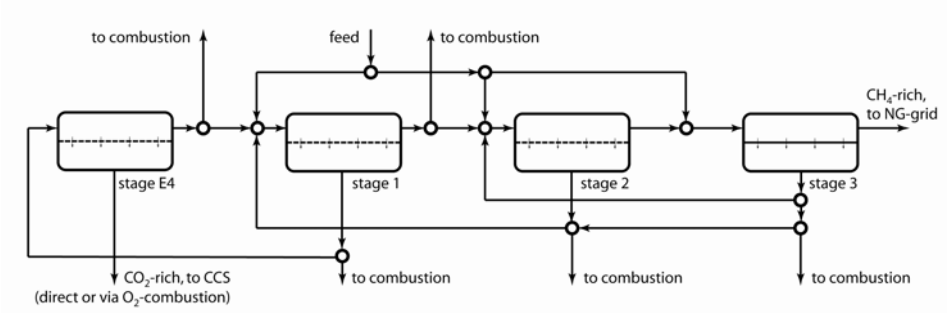


Figure 2: Superstructure of the membrane system.

3.2. Comparison of different levels of process integration

3.2.1. Approach

In order to investigate the influence of the level of integration on the process performance, the separation system is optimised by applying two different strategies. In a first approach, only the isolated performance of the separation system is considered in the optimisation, which represents a sequential design of the production and separation steps without considering the possible synergies resulting from the integration. For this purpose, the thermo-economic indicators are defined to only assess the performance of the separation system. According to its system boundary represented by the dotted line on Figure 1, the performance of the separation is evaluated by its energy efficiency (Eq. 3) and the total cost for the crude SNG separation (Eq. 4):

$$\varepsilon^{sep} = \frac{\Delta h^0_{SNG}\dot{m}^-_{SNG}}{\Delta h^0_{crudeSNG}\dot{m}^+_{crudeSNG} + \dot{E}^{sep,+}} \quad (3); \qquad C^{sep}_p = C^{sep}_{I,d} + C^{sep}_M + C^{sep}_{el} + C^{sep}_{RM} \quad (4)$$

In these equations, Δh^0 terms the heating value of the streams, $\dot{m}$ their mass flows and $\dot{E}$ the electricity consumption. The superscript ^{sep} stands for the separation system, and $^-$ and $^+$ refer to positive output and input terms, respectively. $C_{I,d}$, C_M, C_{el} and C_{RM} represent the depreciated investment, maintenance, electricity and raw material's cost. For the separation system, the latter is calculated by the difference of the energy content between the crude and grid quality SNG. The energy in the depleted streams is therefore considered as lost.
In contrast to this approach targeting only the separation performance, a second set of indicators that represent the whole system performance is defined. The overall system efficiency (Eq. 5) and the production cost for grid-quality SNG (Eq. 6) are used:

$$\varepsilon = \frac{\Delta h^0_{SNG}\dot{m}^-_{SNG} + \dot{E}^-}{\Delta h^0_{wood}\dot{m}^+_{wood} + \dot{E}^+} \qquad (5); \qquad C_p = C_{I,d} + C_M + C_{OL} + C_{el} + C_{RM} \qquad (6)$$

With this formulation, the energy efficiency accounts for the possibility of co-producing electrical power $\dot{E}^-$, and the production cost includes the cost of operating labour C_{OL}. Furthermore, the raw material's cost C_{RM} considers only the expenses for wood, which implies that the depleted gas is treated on-site and might provide useful heat for the system. Since the evaluation of these overall properties in the separation design requires knowledge and data of the whole system, this approach represents a simultaneous design of the crude SNG production and its upgrading. The dotted and dashed parts on Figure 1 are thus considered simultaneously, and advantages due to the process integration with respect to the energy and material streams are accounted.

3.2.2. Optimisation results and discussion

In order to generate a complete set of optimal configurations, a multi-objective optimisation of the product recovery, power consumption and investment cost as objectives is performed, where the membrane system layout, the stage cuts and its operating pressures are used as decision variables. Figure 3 shows the obtained performance indicators for the best system layouts of the superstructure. From these plots, considerable qualitative and quantitative differences in the performance of the isolated membrane cascade and the overall system is observed. The isolated optimisation clearly suggests a 3 stage countercurrent cascade with about 94% of SNG recovery in the separation step. However, from an overall system point of view, it is sufficient to only recover about 84% of the crude SNG in a less costly system with only one recycle loop. For the overall system, a mediocre performance of the separation step is not penalising. With a non-negligible fraction of methane and hydrogen in the depleted stream (20%vol in total), its flame temperature remains sufficiently high to supply heat above the high temperature pinch, and less producer gas must be used for this purpose. Due to the advantageous process integration, considerably less electrical power is consumed in the separation (67 instead of 82 kW_{el}/MW_{SNG}) and a much smaller membrane system (230 instead of 363 $m^2/MW_{crudeSNG}$ in total) is required.

4. Conclusions

By implementing a thermo-economic model for multicomponent membrane gas separation in a process model for SNG production, the process integration of the gas separation with its production has been investigated. It has been shown that a holistic approach provides a qualitatively different and more advantageous process design than separately addressing the crude SNG production and its separation. Using the residual methane and hydrogen in the depleted stream as fuel for the process, it is possible to considerably reduce the size and cost of the gas upgrading system.

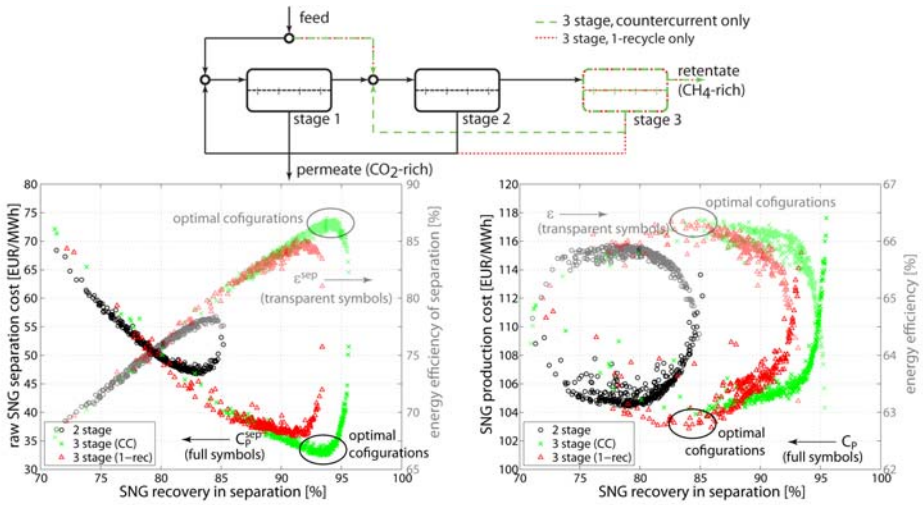


Figure 3: Performance of the separation system (left) and the overall process (right).

References

Belsim SA, 2007, Vali IV, www.belsim.com.

B. Bhide, S. Stern, 1993, Membrane processes for the removal of acid gases from natural gas. I. Process configurations and optimization of operating conditions, Journal of Membrane Science 81, 209-237.

A. Duret, C. Friedli, F. Maréchal, 2005, Process design of Synthetic Natural Gas (SNG) production using wood gasification, Journal of cleaner production 13, 1434-1446.

M. Gassner, F. Maréchal, 2005, Thermo-economic model of a process converting wood to methane, Submitted to Biomass and Bioenergy.

M. Gassner, F. Maréchal, 2008a, Methodology for the optimal thermo-economic, multi-objective design of thermochemical fuel production from biomass, Accepted for publication in Computers and Chemical Engineering.

M. Gassner, F. Maréchal, 2008b, Thermo-economic optimisation of the integration of electrolysis in synthetical natural gas production from wood, Energy 33, 189-198.

S. Heyne, H. Thunman, S. Harvey, 2008, Integration aspects for synthetic natural gas production from biomass based on a novel indirect gasification concept, In: PRES 2008, Prague.

H. Hofbauer et al., 2002, Six years experience with the FICFB-gasification process, In: Proceedings of the 12th European conference and technology exhibition on biomass for energy, industry and climate protection. Amsterdam, Netherlands.

S.-T. Hwang, K. Kammermeyer, 1975, Membranes in separation, Wiley, New York.

J. Luterbacher et al., 2008, Hydrothermal gasification of waste biomass. Sustainable process development using life cycle assessement, Accepted for publication in Environmental Science and Technology.

M. Mozaffarian, R. W. R. Zwart, 2003, Feasibility of biomass/waste-related SNG production technologies, ECN, Petten.

T. Pettersen, K. Lien, 1994, A new robust design model for gas separating membrane modules, based on analogy with counter-current heat exchangers, Computers and Chemical Engineering 18 (5), 427-439.

J. Phair, S. Badwal, 2006, Materials for separation membranes in hydrogen and oxygen production and future power generation, Science and technology of advanced materials 7, 792-805.

S. Stucki, 2005, Projet bois-methane. Rapport sur la clôture de la phase 1 du projet: Preuve de la faisabilité technique à l'échelle du laboratoire, PSI, Villigen, Switzerland.

SVGW, 2007, Richtlinien für die Einspeisung von Biogas ins Erdgasnetz, Entwurf V5, Zürich.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

The application of a task-based concept for design of innovative industrial crystallizers

Richard Lakerveld[a], Herman J.M. Kramer[a], Peter J. Jansens[a], Johan Grievink[b]

[a]*Process & Energy, Delft University of Technology, Leeghwaterstraat 44, 2628CA, Delft, The Netherlands*
[b]*DelftChemTech, Delft University of Technology, Julianalaan 136, 2628BL, Delft, The Netherlands*

Abstract

A new task based approach is applied to design a solution crystallization process unit. Task-based design involves the conceptual built-up of a process (unit) from functional building blocks called tasks, which represent fundamental physical events. The motivation is to get a better control over the physical events governing crystalline product quality. To deliver a proof of concept, two lines of research are followed. First of all, several small scale experiments are designed to demonstrate practical feasibility of the approach. The new lab-scale equipment allows for isolation and manipulation of individual crystallization tasks. Secondly, a model based on the experimentally tested tasks is developed for a crystallizer design and used for dynamic optimization in batch mode of product quality with minimum energy consumption. The results show that a task based crystallizer is capable of keeping the state variables very closely to ideal values, which is the direct result of the ability to control the rates at which individual crystallization tasks are executed as well as the material flows between those tasks.
Keywords: Crystallization, Process Synthesis, Task Based Design, Optimization.

1. Introduction

Crystallization is one of the oldest and economically most important separation and product formation technologies. The design of crystallization processes still poses many challenges despite its wide application. The selection of crystallization equipment is traditionally done from a limited number of state-of-art crystallizers followed by optimization of that particular type of equipment. Present industrial crystallizers harbor many physical phenomena such as primary nucleation, crystal growth, generation of supersaturation, and attrition. Optimization of each of these individual physical phenomena is not well possible because in present industrial crystallizers these phenomena are strongly entangled. This lack of control over physical phenomena causes crystallizers to exhibit complex process behavior and limits flexibility for both design and operation. The physical phenomena are however of key importance as in the end they determine the properties of the crystalline product and also the efficiency in terms of energy consumption. Therefore there is a need for a design approach which considers the important phenomena as starting point for design rather than the equipment itself. This contribution discusses the application of a task based design (TBD)(Menon, 2007).

In TBD an attempt is made to conceptually construct the crystallization process from fundamental building blocks called physical tasks (comparable to Kondili et al. 1993). In principle the TBD approach for solution crystallization can be embedded in existing generalized frameworks for the representation of superstructures and optimization models in process synthesis such as derived by Yeomans & Grossmann

(1999). For crystallization processes it is however difficult to specify physical devices that will execute a single crystallization task as current crystallizers facilitate many of those tasks simultaneously and the control of individual tasks is not well possible. Therefore, the first part of this paper summarizes the key results of small scale experiments with newly built dedicated equipment showing that it is possible to isolate and optimize single crystallization tasks. The results, of which detailed descriptions are described elsewhere (Lakerveld 2008[a,b,c]), demonstrate the practical feasibility of the approach showing the isolation and optimization of the crystallization tasks supersaturation generation, nucleation and growth. In the second part of this paper a model based on the experimentally tested tasks is developed for a task based crystallizer design and used in dynamic optimization of a batch case study tailored towards optimizing product quality. The results show that simply by changing the operational policy, a tight control over product quality is possible.

2. Experimental isolation of single tasks

The control over a single crystallization task means in practice that the operating conditions in a physical device must be tailored such that the resulting driving forces enable a specific task to be dominant. Figure 1 shows for a number of fundamental crystallization phenomena qualitatively at which driving force the kinetic rate is being maximized (task-impact diagram). It reveals partly why it is inherently difficult to optimize the driving force for each of the tasks in current well mixed crystallizers. To tightly control and possibly enlarge the window of operation for a selected key crystallization task, new equipment and methods are needed.

1. For tight control over the generation of supersaturation, a membrane based crystallization process. It gives more flexibility and better control over solvent removal to optimize the driving force of a selected task, most importantly crystal growth. Membranes can be placed at any location in the design of future crystallizers and can be applied outside boiling conditions.
2. To generate new nuclei, ultrasound as a new method to induce nucleation. Ultrasound uses pressure as driving force and can therefore be applied at global low supersaturation. In this way, it helps to disentangle the task-impact diagram (Figure 1) by adding a new driving force, to the diagram.
3. For separation of the task growth and attrition, a bubble column and airlift crystallizer as new equipment for avoiding regions exhibiting large shear forces. It prevents unwanted operation in the upper part of the task-impact diagram where attrition and growth, two counter effective tasks, interfere.

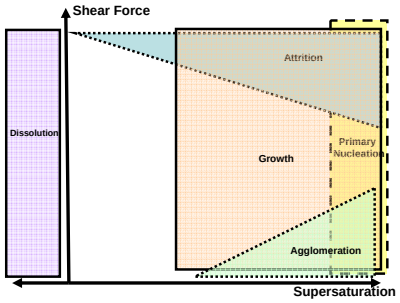


Figure 1: Task- impact diagram qualitatively illustrating for two main driving forces the rate of execution of selected tasks

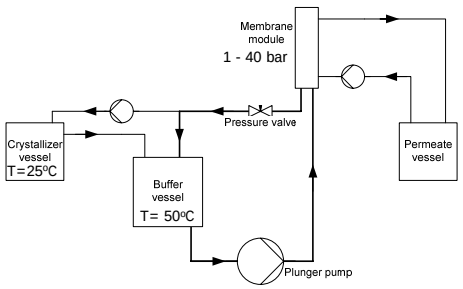


Figure 2: Process flow diagram experimental setup combining crystallization and reverse osmosis

2.1. Tight control over the generation of supersaturation

The first objective of the experimental work is related to tight control of supersaturation at any location in a future crystallizer to prevent spontaneous nucleation bursts and to maximize crystal growth. Membranes can selectively remove the solvent and offer an interesting opportunity to control supersaturation gradients in new crystallizer designs (see for example Tun et al. 2005). Potentially they can also contribute to reduced energy consumption as an expensive vapor liquid equilibrium can be avoided. An experimental setup (Figure 2) was constructed and tested for both an NH_4SO_4 water system and an adipic acid water system to assess the potential application of reverse osmosis (RO) membranes for crystallization processes (Lakerveld et al. 2008[a]). It showed practical feasibility of the concept and typical values for mass flux over the membrane are used in the next section for design.

2.2. Experimental isolation of task Nucleation

Ultrasound is an interesting tool to induce nucleation at low supersaturation in a controlled way (e.g. Virone et al., 2006). The applicability of ultrasound was illustrated in experiments in which an ultrasound generator was placed inside a supersaturated solution of an NH_4SO_4 water system. The supersaturation was kept low to avoid spontaneous nucleation. The objective of the experiments was to relate the power input and insonation time to the number of nuclei produced. It was concluded that nuclei could be generated at low supersaturation. The nucleation rate could be optimized, by changing the power input and insonation time (Lakerveld et al. 2008[b]).

2.3. Experimental isolation of task Crystal Growth in air mixed devices

To isolate the task Crystal Growth a physical environment is needed which can minimize attrition as attrition increases crystal surface area, which competes for growth with the existing crystals. The environment that was selected consisted of a bubble column in which supersaturation was created by simultaneous cooling and evaporation of the solvent by sparging air (Lakerveld et al., 2008[c]). The crystals were kept in suspension by the upward velocity of the bubbles, eliminating the need for a stirrer or a circulation pump. In this way, attrition caused by collisions between crystals and moving mechanical parts was absent. Seeded batch experiments in a 3-l bubble column showed that an initial seed population could grow without an increase in number of crystals. Similar experiments in agitated crystallizers (Hojjati et al. 2005, Lakerveld et al. 2007) showed a clear increase in number of crystals due to attrition. It demonstrated that the concept is very promising to isolate the task nucleation. Further improvements are expected by extending the bubble column to an airlift crystallizer. A dedicated 10-L airlift crystallizer has been designed, constructed and is currently being tested.

3. Dynamic optimization of a task based crystallizer

The starting point of the development of a TBD for solution crystallization, was the development of new process units that could isolate single tasks as described in the previous section. Those new unit operations provide possible building blocks for a task superstructure, which can be optimized resulting in an optimal task structure. In this section, the tasks tested in the experiments, as summarized in the previous section, are combined in a model, which represents a possible outcome of such a task superstructure optimization, and optimized. The aim is to illustrate flexibility for process design resulting in improved product quality with minimum energy demand.

A drawing of the modeled system is given in Figure 3. It contains one compartment with a gas, a liquid and a solid phase (labeled GLS), which behaves as the bubble column. This compartment is used to grow crystals with negligible attrition.

Ultrasound can be applied within the compartment to produce nuclei at low supersaturation with the experimentally determined rate. From this compartment a clear solution is taken to a second compartment (labeled ML) operating at higher temperature, in which the task supersaturation generation is realized. A RO membrane is used here to selectively remove solvent. The system is operated in semi-batch mode.

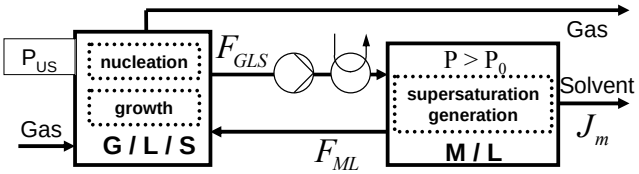


Figure 3: structure of task based crystallizer with 3 tasks

For the development of the model special attention has been paid to proper scaling of the state variables. Furthermore, instead of solving the full population balance describing the dynamic development of the crystal size distribution, which results in a system of PDAE equations, the moment transformation has been used to reduce the system to a set of DAE equations. The growth rate depends linearly on supersaturation

$$G = k_G \cdot \left(w_{GLS} - w_{GLS}^{*} \right) \tag{1}$$

The moment equations are given as follows:

$$\frac{dx_0}{dt} = \eta_U \alpha \tag{2}$$

$$\frac{dx_i}{dt} = i \cdot \left(w_{GLS} - w_{GLS}^{*} \right) \cdot x_{i-1} \quad \text{for i = 1,2,3,4} \tag{3}$$

Where the variable η_U represents the utilization of ultrasound. It can be imagined as the fraction of the total volume that is insonated. Furthermore, $x_0,...,x_4$ represent the scaled moments ($m_0,...,m_4$) and α represents the scaled maximum nucleation rate defined as follows:

$$x_0 = m_0 k_G^3 \;, x_1 = m_1 k_G^2 \;, x_2 = m_2 k_G \;, x_3 = m_3 \;, x_4 = m_4 k_G^{-1} \;, \alpha = B_{max} \cdot k_G^3$$

The component and total mass balance for the GLS compartment are

$$\frac{d\left[V_{GLS} \rho_L w_{GLS} \varepsilon + (1-\varepsilon) V_{GLS} \rho_S \right]}{dt} = F_{ML} w_{ML} \rho_L - F_{GLS} w_{GLS} \rho_L \tag{4}$$

$$-k_V \frac{dx_3}{dt} V_{GLS} \left(\rho_L - \rho_S \right) = \rho_L \left(F_{ML} - F_{GLS} \right) \tag{5}$$

The liquid volumetric flow F_{ML} is flow controlled, the liquid fraction ε is given by

$$\varepsilon = 1 - k_V x_3 \tag{6}$$

The total mass and component balance for the ML compartment complete the model:

$$\frac{dw_{ML} V_{ML}}{dt} = F_{GLS} w_{GLS} - w_{ML} F_{ML} \tag{7}$$

$$\rho_L \frac{dV_{ML}}{dt} = \rho_L F_{GLS} - \rho_L F_{ML} - J_m A_{mem} \tag{8}$$

The energy consumption of the system is defined as

$$E_{tot} = \int_{t=0}^{t=t_f} \left(P_{US} + F_{GLS}(t) \left[\frac{\Delta P(t)}{\eta} + C p_L \rho_L \left(T_{ML} - T_{GLS} \right) \right] \right) \cdot dt \tag{9}$$

The mass flux over the membrane J_m can be manipulated by changing the pressure in the membrane module, which depends linearly on the mass flux. The model was implemented in gPROMS Modelbuilder 3.1.4, (PSE Ltd). The parameter settings corresponded to an NH_4SO_4 water system with initial values in both compartments according to a clear liquid saturated at 25°C. The manipulated variables were the mass flux over the membrane (J_m) and the ultrasound utilization (η_U), which both change the rate at which a specific task is executed. Furthermore, the flow rate F_{ML} between the two

compartments was varied, which manipulated the connection between tasks. The fixed time space was discretized into 100 intervals. The operational policy could be optimized to maximize mean size with minimum energy consumption by introducing two sequential objective functions with (10) having highest priority and constraints

$$\max_{J_M(t),\eta_U,F_{ML}(t),T_{ML}} \frac{m_4\left(t=t_f\right)}{m_3\left(t=t_f\right)} \tag{10}$$

Subject to:

$$\min_{J_M(t),\eta_U,F_{ML}(t),T_{ML}} E_{tot} \tag{11}$$

$$0 < J_M\left(t\right) < 2.5 \tag{12}$$

$$1\cdot 10^{-4} < F_{ML}\left(t\right) < 0.1 \tag{13}$$

$$T_{GLS} < T_{ML} < 373 \tag{14}$$

$$w_{GLS}^{SAT}\left(T_{GLS}\right) < w_{ML}\left(t\right) < w_{ML}^{SAT}\left(T_{ML}\right) \tag{15}$$

$$0.05\% \le \frac{w_{GLS}\left(t\right) - w_{GLS}^{SAT}}{w_{GLS}^{SAT}} \le 1.0\% \tag{16}$$

$$\varepsilon\left(t=t_f\right) < 0.85 \tag{17}$$

for $t \in \left[0, t_f\right]$

The optimal profile for the mass flux over the membrane, which is determined by the pressure difference over the membrane, is depicted in Figure 4 together with the optimal trajectory of the flow rate between both compartments F_{ML}. The trajectory of supersaturation and concentration in ML compartment, which both represent constrained state variables, is shown in Figure 5.

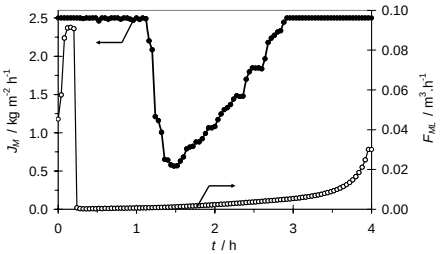


Figure 4: Optimized profile J_M (●) and F_{ML} (○)

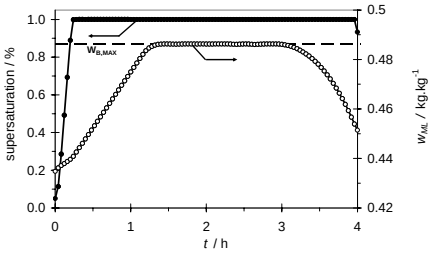


Figure 5: relative supersaturation (●) and solute fraction ML compartment (○)

From Figures 4-5 it can be seen how both the solvent mass flux over the membrane, flow rate between both compartments and utilization of ultrasound work together to maximize the final volume based mean size. Ultrasound is active only for 7% during the first 2 minutes of the batch (not shown) to generate an optimal amount of initial nuclei. The mass flux and flow rate between compartments are high in the beginning (Figure 4) to quickly raise the supersaturation to the constrained value (Figure 5) representing maximum growth (Eq. (1)). The flow rate is reduced as soon as the supersaturation hits the constraint. From this point, the concentration in the ML compartment increases more rapidly (Figure 5) until this concentration becomes saturated (T_{ML} = 81.7 °C). To prevent supersaturation, and therefore scaling on the membrane, the mass flux is reduced quickly at this point to reduce the increase in concentration. As the crystal surface area increases over time more supersaturation can be delivered to the GLS compartment and the mass flux can gradually increase. Note that for minimization of energy the optimizer seeks to minimize the product of T_{ML} and F_{ML} as heating is the most significant contribution to the energy consumption. This is reflected by the profile of W_{ML} (Figure 5), which is at the maximum value, determined by T_{ML}, for a large part of the batch.

It is important to notice that supersaturation is at the maximum value almost throughout the complete batch, showing that the task based crystallizer is very well capable of manipulating supersaturation and therefore product quality (size distribution). Using the same crystallizer design, also a small product with narrow size distribution can be produced (Lakerveld et al., 2008[d]).

4. Conclusions & Outlook

A TBD approach is applied for the design of a crystallization process unit. A two track approach has been followed. First of all, small scale experiments are designed and tested to evaluate technology that is capable of isolating single crystallization tasks. It delivers a proof of principle that dedicated equipment is capable to execute the fundamental crystallization tasks Growth, Nucleation and Supersaturation Generation. Modeling delivers the building blocks for design, which are used in the second part of this paper dealing with dynamic optimization of a task based batch crystallizer for optimization of the final mean size with minimum energy consumption. The results show that a task based crystallizer is flexible and therefore capable of maintaining supersaturation very closely to an optimum value as determined by the designer.

Although the present work is an important step forward in proving the technological viability of TBD for solution crystallization processes, many follow-up questions must be answered. Out of the general engineering abilities reliability and availability have not been discussed. This will be a challenge to the approach since unconventional technology is used. The modeling of individual tasks will also contribute to an improved reliability of the approach and better understanding how to isolate tasks for various systems. Furthermore, the task structure was fixed for the optimization studies performed in this work as it represented the combination of tested experimental setups. However, optimization of the task structure itself could further improve the design of the crystallization process unit. It involves postulation of a task superstructure that is rich enough to incorporate interesting process alternatives on the one hand and removing abundant alternatives on the other hand. Furthermore, optimization of the task superstructure would transform the NLP optimization problem of the current work into an MINLP optimization problem, which is computationally more demanding. Those are challenging questions for future research.

References

[1] H.Hojjati, S.Rohani, 2005, Chem. Eng. Process., 44, 949
[2] A.R.Menon, A.A. Pande, H.J.M. Kramer, J. Grievink, P.J. Jansens, 2007, Ind.Eng.Chem.Res., 46 ,3979
[3] E. Kondili, C.C. Pantelides, R.W.H. Sargent, 1993, Comput.chem. Eng., 17, 211
[4] H. Yeomans, I.E. Grossmann, 1999, Comput.chem. Eng., 23, 709
[5] R. Lakerveld, J. Kuhn, M.A. Bosch, H.J.M. Kramer, P.J. Jansens, J. Grievink, 2008a in ISIC17 Conference Proceedings, Maastricht, The Netherlands, 827
[6] R. Lakerveld, P.G. Verzijden, H.J.M. Kramer, P.J. Jansens, J. Grievink, 2008b in ISIC17 Conference Proceedings, Maastricht, The Netherlands, 835
[7] R. Lakerveld, H.J.M. Kramer, P.J. Jansens, J. Grievink, 2008c in ISIC17 Conference Proceedings, Maastricht, The Netherlands, 819
[8] R. Lakerveld, A.N. Kalbasenka, H.J.M. Kramer, P.J. Jansens, J. Grievink, 2007, in Proceedings of 14th Internationcal Workshop on Industrial Crystallization, 221
[9] C.M. Tun, A.G. Fane, J.T. Matheickal, R. Sheikholeslami, 2005, J. Membr. Sci., 257, 144
[10] C.Virone, H.J.M. Kramer, G.M. van Rosmalen et al.,2006, Journal of Crystal Growth 294, 9
[11] R. Lakerveld, H.J.M. Kramer, P.J. Jansens, J. Grievink, 2008d in ISIC17 Conference Proceedings, Maastricht, The Netherlands, 27

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Viscoelastic flow simulation: development of a methodology of analysis using the software OpenFOAM and differential constitutive equations

Jovani L. Favero,[a] Argimiro R. Secchi,[b] Nilo Sérgio M. Cardozo,[a] Hrvoje Jasak[c]

[a]*DEQUI/UFRGS - Department of Chemical Engineering - Universidade Federal do Rio Grande do Sul, Rua Luis Englert, s/n., Central campus, CEP 90040-040, e-mail:{favero, nilo}@enq.ufrgs.br, Porto Alegre - RS, Brazil*
[b]*PEQ/COPPE-Universidade Federal do Rio de Janeiro, CT, Av. Horácio Macedo, 2030 G116, e-mail: arge@peq.coppe.ufrj.br. CEP 21941-914, Rio de Janeiro - RJ, Brazil*
[c]*FSB/Wikki – Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Croatia and Wikki Ltd, e-mail: h.jasak@wikki.co.uk. United Kingdom. London W8 7PU, United Kingdom.*

Abstract

Viscoelastic fluids are of great importance in many industrial sectors, such as in food and synthetic polymers industries. The rheological response of viscoelastic fluids is quite complex, including combination of viscous and elastic effects and highly non-linear viscous and elastic phenomena. This work presents a new Computational Fluid Dynamics (CFD) tool for the simulation of viscoelastic fluid flows, which consists of a viscoelastic fluid module to be used with OpenFOAM CFD package. The main reasons for using OpenFOAM in the development of this tool are its characteristics with relation to flexibility to deal with complex geometries, unstructured and non orthogonal meshes, moving meshes, large variety of interpolation schemes and solvers for the linear discretized system, and the possibility of data processing parallelization. Several constitutive equations, such as Maxwell, UCM, Oldroyd-B, Giesekus, Phan-Thien-Tanner (PTT), Finitely Extensible Nonlinear Elastic (FENE-P and FENE-CR) and some derivations of Pom-Pom were implemented, in single and multimode form, and in this work the results are presented with the Giesekus model. The proposed methodology was validated by comparing its predictions with experimental and numerical data from the literature for the analysis of a planar 4:1 contraction flow.

Keywords: Viscoelastic flows, planar contraction, numerical simulation, OpenFOAM.

1. Introduction

Understanding and modeling viscoelastic flows are usually the key step in the definition of the final characteristics and quality of finished products in many industrial sectors, such as in food and synthetic polymers industries.

The rheological response of viscoelastic fluids is quite complex, including combination of viscous and elastic effects and highly non-linear viscous and elastic phenomena, such as strain rate dependent viscosity, presence of normal stress differences in shear flows, relaxation phenomena, and memory effects, including die swell (Bird et al., 1987, Macosko, 1994).

In addition to the basic mathematical and computational difficulties associated with the solution of the system of equations resulting from mass, momentum and energy conservation laws, the analysis of viscoelastic flows requires dealing with an additional set of variables and equations related to the constitutive equation used to describe the rheological behaviour of the viscoelastic fluid.
The first constitutive equation developed to describe viscoelasticity was the Maxwell linear model. To get a better representation of real viscoelastic fluids at higher deformation rates, more sophisticated and complex non-linear viscoelastic models should be used. In this class there are models derived from the kinetic theory (such as Oldroyd-B, Giesekus and FENE-P), from network theory of concentrated solutions and melts (Phan-Thien-Tanner - PTT), and from the reptation theory (Pom-Pom and XPP models). A more complete list of models for viscoelastic fluid and some detailed comparisons among them can be found in Bird et al. (1987), Macosko (1994), and Larson (1988).
In the attempt of solving the 'High Weissenberg Number Problem' (HWNP) or minimizing its effects, many strategies as special interpolations schemes (Alves et al., 2003, Muniz et al., 2005) and specific numerical methodologies for the treatment of nonlinear viscoelastic models have been proposed. As well as the 'viscous formulation', the EVSS (Elastic viscous split-stress), the DEVSS (Discrete Elastic viscous split-stress) and the AVSS (Adaptive Viscosity Stress Splitting Scheme) are among the most known of these numerical methodologies.
Another relevant aspect regarding the analysis of viscoelastic flows is the development of software specifically designed for this purpose. Despite the extensive literature on viscoelastic behaviour modelling and viscoelastic flow simulation, the commercial packages meant to be applied for analysis of polymeric flows and polymer processing operations, such as extrusion and injection molding, are actually limited to the description of purely viscous non-Newtonian phenomena. Software with viscoelastic flow analysis capability are still developed and used only in academic environment, for specific applications.
In this sense the insertion of a viscoelastic fluid flow solver in a widely recognized CFD (Computational Fluid Dynamics) package is a relevant matter, since it would bring to viscoelastic fluid flow analysis field the main features of a good CFD package, which include the ability of dealing with complex geometries, unstructured and non orthogonal meshes, moving meshes, large variety of interpolation schemes and solvers for the linear discretized system, and data processing parallelization among others benefits.
Based on these ideas, the goal of this work is the implementation of constitutive models for viscoelastic fluids in OpenFOAM (Open Source Field Operation and Manipulation). OpenFOAM is a free CFD finite volume method (FVM) package with flexible code that, in addition to the many standards solvers already implemented for more general Newtonian fluids problems, also allows users to build their own codes for specific issues. The whole CFD methodology in OpenFOAM can be found in the work of Jasak (1996).

2. Methodology

In this section, the mathematical formulation and the strategy used to the solution of viscoelastic fluid flow problem is presented.

2.1. Mathematical model

The governing equations of laminar, incompressible, and isothermal flow of viscoelastic fluids are the conservation of mass (continuity):

$$\nabla \cdot (U) = 0 \tag{1}$$

and momentum:

$$\frac{\partial (\rho U)}{\partial t} + \nabla \cdot (\rho U U) = -\nabla p + \nabla \cdot \tau_S + \nabla \cdot \tau_P \tag{2}$$

together with a mechanical constitutive equation that describes the relation between the stress and the deformation rate for a given fluid. In the above equations ρ is the density, U the velocity vector, p the pressure, τ_S the solvent Newtonian contribution, and τ_P the polymeric contribution.

The solvent contribution is given by:

$$\tau_S = 2\eta_S D \tag{4}$$

where η_S is the solvent viscosity and D is the deformation rate tensor given by:

$$D = \frac{1}{2}(\nabla U + [\nabla U]^T) \tag{5}$$

The extra elastic contribution, corresponding to the polymeric part, is obtained from the solution of an appropriate constitutive differential equation. A well known and highly used is the Giesekus constitutive model described as (Giesekus, 1982):

$$\tau_P + \lambda \overset{\nabla}{\tau}_P + \alpha \frac{\lambda}{\eta_p}(\tau_P . \tau_P) = 2\eta_p D \tag{6}$$

where λ and η_P are the relaxation time and polymer viscosity at zero shear rate, respectively, and α is the dimensionless 'mobility factor'.

The upper convected derivative of the elastic stress tensor is (Larson, 1988):

$$\overset{\nabla}{\tau}_P = \frac{D}{Dt}\tau_P - \left[(\nabla U)^T \cdot \tau_P\right] - \left[\tau_P \cdot \nabla U\right] \tag{7}$$

with the material derivative of the extra stress tensor defined as:

$$\frac{D}{Dt}\tau_P = \frac{\partial}{\partial t}\tau_P + U . \nabla \tau_P \tag{8}$$

2.2. DEVSS methodology

DEVSS (Guénette and Fortin, 1995) was the methodology chosen in this work to deal with the HWNP, mainly because of the easiness of its extension to different complex constitutive models. This method is based on the introduction of an additional diffusive term on each side of the momentum equation, which is then rewritten as:

$$\frac{\partial(\rho U)}{\partial t} + \nabla \cdot (\rho U U) - (\eta_S + \phi)\nabla \cdot (\nabla U) = -\nabla p + \nabla \cdot \tau_P - \phi \nabla \cdot (\nabla U) \tag{9}$$

where ϕ is a positive number.

2.3. Solver implementation and solution procedure

The procedure implemented to solve the problem of viscoelastic fluid flow can be summarized in the following four steps:

1. Considering an initial known velocity field U^*, given pressure p^* and stress τ^*, the momentum equation is implicitly solved for each component of the velocity vector resulting in U^{**}. The pressure gradient and the stress divergent are calculated explicitly with values of the previous step.
2. With the new velocity values U^{**}, the new pressure field p^{**} is estimated using an equation for the pressure (obtained from the application of divergent operator on momentum equation) and then is corrected with the velocity to satisfy the continuity equation. Here is chosen the PISO algorithm for transient flows.
3. With the corrected velocity field, the stress tensor field τ^{**} is evaluated using a chosen constitutive equation.

4. For transient flows, steps 1, 2 and 3 can be iterated at each time step to a desired accuracy.

The implementation of this algorithm in OpenFOAM was called viscoelasticFluidFoam solver.

2.4. Test geometry

A planar abrupt contraction with contraction ratio *H/h* of 3.97:1 (upstream thickness of *2H = 0.0254 m* and downstream thickness of *2h = 0.0064 m*) was chosen as test geometry, sketched in Figure 1. Quinzani et al. (1994) presented an extensive experimental study on this geometry using LDV (*Laser-Doppler Velocimetry*) and FIB (*Flow Induced Birefringence*) for stress measurements on a polymer solution of 5% wt. of poly-isobutylene (PIB) in tetradecane (C_{14}). Azaiez et al. (1996) compared these experimental results with the predictions of numerical simulation using the finite element method (FEM) and the Giesekus model.

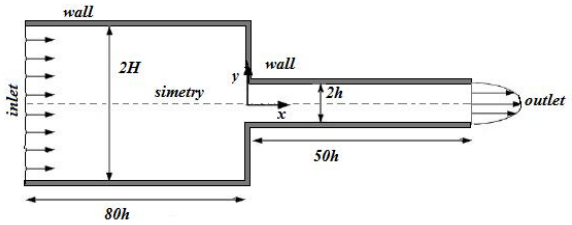


Figure 1: Sketch of geometry and the boundary conditions.

Table 1: Flow conditions.

Q [cm^{3}s^{-1}]	$\langle U \rangle$ [cms^{-1}]	Re_o	De_o
252	15.5	0.56	1.45

The flow conditions are presented in Table 1, where $<U>$ is the average velocity in the downstream section, Q the volumetric flow rate, Re_o and De_o the Reynolds and Deborah number for this case, respectively. The length of the channel before the contraction is *80h* and *50h* for the downstream section.

From the experimental data of Quinzani et al. (1994), the parameters in Table 2 for the Giesekus model were obtained by Azaiez et al. (1996).

Table 2: Parameters of Giesekus model (Azaiez et al., 1996).

Model	α [--]	λ [s]	η_P [Pa.s]	η_S [Pa.s]	ρ [kg m^{-3}]
Giesekus	0.15	0.03	1.422	0.002	*803.87*

The *upwind* and *central differences* approximations were used for the convective terms and the others terms, respectively. *Conjugate gradients* with *AMG* preconditioning were used in the linear system solver for the pressure field and *BiCGstab* with a *Cholesky* preconditioning was used for velocity and stress. The absolute tolerance for pressure was 10^{-7} and the velocity and stress 10^{-6}.

3. Results and Discussion

The obtained results were compared with the simulated data by Azaiez et al. (1996) and experimental data obtained by Quinzani et al. (1994).

Figure 2 illustrates the stress field components τ_{xx}, τ_{xy}, τ_{yy} and the stream lines, showing the effects near the re-entrant corner, where the control volumes are very small. This

region is the most critical region with largest gradients and highest stress values. The stream lines illustrate a vortex appearance in the corner of the geometry.

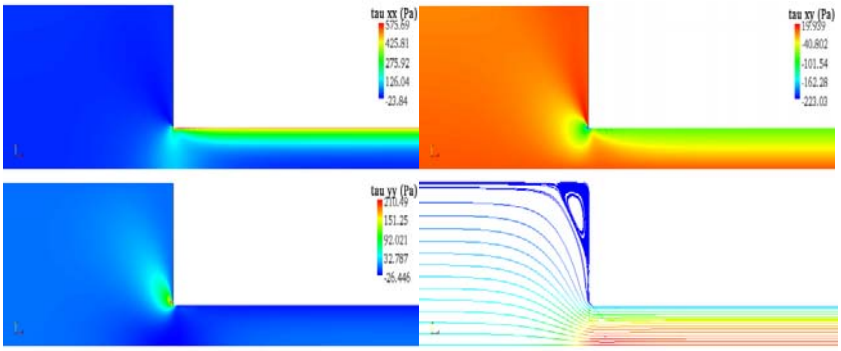


Figure 2: Stress field and stream lines.

In the Figures 3, 4 and 5, a comparison for the velocity profile U, τ_{xy} and the first normal stress difference ($N_1 = \tau_{xx} - \tau_{yy}$), respectively, is made in a lateral and axial scan at the upstream section.

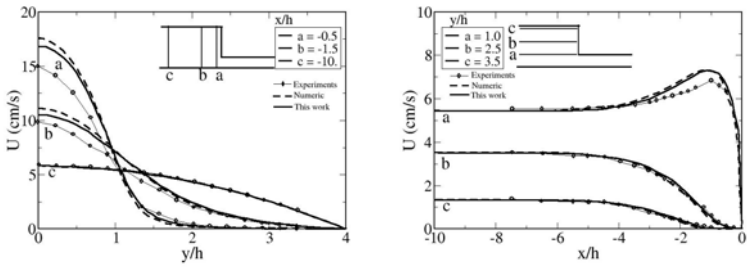


Figure 3: Velocity profile U in the upstream section.

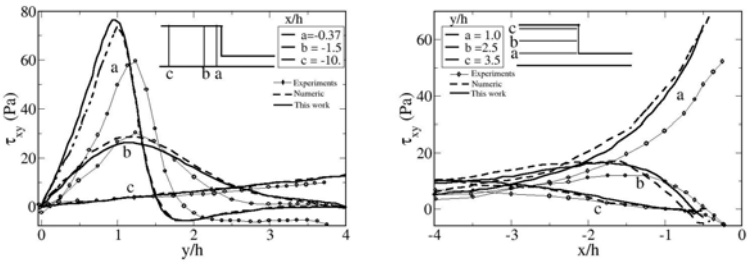


Figure 4: Stress τ_{xy} profile in the upstream section.

It can be observed that the results for the velocity profile are almost identical to those obtained by Azaiez et al. (1996) and both have good agreement with experimental data, mainly in the developed flow region. Higher values of stress were obtained when comparing with the experimental results, but they are in good agreement with the numeric results of the literature. The little difference can be attributed to the different interpolation schemes and meshes used. The difference between the numeric and

experimental data may be attributed to the only one relaxation mode used, and thus, not good enough to completely characterize the stress, mainly in regions where there is a high shear rate and occurs the higher stress values.

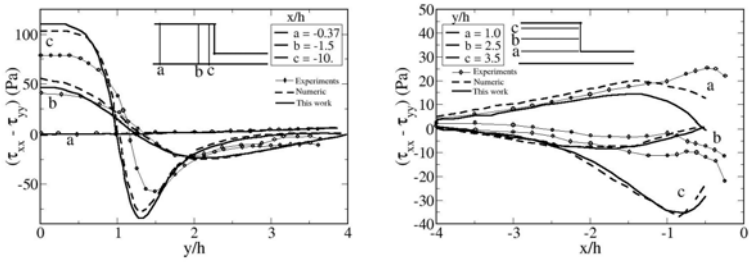


Figure 5: First normal stress difference profile N_1 in upstream section.

The first stress difference (N_1) also show good agreement between the numeric results from literature and the obtained in this work. For the lateral scan, differences between the numeric and experimental data are observed manly at high shear rates. In the developed flow region the numeric and experimental results show good agreement. Analyzing the axial scan, there is a good concordance between the numeric results, but a poor agreement with the experimental data mainly near to the contraction. However, the experimental data may also contain significant errors in that region. Improvements in N_1 prediction may be achieved with the use of multimode models.

4. Conclusion

In this work, a new tool for CFD simulation of viscoelastic fluids was developed. The results showed that the solver leads to consistent results with those found in literature and are in good agreement with experimental data. In order to better represent the values of stress obtained experimentally it would require the use of multimode models. The OpenFOAM is a suitable environment for the development of new solvers for viscoelastic fluid flows, having a great potential for academic and industry researches.

References

R.G. Larson, 1988, Constitutive equations for polymer melts and solutions, Butterworths, Boston.

R.B. Bird, R.C. Armstrong, O. Hassager, 1987, Dynamics of Polymeric Liquids Vol. 1, Fluid Mechanics, John Wiley, New York.

R. Guénette, M. Fortin, 1995, A new mixed finite element method for computing viscoelastic flows, J. Non-Newtonian Fluid Mech., 60, 27-52.

M.A. Alves, P.J. Oliveira, F.T. Pinho, 2003, Int. J. Numer. Meth. Fluids, 41, 47–75.

L.M. Quinzani, R.C. Armstrong and R.A. Brown, 1994, J. Non-Newtonian Fluid Mech., 52, 1-36.

J. Azaiez, R. Guénette, A. Ait-Kadi, 1996, Numerical simulation of viscoelastic flows through a planar contraction, J. Non-Newtonian Fluid Mech., 62, 253-277.

H. Jasak, 1996, Error Analysis and Estimation for the Finite Volume Method with Applications to Fluid Flows. PhD thesis, Imperial College of Science, Technology and Medicine, University of London.

H. Giesekus, 1982, A simple constitutive equation for polymer fluids based on the concept of deformation-dependent tensorial mobility, J. Non-Newtonian Fluid Mech., 11, 69- 109.

C. Macosko, Rheology: Principles, 1994, Measurements and Applications, VCH Publishers.

A.R. Muniz, A.R. Secchi, N.S.M. Cardozo, 2005, Ciência e Tecnologia, 15, 53-58.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling of Multi-Stream Heat Exchangers with Phase Changes for Cryogenic Applications

Ravindra S. Kamath,[a] Ignacio E. Grossmann,[a] Lorenz T. Biegler[a]

[a]*Department of Chemical Engineering,Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh PA 15213, USA*

Abstract

A new equation-oriented process model for multi-stream heat exchangers (MHEX) is presented with a special emphasis on phase changes. The model internally uses the pinch concept to ensure the minimum driving force criteria. Streams capable of phase change are split into sub-streams corresponding to each of the phases. A novel disjunctive representation is proposed which identifies the phases traversed by a stream during heat exchange and assigns appropriate heat loads and temperatures for heat integration. Model is suitable for optimization studies particularly when the phases of the streams at the entry and exit of the MHEX are not known *a priori*. The capability of the model is illustrated using a case study on liquefaction of natural gas.

Keywords: multi-stream heat exchangers, heat integration, complementarity

1. Introduction

A multi-stream heat exchanger (MHEX) is a single process unit in which multiple hot and cold streams exchange heat simultaneously. MHEXs are very common in cryogenic applications where heat transfer equipment need to be kept compact and well-insulated while recovering heat from streams at very small temperature driving forces. Use of a MHEX to perform such heat transfer tasks often leads to substantial savings in both energy and capital cost.

Unlike conventional two stream heat exchangers, MHEXs involve many hot and cold streams; it is not clear how to enforce the minimum driving force constraint because the matches between the streams are not known *a priori*. Even performing an energy balance can be non-trivial when some streams involved in the MHEX change phase during heat transfer. This is because the correlations used to calculate enthalpy depend on the phase and since the outlet temperature is not known *a priori*, it implies that the phase is not known *a priori*. This is even more challenging in the case of process optimization because pressure and composition of some or all the streams in the flowsheet are treated as variables, which cause the phase boundaries to move during the optimization. Thus, it is not surprising that there are hardly any process models for MHEX in the open literature. As a result, design and operating conditions of many cryogenic processes like natural gas liquefaction and air separation are often based on rule of thumb or heuristics. Therefore, there is scope to optimize the operating conditions and even the state of the streams at the entry and exit points of the MHEX in order to further improve the process efficiency.

In this work, we propose a general nonlinear equation-oriented model for MHEX that addresses all the issues mentioned above. The model is based on the Duran and Grossmann (1986) model, which relies on pinch analysis for heat integration. The issue

of phase change is dealt with by splitting the physical stream into three sub-streams corresponding to superheated, two-phase and sub-cooled regions. This splitting is performed using a disjunctive representation involving Boolean variables. The disjunctions are reformulated in order to avoid binary variables by solving an inner minimization problem with complementarity constraints. The capability of the model is demonstrated using a case study from the cryogenic industry.

2. Model for MHEX in the absence of phase changes

The model for MHEX is equivalent to the following problem statement:

Given an MHEX that does not consume any heating and cooling utilities, determine feasible temperatures and heat capacity flowrates for the involved streams

Based on the above problem statement, we can modify the Duran and Grossmann (1986) model for simultaneous optimization and heat integration and apply it for MHEXs. The Duran and Grossmann (1986) model uses the heat integration constraints to calculate the utility targets, which are embedded with appropriate cost coefficients in the objective function of the overall nonlinear optimization problem for the flowsheet. Our proposed modification involves setting the hot and cold utility loads in their heat integration constraints to a constant value of zero. This forces the heat integration constraints to treat MHEX as an adiabatic device, i.e. net heat lost by all the hot streams will be matched to the net heat gained by all cold streams. Since the MHEX does not require hot and cold utilities, it does not contribute any utility cost to the objective function of the parent flowsheet. Thus, the final model for MHEX is the modified set of heat integration constraints that only need to be embedded as additional constraints in the nonlinear programming model of the overall flowsheet. This model has the following form:

$$
\begin{aligned}
&\min\ \phi(w,x)\\
&s.t.\ \ h(w,x)=0\\
&g(w,x)\leqslant 0\\
&\Omega(x)=\sum_{i\in H}F_i\left(T_i^{in}-T_i^{out}\right)-\sum_{j\in C}f_j\left(t_j^{out}-t_j^{in}\right)=0\\
&AP_C^p(x)-AP_H^p(x)\leqslant 0 \quad p\in P\\
&AP_H^p(x)=\sum_{i\in H}F_i\left[\max\left\{0,T_i^{in}-T^p\right\}-\max\left\{0,T_i^{out}-T^p\right\}\right] \quad p\in P\\
&AP_C^p(x)=\sum_{j\in C}f_j\left[\max\left\{0,t_j^{out}-(T^p-\Delta T_{min})\right\}-\max\left\{0,T_i^{out}-(T^p-\Delta T_{min})\right\}\right]
\end{aligned}
\tag{1}
$$

where H and C are index sets for the hot and cold streams which are involved with the MHEX and vector x given by $x=\left(F_i,T_i^{in},T_i^{out}:\text{all } i\in H;\ f_j,t_j^{in},t_j^{out}:\text{all } j\in C\right)$ represents the corresponding temperature and flowrate of these streams. $P=H\cup C$ is the index set of all the pinch point candidates i.e. the inlet of the involved streams. The vector w represents all variables in the process flowsheet not associated with heat integration while $\phi(w,x)$, $h(w,x)$ and $g(w,x)$ represent the objective function, mass and energy balances, design equations and other specifications of the process.

3. Dealing with phase changes in the MHEX

Although the concept of pinch and heat integration assumes that the heat capacity-flowrate of the streams is constant during heat transfer, this seems unreasonable when the stream changes phase or traverses through the 2-phase region. Use of a constant heat-capacity flowrate is likely to cut off candidate pinch locations near the dew and bubble points leading to a violation of the minimum driving force criteria.

In this work, we propose a new strategy for handling streams undergoing phase changes where it is not known *a priori* whether a stream changes phase or not. We propose to classify *a priori* the streams involved in heat integration into two mutually exclusive sets of streams which are capable of changing phase and those that do not change phase. Streams belonging to the former set, which are denoted as parent streams, are sub-divided into sub-streams corresponding to superheated (SUP), 2-phase (2P) and the sub-cooled (SUB) regions. From the point of view of heat integration, the parent stream is disregarded and instead, each of its sub-streams is treated as a separate stream with corresponding inlet and outlet temperatures and associated heat loads.

As per our strategy, sub-streams corresponding to all three phases exist in the model irrespective of actual traversal of phases by the parent stream at any feasible solution for the problem. Obviously, when a particular phase does not exist, we need to ensure that the corresponding sub-stream does not contribute to the heat integration calculations. This is done by setting the heat load of the sub-stream to zero and forcing the corresponding outlet and inlet temperature to be equal to a default value. These tasks are performed using a novel disjunctive model which is discussed in the next section.

4. Disjunctive Model for Phase Detection

The phases traversed by the parent stream while exchanging heat can be determined if the state of this stream at the inlet and outlet of the MHEX is known. Our model detects the state of the parent stream by comparing its temperature with dew and bubble point temperatures. The task of phase detection and making appropriate assignments involves combinatorial decision-making and this can be best accomplished using disjunctions and logic propositions (Raman and Grossmann, 1994). The complete model for phase detection consists of three components: disjunctions for phase detection at inlet and outlet, flash calculations for the 2-phase sub-stream that integrate with the disjunctions, and enthalpy calculations and evaluation of heat loads of the sub-streams.

The disjunction for phase detection consists of three terms corresponding to the three possible states. The disjunctions corresponding to the inlet and outlet state of the parent stream involved in the MHEX have the following form:

$$\begin{bmatrix} Y_{IN}^{SUP} \\ T_{IN} \geqslant T_{DP} \\ T_{in}^{sup} = T_{IN} \\ T_{in}^{2p} = T_{DP} \\ T_{in}^{sub} = T_{BP} \end{bmatrix} \bigvee \begin{bmatrix} Y_{IN}^{2P} \\ T_{BP} \leqslant T_{IN} \leqslant T_{DP} \\ T_{in}^{sup} = T_{DP} \\ T_{in}^{2p} = T_{IN} \\ T_{in}^{sub} = T_{BP} \end{bmatrix} \bigvee \begin{bmatrix} Y_{IN}^{SUB} \\ T_{IN} \leqslant T_{BP} \\ T_{in}^{sup} = T_{DP} \\ T_{in}^{2p} = T_{BP} \\ T_{in}^{sub} = T_{IN} \end{bmatrix} \tag{2}$$

$$\begin{bmatrix} Y_{OUT}^{SUP} \\ T_{OUT} \geqslant T_{DP} \\ T_{out}^{sup} = T_{OUT} \\ T_{out}^{2p} = T_{DP} \\ T_{out}^{sub} = T_{BP} \end{bmatrix} \bigvee \begin{bmatrix} Y_{OUT}^{2P} \\ T_{BP} \leqslant T_{OUT} \leqslant T_{DP} \\ T_{out}^{sup} = T_{DP} \\ T_{out}^{2p} = T_{OUT} \\ T_{out}^{sub} = T_{BP} \end{bmatrix} \bigvee \begin{bmatrix} Y_{OUT}^{SUB} \\ T_{OUT} \leqslant T_{BP} \\ T_{out}^{sup} = T_{DP} \\ T_{out}^{2p} = T_{BP} \\ T_{out}^{sub} = T_{OUT} \end{bmatrix} \tag{3}$$

where T_{IN} and T_{OUT} are inlet and outlet temperature of parent stream, T_{DP} and T_{BP} are the dew and bubble point temperature of parent stream and the remaining variables correspond to inlet and outlet temperatures of the proposed sub-streams.

The sub-stream corresponding to the 2-phase region has to be treated separately when the outlet of the parent stream lies in the 2-phase region, because both vapor and

liquid phases exist and flash calculations are required to evaluate the vapor and liquid fractions and their compositions. However, if the parent stream does not traverse the 2-phase region, it is well-known that the VLE equations for flash calculation have no feasible solution when operating in the single region. To handle this issue, we enforce the flash calculations even when the 2-phase region is not traversed, but operate it as a single saturated (vapor or liquid) phase at the boundary of the 2-phase region. Consequently, the VLE equations do not need to be relaxed at all irrespective of state of the parent stream at the inlet and outlet conditions. The inputs to the flash model change depending upon the values of variables supplied by the disjunctions as follows:

$$\begin{aligned} H_{Flash\ inlet} &= H_{IN} - H_V\left(T_{in}^{sup}\right) + H_{DP} + H_{BP} - H_L\left(T_{in}^{Sub}\right) \\ T_{Flash} &= T_{out}^{2p} \end{aligned} \tag{4}$$

Here, H_{IN} is the enthalpy of the parent stream at the inlet of the MHEX and is assumed to be supplied by the upstream process unit. H_{DP} and H_{BP} are the enthalpy of the parent stream at dew and bubble point conditions, respectively. H_V and H_{L} are enthalpy correlations corresponding to the vapor and liquid phases. The flash model operates in (P, T) mode, with the heat duty Q_F being calculated as an output variable.

The advantage of our proposed formulation is that all enthalpy calculations can be performed outside the disjunctions while picking appropriate enthalpy correlations for the vapor and liquid phases. The heat load for the sub-streams is given by following set of equations:

$$\begin{aligned} Q^{sup} &= H_V\left(T_{in}^{sup}\right) - H_V\left(T_{out}^{sup}\right) \\ Q^{sub} &= H_L\left(T_{in}^{sub}\right) - H_L\left(T_{out}^{sub}\right) \\ Q^{2p} &= \ -Q_F \end{aligned} \tag{5}$$

Equation (5) is applicable if the parent stream is a hot stream. For cold streams, we simply reverse the signs of the terms in the RHS of (5).

5. Reformulation of disjunctions

The disjunctions for phase detection can be reformulated as a mixed-integer nonlinear programming (MINLP) problem. However, it is also possible to reformulate it as a nonlinear programming problem and thus avoid the use of binary variables. This reformulation involves solving an inner minimization problem of the following form:

$$\begin{aligned} \text{Min} \quad & -\left[Y_{\text{sup}}\left(T - T_{DP}\right) + Y_{2p}\left(T_{DP} - T\right)\left(T - T_{BP}\right) + Y_{sub}\left(T_{BP} - T\right)\right] \\ \text{s.t.} \quad & Y_{sup} + Y_{2p} + Y_{sub} = 1 \\ & Y_{\text{sup}} \geq 0,\ Y_{2p} \geq 0,\ Y_{sub} \geq 0 \end{aligned} \tag{6}$$

which sets exactly one of Y_{sup}, Y_{2P}, and Y_{sub} to unity. The optimality conditions corresponding to (6) are given by following constraints:

$$\begin{aligned} Y_{sup} + Y_{2p} + Y_{sub} &= 1 \\ -\left(T - T_{DP}\right) - \mu_{\text{sup}} + \lambda &= 0 \\ -\left(T_{DP} - T\right)\left(T - T_{BP}\right) - \mu_{2p} + \lambda &= 0 \\ -\left(T_{BP} - T\right) - \mu_{sub} + \lambda &= 0 \\ 0 \leq Y_{sup} \perp \mu_{sup} &\geq 0 \\ 0 \leq Y_{2p} \perp \mu_{2p} &\geq 0 \\ 0 \leq Y_{sub} \perp \mu_{sub} &\geq 0 \end{aligned} \tag{7}$$

where $\lambda$and μ are multipliers corresponding to equality and non-negativity constraints respectively of (6). The last three equations of (7) are complementarity constraints and are handled using the penalty formulation (Baumrucker et al., 2008).

6. Case Study: Natural Gas Liquefaction

We demonstrate our proposed model for multi-stream heat exchangers using the commercial Poly Refrigerant Integrated Cycle Operations (PRICO) process for natural gas liquefaction (Price and Mortko, 1996). The PRICO process, shown in Figure 1 employs a single-stage mixed refrigerant (MR) system. Natural gas stream enters at 55 bar and 25°C and is liquefied by cooling it to -163°C. This cooling task is accomplished by using a refrigerant of certain composition which circulates in a closed loop. The objective is to determine the operating conditions and composition of the refrigerant that minimizes the compressor work. The posed problem has several challenges which are noted below:

1) Outlet of the cooler (stream S5) can be either superheated or 2-phase.
2) High pressure refrigerant outlet (stream S6) can be either 2-phase or subcooled.
3) Outlet of the throttle valve (stream S7) can be either 2-phase or subcooled.
4) All pressures and temperatures in the flowsheet are free and can be optimized.

The objective of minimizing the compression work in the PRICO process has been studied previously by Del Nogal et al. (2008). Their strategy involves two steps. In the first step, a genetic algorithm proposes a set of candidate solutions over a discretized space which is assessed and refined using an in-house simulator. In the second step, this set of solutions is used as initial guess for determining the optimal solution over the continuous space using standard NLP optimization. On the other hand, our formulation allows the use of an equation-oriented strategy for heat integration, phase detection as well as handling of nonexistent phases. Consequently, we can solve the same problem as a single medium-size equation-oriented NLP problem.

The comparison of our optimization results with that of Del Nogal et al. (2008) is shown in Table 1. As can be seen, our new methodology is able to find a better optimal solution that features more than 10% reduction in power consumption. It is also worth mentioning that our equation-oriented optimization strategy requires only two minutes as compared to 410 minutes by the integrated genetic optimizer-simulator framework of Del Nogal et al. (2008) on similar computer hardware.

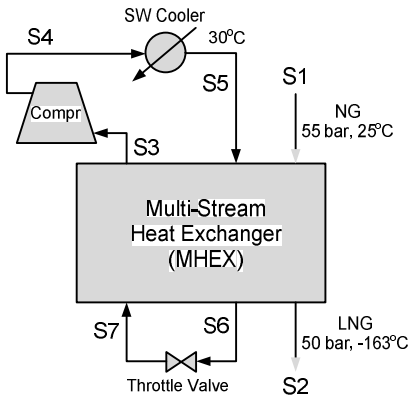


Fig. 1. PRICO process for liquefaction of natural gas

Table 1. Comparison of optimization results for the PRICO problem

	ΔT_{min} = 1.2°C		ΔT_{min} = 5°C	
	Nogal's work	This work	Nogal's work	This work
Power (MW)	24.53	21.51	33.49	28.63
Flow (kmol/s)	3.53	2.928	3.47	3.425
P_{Lower} (bar)	4.84	2.02	2.4	1.68
P_{Upper} (bar)	43.87	17.129	36.95	26.14
Refrigerant (mol %)				
N_2	10.08	5.82	15.32	12.53
CH_4	27.12	20.62	17.79	19.09
C_2H_6	37.21	39.37	40.85	32.96
C_3H_8	0.27	0.0	0.41	0.0
n-C_4H_{10}	25.31	34.19	25.62	35.42

7. Conclusions

Developing a process model for MHEXs is not trivial owing to issues like ensuring minimum driving force criteria and accounting for heat load of streams with or without phase changes particularly when the matches between the streams are not known *a priori*. This paper describes a new equation-oriented process model for MHEX which addresses all the previously mentioned issues. The process modeling is based on the fact that a MHEX can be regarded as a special case of a heat exchanger network that does not consume any utilities. Consequently, the model by Duran and Grossmann (1986) for simultaneous optimization and heat integration can be tailored for modeling MHEXs in the absence of phase changes. To handle phase changes, a novel strategy is proposed in which the streams involved in the MHEX are split into three sub-streams corresponding to the super-heated, two-phase and the subcooled regions. This is accomplished by using a disjunctive formulation which automatically detects phases and performs the appropriate flash and enthalpy calculations irrespective of the actual phases traversed by the streams. The model is demonstrated using a small case study involving liquefaction of natural gas. It is shown that this equation-oriented optimization strategy can lead to a significant reduction in computation time and even provide better solutions as compared to strategies used in previous work. It is expected that this model will find useful applications in process simulation and optimization of flowsheets having one or more MHEXs where the state of the streams entering and/or leaving the MHEX are not known *a priori* and can be optimized to achieve a desired objective.

References

Duran, M. A. and Grossmann, I. E. (1986). Simultaneous Optimization and Heat Integration of Chemical Processes. AIChE J., 32, 123-138.

Raman, R. and Grossmann, I. E. (1994). Modelling and Computational Techniques for Logic based Integer Programming. Comput. Chem. Eng., 18, 563-578.

Price, B. C. and Mortko, R. A. (1996). PRICO - A simple, flexible proven approach to natural gas liquefaction. In Proceedings of the 17th International LNG/LPG Conference, Gastech '96. Austria Centre, Vienna.

Baumrucker, B. T., Renfro, J. G., and Biegler, L. T. (2008). MPEC problem formulations and solution strategies with chemical engineering applications. Comput. Chem. Eng., 32, 2903-2913.

Del Nogal, F., Kim, J., Perry, S. and Smith, R. (2008). Optimal Design of Mixed Refrigeration Cycles. Ind. Eng. Chem. Res., 47, 8724-8740.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Steady-state optimization of a continuous pharmaceutical process

Dimitrios I. Gerogiorgis and Paul I. Barton

Department of Chemical Engineering, Massachusetts Institute of Technology (M.I.T.)
Landau Building (66-569), 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Abstract

Continuous Pharmaceutical Manufacturing can foster the production of key molecules, because it has a strong potential to improve product quality and suppress operating cost. Novel pharmaceutical production lines are thus ideal for exploring this paradigm shift, by introducing systematic process design, modern analytics and advanced control. Microreactor and microseparator engineering appears particularly advantageous in this high-value and low-capacity context, both at a pilot-plant and a full-production scale, and it can be combined with a plethora of proven continuous downstream technologies, facilitating the reduction of total organic waste emission and environmental footprint. The present paper focuses on flowsheet synthesis and optimization of an upstream process which is targeted at the production of a novel active pharmaceutical ingredient. The process flowsheet comprises four organic synthesis reactions and three separations, considering microsystems (microreactors and microseparators) for the respective steps. Conceptual synthesis allows for efficient separation and solvent integration, and indeed improves process economics, in contrast to wasteful batch pharmaceutical processing. Steady-state mass balances have been determined for the desired throughput and purity.

Keywords: Continuous Pharmaceutical Manufacturing (CPM), modeling, optimization.

1. Introduction and motivation

Continuous Pharmaceutical Manufacturing (CPM) attracts significant attention today, because of the expanding profitability gap experienced by the pharmaceutical industries; this is due to increasing R&D and operating cost and decreasing drug prices (Behr, 2004). Despite the strict licensing requirements for product purity and stability, and the established batch manufacturing processes, the FDA has recognized that CPM has a strong potential to improve product quality and suppress operating cost (Plumb, 2005). Comparative studies have confirmed this potential experimentally (Betz et al., 2003). Corporations must focus on introducing modern process engineering design principles, advances in instrumentation and control, and data management systems (FDA, 2004). The vast majority of *Active Pharmaceutical Ingredients (APIs)* are produced by batch processes, whose reconfiguration implies long, expensive clinical trials and revalidation. Nevertheless, novel APIs provide opportunities to explore a CPM paradigm, addressing systematically the design and operation of continuous pharmaceutical production lines. *Microreactor technology* offers a convenient, efficient, cost-effective way to conduct many organic synthesis reactions (Jensen, 2001); its advantages include miniaturization (thus space and fixed investment cost savings), high heat and mass transfer coefficients, scaleability, quick process condition screening, rapid manufacturing and customization. *Microseparators* are fabricated based on similar principles (Kralj et al., 2007), and can accelerate R&D for CPM processes, from laboratory to pilot-plant and production scale.

Organic synthesis is the cornerstone of CPM process development and flowsheeting, since a robust and efficient chemical reaction pathway is essential for proof of concept. In contrast to standard batch process prerequisites (high purity, conversion and yield), CPM processes can be flexible and viable at a much lower conversion and yield, as long as (a) continuous flow without precipitation and (b) efficient separations are ensured. Batch pharmaceutical manufacturing often relies on recipe-oriented R&D, conducted via many campaigns of arduous, expensive and time-consuming laboratory experiments.

Advances in unit operations engineering can also positively impact CPM processes, by improving efficiency as well as reducing solvent emissions and environmental footprint. Nanofiltration (Lin & Livingston, 2007), steady-state recycle (Kennedy et al., 2004) and continuous granulation (Vervaet et al., 2005) are examples of proven technologies which enhance the upstream (API production) and downstream (drug formulation) part of many pharmaceutical production lines, accelerating the transition to CPM processes. Concurrently, breakthroughs in Process Analytical Technology (PAT) (Munson et al., 2006), novel sensors and actuators provide in-depth understanding of chemical kinetics, thus enabling first-principles modeling and more efficient multivariate process control.

The present paper focuses on plantwide synthesis and optimization of an upstream process targeted at the production of a novel Active Pharmaceutical Ingredient (API). The conceptual flowsheet comprises 4 organic synthesis reactions and 3 separations, considering microsystems (microreactors and microseparators) for the respective steps. Conceptual synthesis allows for solvent integration (Ahmad & Barton, 1999) which can improve process economics, in contrast to wasteful and expensive batch processing. Steady-state upstream process optimization allows the determination of optimal stream flows and liquid-liquid extraction solvent flow, for specified API throughput and purity.

2. Pharmaceutical synthesis chemistry

The present paper considers a chemical reaction pathway for the continuous production of an API whose batch manufacturing has been successfully demonstrated and relies on 12 reaction steps and 11 intermediate separation steps, with the use of several solvents. The continuous version of the process comprises only 4 reaction steps, which occur in a polar solvent (acetonitrile, CH_3CN), under atmospheric pressure and room temperature. An aqueous API stream is significantly more preferable for downstream processing, due to the appreciable toxicity and strict limits for CH_3CN, a Class 2 solvent (EMEA, 2001). The use of relatively unstable diazomethane (epoxidation agent in the second reaction) dictates the installation of a sizeable adjacent continuous plant (Proctor & Warr, 2002). The first (addition) reaction combines 2 organic molecules to the solid intermediate C3:

$$C1 + C2 \xrightarrow[CH_3CN]{Pd,\ H_2O} C3 + HCl \quad (1)$$

The second (epoxidation) reaction activates the C3 polar end towards intermediate C4:

$$C3 + CH_2N_2 \xrightarrow[CH_3CN]{Et_2O} C4 + N_2 \quad (2)$$

The third (conversion) reaction nitrifies the same C4 polar end towards intermediate C5:

$$C4 + CH_3CN \xrightarrow{HCl} C5 \quad (3)$$

The fourth (hydrolysis) reaction occurs in acidic environment towards the final product:

$$C5 + 2H_2O + HCl \longrightarrow API + CH_3COOH \quad (4)$$

Evidently, API synthesis reactions are fostered by the acidic polar solvent environment and CH_3CN has a vital role (solvent as well as reagent) in the first 3 synthesis reactions.

3. Process synthesis and flowsheeting

The conceptual process flowsheet that has been synthesized for the production of API considers that organic synthesis can proceed continuously within a minimal number of unit operations (*4 reactors*, *3 separators*) and comprises a total of *20 streams* (Figure 1). *Microreactors* (R-001, R-002 R-003 and R-004) produce the first, second and third organic solid intermediates (C3, C4 and C5) and the final product (API), respectively. *Separator 1* (S-001) is the initial separator, removing the H_2O (used as proton acceptor) or the precipitating salt (if an organic base is used) during the first synthesis reaction. *Separator 2* (S-002,003) is the intermediate system separating the gas (N_2) from R-002, and subsequently removes (via microdistillation) the CH_2N_2 stream carrier (ether, Et_2O). *Separator 3* (S-004) is the final separator, removing CH_3CN to yield a pure API stream.

A detailed description of all streams encountered in the flowsheet (Figure 1) follows.

- *Flow 1* (F-001) contains the reagent (C1) solution in CH_3CN which is introduced into R-001.
- *Flow 2* (F-002) contains the reagent (C2) solution in CH_3CN which is also fed into R-001.
- *Flow 3* (F-003) contains the base (H_2O) that is necessary for HCl removal in the first reaction.
- *Flow 4* (F-004) contains all first reaction (R-001) compounds towards separation (S-001).
- *Flow 5* (F-005) contains the entrainer used for efficient separation of the base used in R-001.
- *Flow 6* (F-006) contains the bottoms which is drawn from S-001 with reaction by-products.
- *Flow 7* (F-007) contains the purified product stream which is introduced into R-002.
- *Flow 8* (F-008) contains the CH_2N_2 solution in ether (Et_2O) which is fed to R-002 from a line.
- *Flow 9* (F-009) contains all second reaction (R-002) compounds towards separation (S-002).
- *Flow 10* (F-010) contains nitrogen (N_2), which leaves the product using a degasser (S-002).
- *Flow 11* (F-011) contains a gas-free stream which is fed to S-003 for solvent (Et_2O) removal.
- *Flow 12* (F-012) contains the ether (Et_2O) which is separated out of the R-002 product stream.
- *Flow 13* (F-013) contains the purified stream which is fed to R-003 for the third reaction.
- *Flow 14* (F-014) contains the HCl which is introduced to R-003 to catalyze the third reaction.
- *Flow 15* (F-015) contains all third reaction compounds (from R-003) that are fed to R-004.
- *Flow 16* (F-016) contains the water that is fed to R-004 for the final hydrolysis reaction.
- *Flow 17* (F-017) contains all fourth reaction (R-004) compounds towards separation (S-004).
- *Flow 18* (F-018) contains the entrainer used for efficient organic solvent (CH_3CN) recovery.
- *Flow 19* (F-019) contains the bottoms drawn from S-004 (extraction solvent + most CH_3CN).
- *Flow 20* (F-020) contains the aqueous API solution, which is the upstream flowsheet target.

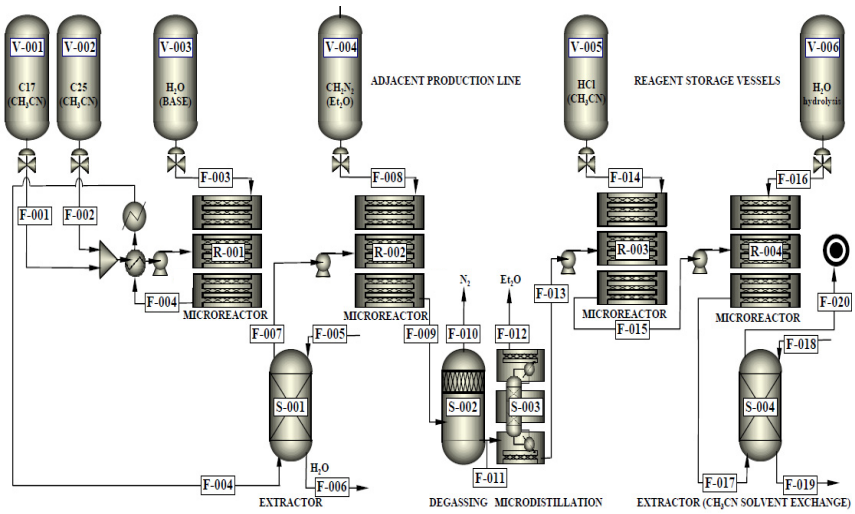


Figure 1: A conceptual upstream process flowsheet developed for the production of a novel API.

4. Mass balances for the upstream process

4.1. Mass balances: Necessary assumptions

A series of assumptions are necessary in order to determine the upstream mass balances:

4.1.1. Reactors – Kinetics

Organic synthesis reactions are irreversible without by-products, and a stoichiometric feed ratio is assumed for reaction 1 (R-001), to avoid carry-over of excess feedstocks. Both C1 and C2 solutions in acetonitrile are fed at a concentration of **3.75 M** (1:5 mol). Diazomethane (CH_2N_2) is fed at 100% excess to R-002 (**3.15 M**), to ensure a high yield. Hydrochloric acid (HCl) captured in S-001 is fed to R-003 at a concentration of **2 M**. Literature values of reaction yields are: $\mathbf{r_1}$**=0.53**, $\mathbf{r_2}$**=0.74**, $\mathbf{r_3}$**=0.90**, $\mathbf{r_4}$**=0.90**, respectively. The use of other CPM auxiliary solvents is not necessary, in contrast to batch schemes; reagent excess use is avoided, thus effectively reducing separator duties and total waste.

4.1.2. Thermodynamics – Separators

Nitrogen (N_2) is insoluble in ether and escapes the degasser (S-002) almost completely. *Water (H_2O)* is the inorganic basis used (R-001), and undergoes very efficient removal. For both, we assume that **95%$_{mol}$** escapes (only **5%** of material remains in the stream). Separator efficiencies are assumed as: $\mathbf{s_1}$ **= 0.70**, $\mathbf{s_3}$ **= 0.90** and $\mathbf{s_4}$ **= 0.90**, respectively. Thus, at separator **i**, $\mathbf{s_i}$% of undesired and (**1-**$\mathbf{s_i}$)% of valuable compounds are removed. Portions of solvents and carry-over chemicals remain inert in all subsequent reactions. *Solvent and reagent recycling* is not considered due to low (pilot plant) stream volumes. All unit operations (microreactors, separators) operate at atmospheric pressure, and all thermophysical properties for reagents and solvents are calculated at room temperature. The target API flowrate and purity in the final stream are given (**48 g.h^{-1}** and **> 40%$_{wt.}$**).

4.2. Mass balances: Results

The upstream process mass balances for the main streams are illustrated in Figure 2:

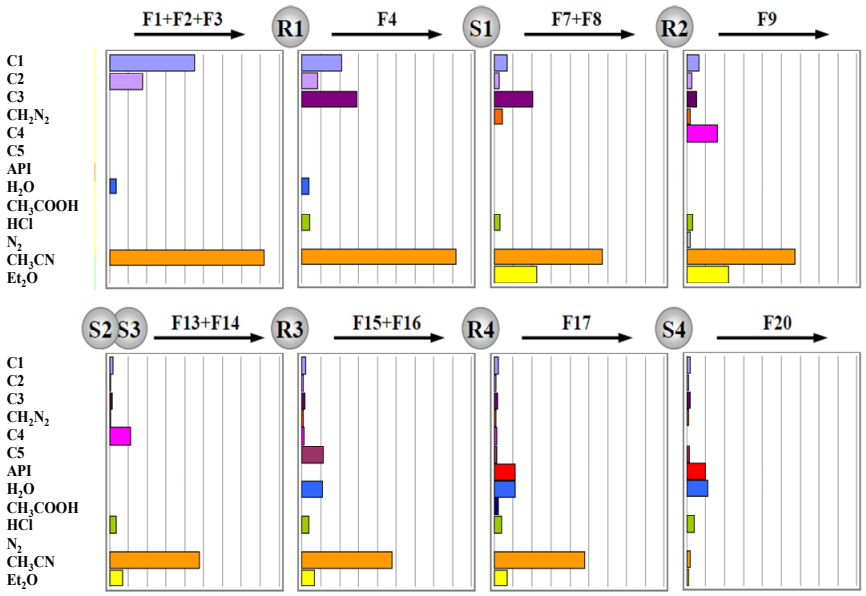


Figure 2: Schematic illustration of mass balances for the principal process streams (unit: 50 g.h^{-1}).

5. Liquid-liquid extraction (LLE)

The final separator (S-004) is the most challenging separation step of the novel process: the inlet stream contains the API (whose solution stability must be ensured, maximizing its outlet stream retention) dissolved with impurities in an acidic CH_3CN-H_2O mixture. This flow (F-017) contains traces of unreacted intermediates and reaction by-products. High API solubility is assumed therein, despite a known low solubility in pure CH_3CN, as it is very soluble in CH_3CN-H_2O-$HClO_4$ mixtures, so low pH is probably a stabilizer. A dense effluent (F-020) reduces operating cost and facilitates downstream processing.

5.1. LLE solvent selection

The ideal LLE solvent has flow and heat stability (low melting and high boiling point), acceptably low viscosity (for flow), high density (for mixing and reduced fixed cost) and most importantly, a high distribution factor (m) for acetonitrile (CH_3CN) extraction. Furthermore, a low aqueous saturation concentration and a low toxicity are desirable. Industrial CH_3CN-H_2O separations rely on butyl acetate, which is prohibitively toxic. Evaluating the respective properties of several organic solvents leads to the selection of *toluene* (C_7H_8), which has an acceptable toxicity and an outstanding distribution factor, determined by linear approximation from literature data (Di Cave & Mazzarotta, 1991). A compilation of important physicochemical properties of candidate solvents and the ternary CH_3CN-C_7H_8-H_2O phase diagram by Rao et al. (1979) are presented in Figure 3.

CANDIDATE SOLVENT	MW (g/mol)	MP (°C)	BP (°C)	ρ (g/cm³)	μ (cP)	$C_{sat}^{H_2O}$ (g/l)	m_{ACN} (i / H_2O)	CLASS	TOXIC (< ppm)
n-HEXANE	86	-95	69	0.654	0.29	~0	()	2	290
CYCLOHEXANE	84	7	81	0.779	1.02	0	()	2	3880
BENZENE	78	6	80	0.878	0.65	0.80	2.50	1	2
TOLUENE	92	-93	111	0.866	0.59	0.47	2.22	2	890
o-XYLENE	106	-25	144	0.880	0.81	~0	0.50	2	2170
m-XYLENE	106	-48	139	0.860	0.62	~0	0.66	2	2170
p-XYLENE	106	13	138	0.860	0.34	~0	0.86	2	2170
ACRYLONITRILE	53	-84	77	0.810	1.30	70	E.D., T	1	1
BUTYL ACETATE	116	-74	126	0.880	0.73	7	1.80	3	5000

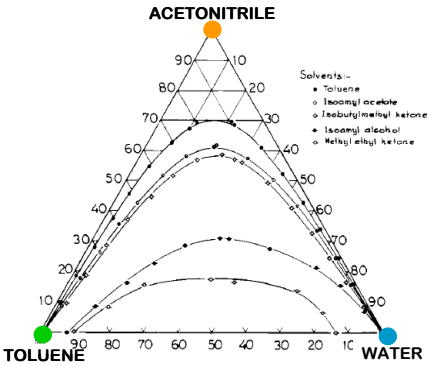


Figure 3: Liquid-liquid extraction: solvent selection and the CH_3CN–C_7H_8–H_2O ternary system.

5.2. Determination of LLE solvent duty

The liquid-liquid extractor (S-004) can operate in crosscurrent or countercurrent mode. The first offers higher flexibility, faster dynamics and is thus suitable for a pilot plant; the second balances mass transfer, economizes on solvent and is suitable for production. Considering a pure solvent feed (S), solvent equidistribution (crosscurrent), no solvent blowdown (countercurrent) and solvent-feed immiscibility, the reduction ratio X_F/X_R is:

$$\frac{X_F}{X_R}=\left(1+\frac{mS}{nF}\right)^n \quad \textit{(crosscurrent LLE)} \quad (5) \qquad \frac{X_F}{X_R}=\frac{\left(\frac{mS}{F}\right)^{n+1}-1}{\left(\frac{mS}{F}\right)-1} \quad \textit{(countercurrent LLE)} \quad (6)$$

The pure solvent duty (S) can be determined as a function of feed (F) and stages (n) for a high reduction ratio ($5\cdot10^4$), due to the stringent residual solvent limits (EMEA, 2001). For $\{2,3,4\}$ stages, the crosscurrent solvent duty is $\mathbf{S=\{200.5,48.4,25.1\}\,g\,C_7H_8/g\ feed}$, while the countercurrent solvent duty is much lower: $\mathbf{S=\{100.5,16.4,6.6\}\,g\,C_7H_8/g\ feed}$. Hence, S-004 requires several stages in both pilot-plant and production scale operation.

6. Conclusions

A conceptual upstream CPM process and its mass balances clearly indicate that the API of interest can be continuously manufactured at the prescribed flow rate and purity, and that the use of acetonitrile and toluene as reaction and separation solvent, respectively, have a clear impact on minimizing the number of process steps, reducing flowsheet size (fixed investment cost), and suppressing feedstock and solvent duties (operating cost). The proposed continuous flow chemistry is based on a concise reaction pathway, and *toluene* is the most efficient extraction agent among all 9 candidate solvents considered. A CPM flowsheet is presented here, and detailed mass balances have been calculated. The API capacity considered is **48 g** per hour of operation **(F-020: $\mathbf{m_{API} = 48\ g.h^{-1}}$)**, and *the product stream* (F-020) contains **48.19 g $\mathbf{h^{-1}}$** of API in a total flow of **109.02 g $\mathbf{h^{-1}}$**, with several compound traces (unreacted raw materials, intermediates, acids, solvents) which can be removed in the next downstream processing steps (crystallization, drying). The *total effluent waste* is: **$\mathbf{F_W = 15.58\ mol.h^{-1}}$ ($\mathbf{830.13\ g.h^{-1}}$) ($\mathbf{76\%_{mol}}$=$\mathbf{64\%_{wt}}$** solvents), thus *an order of magnitude lower* than that generated by a batch manufacturing process. The waste is almost 8 times the product stream in a flowsheet without solvent recycling (1.46 mol=**$\mathbf{107.6\ g.h^{-1}}$ $\mathbf{Et_2O}$**, 10.36 mol=**$\mathbf{424.6\ g.h^{-1}}$ $\mathbf{CH_3CN}$**); the latter is not pursued, because installing additional unit operations can adversely affect pilot-plant economics. The cost savings shown here will be augmented in the case of a production-scale plant.

Acknowledgements

The authors hereby acknowledge gratefully financial support from Novartis Pharma AG.

References

B.S. Ahmad, P.I. Barton, 1999, Process-wide integration of solvent mixtures, *Comput. Chem. Eng.* **23**(10): 1365-1380.

A. Behr, 2004, New developments in chemical engineering for the production of drug substances, *Eng. Life Sci.* **4**(1): 15-24.

G. Betz, P. Junker-Purgin, H. Leuenberger, 2003, Batch and continuous processing in the production of pharmaceutical granules, *Pharm. Dev. Technol.* **8**(3): 289-297.

S. Di Cave, B. Mazzarotta, 1991, Liquid-liquid equilibria for ternary systems formed by acetonitrile, water, and aromatic hydrocarbons, *J. Chem. Eng. Data* **36**(3): 298-303.

European Medicines Agency (EMEA), 2001, *Impurities: Residual solvents*, VICH Topic GL18 (plus annexes) (Web: http://www.emea.europa.eu/pdfs/vet/vich/050299en.pdf)

U.S. Food and Drug Administration (FDA-CDER), 2008, *Innovation & continuous improvement in pharmaceutical manufacturing*. (Web: www.fda.gov/cder/gmp/gmp2004/manufSciWP.pdf)

K.F. Jensen, 2001, Microreaction engineering: is small better?, *Chem. Eng. Sci.* **56**(2): 293-303.

J.H. Kennedy, M.D. Belvo, V.S. Sharp, J.D. Williams, 2004, Comparison of separation efficiency of early phase active pharmaceutical intermediates by steady state recycle (SSR) and batch chromatographic techniques, *J. Chromatogr. A* **1046**(1-2): 55-60.

J.G Kralj, H.R. Sahoo, K.F. Jensen, 2007, Integrated continuous microfluidic liquid-liquid extraction, *Lab Chip* **7**: 256-263.

J.C.T. Lin, A.G. Livingston, 2007, Nanofiltration membrane cascade for continuous solvent exchange, *Chem. Eng. Sci.* **62**(10): 2728-2736.

J. Munson, C.F. Stanfield, B. Gujral, 2006, A review of process analytical technology (PAT) in the US pharmaceutical industry, *Curr. Pharm. Anal.* **2**(4): 405-414.

K. Plumb, 2005, Continuous processing in the pharmaceutical industry - Changing the mind set, *Chem. Eng. Res. Des.* **83**(A6): 730-738.

L.D. Proctor, A.J. Warr, 2002, Development of a continuous process for the industrial generation of diazomethane, *Org. Proc. Res. Dev.* **6**(6): 884-892.

D.S. Rao, K.V. Rao, A.R. Prasad, C. Chiranjivi, 1979, Extraction of acetonitrile from aqueous mixtures. 2. Ternary liquid equilibria, *J. Chem. Eng. Data* **24**(3): 241-244.

C. Vervaet, J.P. Remon, 2005, Continuous granulation in the pharmaceutical industry, *Chem. Eng. Sci.* **60**(14): 3949-3957.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

"Smelly" Pinch

Toshko Zhelev[a], Natasha Vaklieva-Bancheva[b], Jaime Rojas-Hernandes[a], Tony Pembroke[a]

[a]*University of Limerick, Ireland*
[b]*IChE, Bulgarian Academy of Sciences*

Abstract

This paper attempts to make one further step in energy efficiency improvement of wastewater treatment plants connecting it with environmental considerations, namely the suppression of unpleasant aromas. Solution of the optimum heat integration when the off-gas from reactors is considered as heat source together with the effluent heat, is presented. Further, the thermal management is linked to absorption efficiency through the classical concept of simultaneous heat and mass transfer. The conceptual approach of heat integration applied through Pinch analysis is extended to consider the mass exchange through a combined operating composite and equilibrium curve plot allowing for adjustment of uncertain variables, such as streams target temperatures and stripping water flowrate. An case study from an autothermal anaerobic digesting plant (ATAD) in Ireland is presented and the future plans of this ongoing project are discussed.

Keywords: ATAD, Pinch, efficiency, heat integration, aroma.

1. Introduction

The new addition to the bio-energy utilisation problem for ATAD treatment, discussed earlier by Rojas-Hernandes et al., 2008 is the source of heat associated with the gas, drawn from all reactors and from the sludge storage, which is supplied for ammonia absorption.

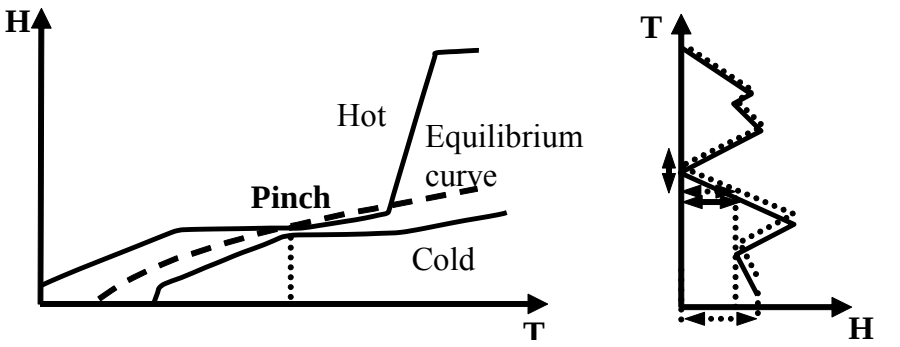
The specifics of the problem of bio-energy recovery justify certain uncertainties in the estimation of the final temperatures of majority of streams participating in the heat integration task. The cooling level of the effluent stream can be specified as "as low as possible". The heating of the fresh sludge stream – as "as high as possible". The same is the case with the air for aeration and the desired temperature of the exhaust gas.

Naturally the thermodynamic feasibility restricts the extreme level of these temperatures. The cold sludge temperature can not be higher than the heating effluent as the effluent target temperature can not be lower than the cold sludge supply temperature. The exhaust gas temperature can not be lower than the wastewater ambient conditions justified temperature or the temperature of the aeration air can not get higher than the gas stream extraction temperature.

2. Simultaneous Heat and Mass Transfer

2.1. Background

Apart from the obvious grass-root optimisation solution to the energy integration problem we propose to apply in parallel an attractive graphical approach, which was proposed by Zhelev & Semkov, 2004 for the purpose of efficiency management of an economiser system for flue gas energy recovery.



(a) Reversed Composite curve (b) Grand Composite Curve

Figure 1: Composite curves for simultaneous heat and mass integration

Combined heat-mass transfer targeting: - maximum energy recovery; - maximum potential of recovered heat

According to the concept of the H/T composite curve, streams' temperatures have to be adjusted to ensure the equilibrium curve does not cross any of the hot or cold composites. The Grand composite curve, constructed as suggested by Pinch analysis (Smith, 2001) based on "shifted composites" can be extended and used to target the maximum energy interchange and the level at which the cold sludge can be preheated. It will correspond to the highest level of the Pinch (actual level – highest temperature of "cold utility"), restricted by the beginning of the opening of the rand composite at the top.

3. Smell

Now, what about the smell?
The function representing the change of the absorbability of ammonia in water with temperature is presented below. The ideal gas method can predict well the mass absorption rate.

$$P_v(t)V = m_v(t)RT_i \Rightarrow m_v(t) = \frac{P_v(t)V}{RT_i} \quad (1)$$

Where the total vapor volume, $V = 1.55\text{x}10^{-3}$ m^3.
If we substitute the existing temperatures of the gas stream with the temperatures corresponding to the environmental standards constrained ammonia concentration (less than 1 ppm as mentioned by Taghirpour, 2008), the solution of the heat integration problem in general will change and then the recovery of heat will be aroma-driven.

4. Heat Integration

Two options were analysed:

(a) heat integration utilising the sensible heat of the exhaust gas;
(b) heat integration utilising the sensible and the latent heat of the exhaust gas.

The second option envisages the inclusion of the absorber in the integration scheme. Because of the simultaneous heat and mass transfer this increases the complexity of the formulated problem.
This will allow us to consider the efficiency of the aroma absorption in parallel to the heat recovery.

4.1. Sensible heat utilization

The proposed heat integration scheme is presented in Figure 2.

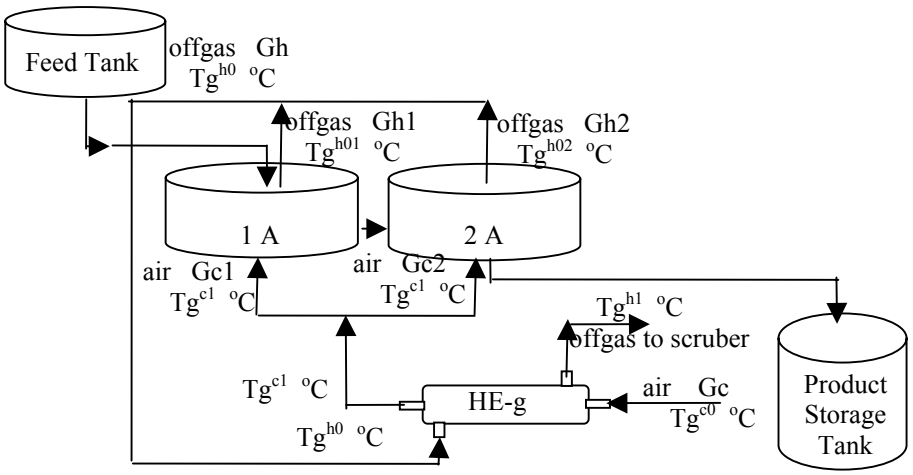


Figure 2: Heat integration scheme including the gas and the aeration air streams

The gas exiting reactors 1-A and 2-A contains air, water vapour, ammonia and hydrogen sulphide. Next it enters a scrubber for absorption of aromas. The gas leaves the reactors at the operational temperature of the sludge there, i.e. around 55°C, which represents a promising source of energy. We propose the following recovery scheme: Continuous stream $Gh1$ [m^3/day] of a warm gas at temperature $Th1$ [^oC], leaving 1-A is merging the warm gas extracted from reactor 2-A at $Gh2$ [m^3/day], $Th2$ [^oC]. The joined stream has a flow of Gh and temperature of $Th0$ and is used to preheat the joined stream of cold air (Gc [m^3/day], $Tc0$ [^oC]), used for aeration of both reactors. The data for the streams involved in the heat integration are given in the table below.

Stream	Total flow G [m^3/day]	Integrated G [m^3/day]	Process duration	Duration of integration τ	T^{o}C
Offgas 1-A	6210	5807,75 (Gh1)	24 h	22h 31 min	50
Offgas 2-A	6390	5982,62 (Gh2)	24 h	22h 31 min	60
Air to 1-A	6210	6110,57 (Gc1)	22h 54 min	22h 31 min	12
Air to 2-A	6390	6364,02 (Gc2)	23h 16 min	22h 31 min	12

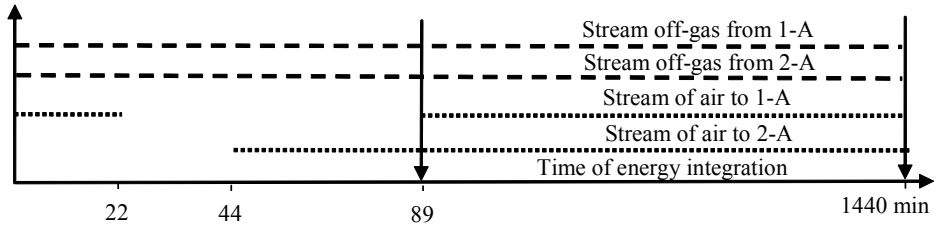


Figure 3: Gantt chart of the process

The specific heat capacity and the density of the two gas streams are ($Cph = Cpc = 1012$ [J/kg.°C]; $\rho h = \rho c = 1{,}225$ [kg/m^3]. The relative humidity of the warm stream is $\varphi = 100\%$, when that of the cold air is assumed as $\varphi = 80\%$.

Determined are: (1) The lowest achievable temperature of cooling the reactors' exhaust gas; (2) Energy saving as result of the heat recovery from the exhaust gas.

4.1.1. The lowest achievable temperature of cooling the reactors' exhaust gas

4.1.1.1. Mathematical model of the heat recovery

As the inlet streams temperatures are known, the target temperatures are determined as a function of the heat exchange surface area as follows:

$$\mathrm{Th1} = \mathrm{Th0} - (\mathrm{Th0} - \mathrm{Tc0}).\frac{1 - e^{-mUA}}{1 - \frac{Wh}{Wc}e^{-mUA}}, \quad (2)$$

and that of the preheated cold air is:

$$\mathrm{Tc1} = \mathrm{Tc0} - \frac{Wh}{Wc}(\mathrm{Tc0} - \mathrm{Th0}).\frac{1 - e^{-mUA}}{1 - \frac{Wh}{Wc}e^{-mUA}}, \quad (3)$$

where:

$$Wh = \frac{Gh1 + Gh2}{\tau}.Cph\,,\ Wc = \frac{Gc1 + Gc2}{\tau}.Cpc\,,\ \text{and}\ m = \frac{1}{Wh} + \frac{1}{Wc}.$$

4.1.1.2. Constraints

The heat exchange has to be physically feasible:

$$\Delta Th = Th0 - Tc1;\ \Delta Tc = Th1 - Tc0;\ \Delta T\_Eg = \min\{\Delta Th,\ \Delta Tc\},$$
$$\Delta T\_Eg = \Delta Tmin\,. \quad (4)$$

The heat exchange surface area is fixed within preliminary given upper and lower bounds (restricted by manufacturing/engineering feasibility):

$$A^{min} \le A \le A^{max}. \quad (5)$$

4.1.1.4. Objective function

The goal is to cool the reactors' exhaust stream to the lowest possible temperature. Thus, equation (2) of the model will be used as objective (efficiency) function.

$$\underset{A}{F} = \min\ Th1. \quad (6)$$

4.1.1.5.Results

To study the influence of $\Delta Tmin$ the optimisation problem was solved at three different values for the economically justifiable $\Delta Tmin$ equal to 5, 7 и 10 degrees. The interval of acceptable changes of the heat exchange surface area is assumed to be between 0 and 50 [m^2]. The results are given in table below.

$\Delta Tmin$	Th1 [°C]	Tc1 [°C]	A [m^2]
5	17	48	5.245
7	19	46.1	3.741
10	22	43.3	2.505

It is clear that a substantial decrease of the temperature of the exhaust supplied to the scrubber is achievable (from 55.1 to 17.2 or 22.0°C). Achieved lower temperature will improve the conditions of the scrubber. If assumed that the water temperature in the scrubber is in the range of 10-20°C, then the solubility of the ammonia will be between

59.7 and 52.6 g in 100 g water, which compared to the solubility at 40.3-30.7 at temperature 30-40°C will demonstrate the true ecological benefit of proposed heat integration. For the hydrogen sulphide the situation is similar: at a temperature of 10-20°C the volume of the gas resolved in a unit of pure water would be between 2.9-2.6, when at 30-40°C, it would be between 2.0-1.7.

4.1.2. Energy savings resulting from the heat recovery from the exhaust gas

The target (final) temperatures of the preheated cold aeration air are the base for calculation of the energy saving in the reactors resulting from heat integration. Assuming the relative humidity at 80% (average for Ireland) and using Ramsin's diagrams of most air we read the humidity x=0.007 [kg water/kg of dry air] at the temperature $Tc0$ of the cold air. Next, the enthalpy of the moist air will be:

$$Ic0 = \left(1000 + 1.97 * 10^3 * x\right) Tc0 + 2493 * 10^3 * x \text{ [J/kg of dry air].} \quad (7)$$

Analogically the enthalpy of preheated in the heat exchanger air will be:

$$Ic1 = \left(1000 + 1.97 * 10^3 * x\right) Tc1 + 2493 * 10^3 * x \text{ [J/kg of dry air].} \quad (8)$$

	$Tc0$=12	$Tc1$= 48 at $\Delta Tmin$ =5	$Tc1$=46.1 at $\Delta Tmin$ =7	$Tc1$ = 43.3 at $\Delta Tmin$ =10
$Ic0$	$2.962x10^4$			
$Ic1$		$6.611x10^4$	$6.419x10^4$	$6.135x10^4$

The difference between $Ic1$ (eq. (8)) and $Ic0$ (eq.(7)) multiplied by the flow of the supplied for aeration air will give the amount of the saved heat for the period of integration in [J].

	$Tc1$ at $\Delta Tmin$ =5	$Tc1$ at $\Delta Tmin$ =7	$Tc1$ at $\Delta Tmin$ =10
Energy-saved [J]	$5.577x10^8$	$5.283x10^8$	$4.849x10^8$

Knowing that, 1 kWh = $3.6x10^6$ J, the energy, saved in each of the three cases of assumed $\Delta Tmin$ will be:

	$Tc1$ at $\Delta Tmin$ =5	$Tc1$ at $\Delta Tmin$ =7	$Tc1$ at $\Delta Tmin$ =10
E-saved [kWh]	154.921	146.745	134.695

It is known that the compensation of the temperature shock in the reactors due to the cold feed load for the winter period requires extra energy associated with prolonged agitation equivalent to some 1000 kWh for the two reactors. Therefore the additional investment for heat exchange equipment can be compensated by achieved daily electrical saving of 13-15%.

4.2. Sensible and Latent heat utilization (Conceptual approach)

As mentioned by (Zhelev, 2008), the Enthalpy-Temperature plot can help to reasonably manage the mass and energy targeting escaping from complex rigorous optimisation. The lack of precision in finding optimum solution in this case is compensated by better control of solutions using engineering inside.

Note, the scrubber/absorber may play the role of humidifier or dehumidifier, which depends on the internal conditions. At present it may very well be a dehumidifier, because the temperature of the gas is higher than the water temperature and in winter conditions the water temperature can very well reach temperatures below the dew point of the water vapour. If the temperature of the gas is lowered by utilisation of its energy through heat integration, then this will represent a humidification scenario.

In the gas-heat integration case the pre-heated air for aeration will be supplied to the hotter reactor sludge, which will cause condensation of moisture contained in the air, accompanied by release of extra heat (compensated by the sensible air heating by sludge). In any case this process has to be carefully monitored because of the danger of foam formation – another not very well-understood phenomenon.
The composite curves constructed for the case study demonstrate the specifics of the heat integration task. It appears to be a typical threshold Pinch problem from both sides (absence of hot or cold utility requirements). This is expected because the wastewater treatment plant naturally does not use external heating, and at the same time no cooling would be expected.
Proposed graphical methodology allows manipulation of control variables. The complete set of "free" variables is: Target temperature of the effluent; Cold air (for sludge aeration) target temperature; Cold sludge target temperature; Stripping water target temperature; Stripping water flowrate.
Proposed manipulations are related to the amount of energy recovery and the efficiency of absorption, which would depend on the stripping water flowrate maintaining the exit ammonia concentration at 1 ppm as per environmental standards. The operating conditions will be influenced by the ambient conditions and this can be taken care of by the control system. Our manual evaluation system is based on the reversed composite curves plotted together with the equilibrium curve. The starting point of this plot needs fixed supply temperatures, target temperatures and flowrates.
These are fixed as follows: Target temperature of the effluent = 19°C (max cold sludge temp + ΔT_{min} (5°C)); Cold air (for sludge aeration) target temperature (50°C); Cold sludge target temperature (57°C); Stripping water target temperature (16°C); Stripping water flowrate (from the energy balance at all fixed temperatures).
"Dragging" the kinks of the composite curve one can try escape from crossing or touching and to achieve as much as possible favourable distribution of driving forces along the stripper. An important part of this temperature/flow re-adjustment procedure would be not to affect the earlier assigned minimum level of ammonia concentration. Here we have to stress that the Enthalpy/Temperature drawing would demonstrate in more details of gas/water mixture condition, then the gas concentration changes.

5. Conclusion and Future Work

Presented work combines rigorous mathematical optimisation with conceptual targeting in energy efficiency improvement accounting for the unpleasant aromas associated with wastewater treatment. Our future work is related to the discrete character of the reactor feeding at present. The feeding pattern, i.e. the frequency and the volume of feeding portions are mixed with the continuous character of aeration/gas extraction and absorption.

References

Zhelev, T.K. and K.A. Semkov, 2004, More Efficient Flue Gas Energy Recovery through Pinch, Journal of Cleaner Production, 12, 165-170.
Zhelev, T., N. Vaklieva-Bancheva and D. Jamniczky-Kaszás, 2008, About Energy Effic. Improvement of Auto-thermal Thermophilic Aerobic Digestion Processes, Com. Aided Chem.Eng, 25, Elsevier, ISBN: 978-0-444-53228-2, T6: CAPE and Society, 1-6.
Taghipour, H., M.R. Shahmansoury, B. Bina and H. Movahdian, 2008, Operational parameters in biofiltration of ammonia-contaminated air streams using compost–pieces of hard plastics filter media, Chemical Engineering Journal, 137, 2, 198-204.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Supporting Chemical Process Design under Uncertainty

Achim Wechsung[a,b], Jan Oldenburg[a], Joanna Yu[a], Axel Polt[a]

[a]*BASF SE, 67056 Ludwigshafen, Germany*
[b]*currently: MIT, Dept. of Chem. Eng., 77 Mass. Ave, Cambridge, MA 02139, USA*

Abstract

A major challenge in chemical process design is to make design decisions based on partly incomplete or imperfect design input data. Still, process engineers are expected to design these processes safe, dependable and cost-efficient. The complexity of typical process models limits intuitive engineering estimates to judge the impact of uncertain parameters on the proposed design. In this work, we present an approach to quantify the effect of uncertainty on a process design in order to enhance comparisons among different designs. To facilitate automation, a heuristic to differentiate between numerical and physical infeasibility when simulations do not converge is introduced. It is shown how this methodology yields more details about limitations of a studied process design.

Keywords: Chemical process design, Modeling and simulation, Uncertainty, Process performance measures.

1. Introduction

Process engineers are expected to design safe, dependable and cost-efficient processes. During the stage of conceptual process design, most design decisions are yet to be made, as the design will become more detailed. As each decision at this early stage will limit future choices and also reduce scope of future decisions, one still has a comparable large leverage on overall cost at this moment, but also lacks detailed information. On top, chemical processes are typically complex networks where effects of modifications at one unit operation on the overall process performance are not intuitive. If uncertainty of a subset of model determining parameters is incorporated in the modeled process, complexity increases further. In the end, impact of changes to underlying assumptions of the process model is even less intuitive—even to an experienced process engineer. Hence, providing engineers with a systematic method to foster understanding of implications of uncertain parameters on their current design will facilitate better decision-making in process design.

Uncertainties are introduced in process design in many ways, e.g. by insufficient knowledge about chemical reactions, by limited thermodynamic models, by varying raw material purities or by inadequate cost estimates for raw materials. At the same time, the designed process is nowadays expected to be less accommodating to unforeseen risks or deviations from design conditions using costly means of equipment overdesign.

Commonly, commercial deterministic process simulation and/or optimizing software (e.g. Aspen Plus) is used to design processes beginning at very early stages of the design process. As these simulators require exact values, customarily, engineers first approximate uncertainties using likeliest values to set up the process models. In a possible second step, they can estimate the impact of their limited knowledge about several model parameters by manually perturbing these and studying the effect on their design in a

very laborious and time-consuming matter, esp. since one is generally interested in more than a single uncertain parameter. As early as in the 1970's, first publications aimed at moving beyond simple perturbation schemes. These were mainly motivated by better justifying—and possibly reducing—overdesign factors used in the design of chemical processes (Freeman and Gaddy, 1975; Grossmann and Sargent, 1978; Grossmann *et al.*, 1983).

Building upon Freeman and Gaddy's approach, we employ a Monte Carlo scheme to study the impact of uncertain parameters on the design of a process plant. Utilizing i-TCM (Wiesel and Polt, 2006), processing equipment and its limitations is modeled in Aspen Plus using short-cut sizing methods. The simulation runs are controlled via an Excel interface. One of the greatest obstacles of automated process simulations is the difficulty to differentiate between physical infeasibility and numerical difficulties within the solver. We propose a heuristic priority scheme that supports the user in significantly increasing the number of converged runs by augmenting the optimization with a tailored relaxation scheme. With this procedure, the process engineer gains insight into the limitations of the process without the need of interpreting the results of failed simulation runs–in itself a difficult and sometimes even hopeless task in times of equation-oriented simulation.

The present contribution is organized as follows: First, the dependability analysis is presented in Section 2. Then, a case study will illustrate the application of the method in Section 3. Our work is summarized in Section 4 and possible extensions are discussed.

2. Dependability Analysis

When all parameters to a process are known exactly, the optimal design for this process can be obtained performing a single optimization. In case of uncertain parameters with a known probability distribution, computing the optimal design is more involved as design of a chemical process turns into a two-stage problem (Malik and Hughes, 1979). Expected values will have to be computed numerically, e.g. using stochastic techniques, leading to the repeated requirement to solve the inner optimization problem that selects for the operating conditions. Hence, determining an optimal design using rigorous optimization methods becomes computationally prohibitive when studying complex processes. One usually abandons the requirement to solve both optimization problems rigorously. It is common industry practice to perform the outer optimization, which selects the best flowsheet and sizes the equipment, manually based on engineering insight while using commercial process simulation and/or optimization tools, e.g. Aspen Plus, to solve the inner optimization problem.

When several uncertain parameters are involved, it becomes increasingly difficult to compare designs manually. Especially the need to compute expected values as a basis of comparison is cumbersome when performing this task manually. Also, resorting to perturbation analysis will only yield a qualitative understanding of differences between competing designs. Hence, being able to use a quantitative measure as a result of automatically performed calculations is advantageous in designing an optimal process.

Freeman and Gaddy (1975) propose such a measure that they named dependability, defined as the probability that the process meets its specification,

$$D(x_d, s) = \int_{x_0} P(x_0)\, \delta(x_0, x_d, s)\, dx_0 \ , \qquad (1)$$

where x_d notes the design variables, s the process specifications, x_0 the uncertain parameters, $P(x_0)$ the probability density function for x_0 and $\delta(x_0, x_d, s)$ states if the design meets the specifications under the assumed uncertain parameters $(\delta = 1)$, or if does not ($\delta = 0$). $P(x_0)$ will be known in advance, whereas $\delta(x_0, x_d, s)$ is the result of solving the inner optimization problem for one sample of the uncertain parameters and one particular design; it is not trivially accessible a priori.

D can also be interpreted as the percentage of time when the plant is operating in specification (Freeman and Gaddy, 1975). In many cases, $D < 1$ does not necessitate that the process will not be able to meet its specification on average. On the contrary, a plant with a smaller additional design margin may still be able to make up for times of reduced production when the varying parameters are in its advantage at a later time. Also, $D < 1$ does not entail that the plant is not down at some instances throughout the year, though product quality or quantity may be reduced. In contrast, methods requiring the process to stay within its specifications at all times (e.g. Grossmann *et al.*, 1983) lead to greater equipment overdesign to fulfill these conditions.

2.1. Determining Dependability using a quasi Monte Carlo Method

In order to obtain an estimate for the dependability, the integral in Eq. 1 has to be solved. When examining the integrand, one notes that δ is a complicated function and it is in most cases not known upfront. On the other hand, the probability density function for the uncertain parameters is specified when the problem is set up. Though the integral is too complex to be solved analytically, it can be solved by numerical means. Monte Carlo (MC) methods are widely used to solve such—possibly multi-dimensional—integrals in supplementing decision-making (Sprow, 1967).

In contrast to random MC sampling which use randomly selected samples in parameter space according to the probability distribution function, quasi-Monte Carlo (QMC) methods uses a deterministic number sequence that is chosen explicitly to increase rate of convergence. Latin hypercube sampling, an example for a QMC method, divides parameter space in each dimension in equally probable sections and samples only once from each section. It has been shown that QMC methods are capable of outperforming MC methods in terms of obtaining more accurate results with the same number of evaluations (Morokoff and Caflisch, 1994). For this contribution, a Hammersley sequence was used to obtain the number sequence for the QMC method. Diwekar and Kalagnanam (1997) proposed this sequence as it can easily be expanded to higher dimensions. The authors also presented evidence for favorable convergence properties of their proposed sequencing method.

2.2. Implementation of dependability analysis

Dependability is accessible to determination with the tools available to process engineers. In contrast to rigorous methods that depend on specialized optimization software, the stochastic approach can be carried out using process simulation tools in combination with standard office software. The above outlined method can be implemented in commercial process optimization tools, in our case Aspen Plus, in connection with an external controller for the MC simulations, here Microsoft Excel. Simulations and their data in Aspen Plus are accessible from Excel using an ActiveX automation server interface. Additionally, routines to determine the Hammersley sequence, to transform this uniform sequence to the specified probability distribution of the uncertain parameter, and to control runs in Aspen Plus are implemented in Visual Basic for Applications (VBA) in Excel. Furthermore, post-processing and visualizing of accumulated data is also included in the Excel workbook. In detail, for each individual Aspen Plus run, information on the

values of the sampled parameters, selected computed variables, and status of the run are reported to the user as a reference. These data are aggregated and reported for each variable as mean and standard deviation. Furthermore, dependability and fraction of converged simulations are provided.

2.3. Aid in Differentiating between Numerical and Physical Infeasibility

One of the major drawbacks in early attempts was the lack of robustness when solving the optimization problem for the process flowsheet. Since only converged results—regardless of their success—provide insight into physical constraints of the problem, it is essential to reduce infeasibility due to numerical reasons. Otherwise, when not converged runs are excluded, the statistic can be severely biased and report too optimistic values as simulations are more likely to converge for perturbation towards physically "easier" parameter values.

Non-convergence of a flowsheet can be either due to numerical issues when solving the flowsheet or physical infeasibility of the designed process subject to the assumed conditions. While latter provides the engineer with useful information about the design, former is unfavorable as numerical problems conceal the question of physical feasibility. In general, it is a very time-consuming, manual task to converge a process flowsheet. In this case, this path is infeasible as there will be of the order of hundreds or thousands simulations for one specific process design. To boost convergence—key to providing reliable information about the process—we introduce a heuristic to prioritize competing restrictions using insights from daily operations. According to this ranking the most important priority is to produce substances of desired purity, the second priority is to achieve the desired quantity of product, and lastly, the process is to remain cost-effective. Hence, the optimization problem is set up to maximize product subject to purity constraints and includes an upper limit on the produced amounts.

Furthermore, bounds of the optimization problem can be selectively relaxed when the original problem does not converge. If the problem converges in a second attempt with the relaxed bounds, this solution is used as the starting point for a homotopy method. After each successful convergence of the optimization problem, the relaxed bounds will be consecutively retightened until the original constraints are reached. If convergence is lost as the bounds are retightened, this is noted in the detailed output to the user. Apparently, one is attempting to operate the process just beyond its physical limit, which is useful information when studying the feasibility limits of a proposed process design.

Dependability can be viewed as a condensed measure of the ability of the process design to cope with varying conditions. In combination with heuristics to improve convergence of the flowsheet, it can be determined automatically. Certainly, more converged runs increase the statistical foundation for the measure of dependability. Furthermore, more detailed information about feasibility limits also gives the process engineer a better understanding where physical limitations of the process are.

3. Case study on uncertainty in process design for ethylene oxide synthesis

To illustrate the above-described method, we study a process to synthesize ethylene oxide that was set up based on available information in the literature (Onken and Behr, 1996; Rebsdat and Mayer, 2005). The process is modeled and solved in Aspen Plus using the equation-oriented (EO) simulation mode.

3.1. Setting up simulation for dependability analysis

First, an optimal process design—in terms of minimized overall cost—is found using short-cut equipment sizing and costing methods to provide the baseline for the analysis

of effects of uncertain parameters (Wiesel and Polt, 2006). The dimensions resulting from this optimization are subsequently adjusted to account for empirical overdesign factors. Then, physical limitations of used equipment are included in the model using simplified physical measures, such as F-factors for distillation columns, similar to those used in the sizing procedures. The optimization problem to solve for the dependability analysis is to maximize product quantity subject to unbounded feed streams, product purity restrictions and the above designed plant with its modeled physical limitations.
In this case study, uncertainty in conversion rates used to model the reaction kinetics as well as fluctuating raw material purities are regarded. It is assumed that all parameters follow a triangular distribution with specified lower and upper limits as well as a mode. In order to obtain trustworthy results, 1000 simulations are performed sampling the uncertain parameters according to the above-described QMC technique. Overall, the simulations require computation time of the order of several hours.

3.2. Results of dependability analysis

In the studied case, more than 95% of all simulations converged; 18% required lower product quantity to converge whereas more than 77% complied with the process specifications. Therefore, the dependability of the design is 0.77. Figure 1 shows that these different result areas are rather cleanly separated from each other except for few numerical artifacts.

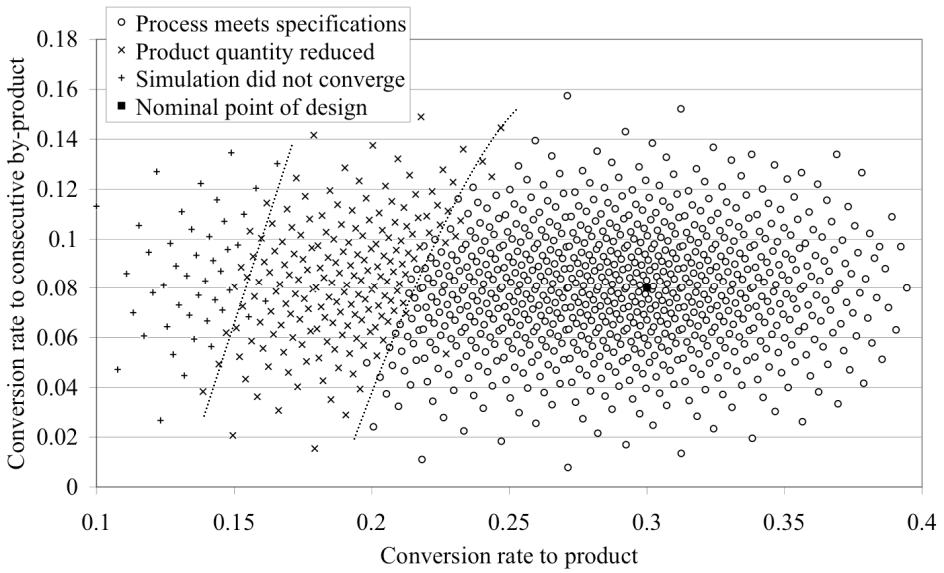


Figure 1: Plot of simulation results show the influence of two uncertain parameters on process dependability. Bounds of the dependable range are approximated and included for illustration.

In the reported case, the plant is capable of producing on average close to 98% of nameplate product quantity when neglecting samples for which no information is available due to not converged simulation runs.

3.3. Discussion

As fig.1 shows, the process is well behaved in proximity to its nominal point of design. However, as conversion rates deviate in unfavorable directions, the limits of the process are tested and exceeded as denoted by decrease of product. When the studied parameters vary even more, numerical issues inhibit further studies, as convergence is lost.

Although the dependability of the designed process appears to be fairly low, the design is capable of nearly achieving nameplate product quantity. The discrepancy between low dependability and high actual product quantity can be explained by the only gradual decline of product output as the process is operated at points beyond its name plate production bound. Here, the importance of as many as possible converged simulations becomes obvious since conclusions can only be drawn from results of converged results. Therefore, if one were to not account for and not to include this region between fulfilling specifications and no convergence in numerical results for averaged process state variables, the actual performance of the plant would be greatly underestimated resulting in the seemingly need for greater equipment overdesign.

4. Conclusion

An approach to support process design when some parameters are not known exactly has been proposed in this work. Important aspects of the methodology are interconnection of process simulator and simple user interface to encapsulate and automate repeated Aspen Plus runs. Equally important are the proposed heuristic means to differentiate between numerical problems and physical limitations. These lead to an increase in converged runs which allow this approach to generate non-trivial insight into complex chemical processes and to provide a quantitative measure of sensitivity of the designed process to uncertain parameters. Hence, the proposed method can provide help to process engineers when designing processes to meet uncertain process conditions on one hand while on the other hand limiting unnecessary process overdesign.
The method of dependability analysis can be applied to study effects of manifold uncertain influences to a process. Though in the presented case study, uncertainties have been restricted to process parameters, the method can be easily extended to incorporate uncertain cost coefficients, e.g. to study the impact of fluctuating raw material costs. Likewise, the proposed priority ranking need not be applicable for all processes; it can be adjusted to meet different priorities without further implications to the methodology.

References

Diwekar, U.M., Kalagnanam, J.R., 1997, Efficient sampling technique for optimization under uncertainty, AIChE J 43(2):440-447.
Freeman, R.A., Gaddy J. L., 1975, Quantitative overdesign of chemical processes, AIChE J 21(3):436-440.
Grossmann, I.E., Sargent, R.W.H., 1978, Optimum design of chemical plants with uncertain parameters, AIChE J 24(6):1021-1028.
Grossmann, I.E., Haleman, K.P. Swaney, R.E, 1983, Optimization strategies for flexible process design, Comput Chem Eng 7(4):439-462.
Malik, R.K., Hughes, R.R, 1979, Optimal design of flexible chemical processes, Comput Chem Eng 3:473-485.
Morokoff, W.J., Caflisch, R.E., 1994, Quasi-random sequences and their discrepancies, SIAM J Sci Comput 15(6):1251-1279.
Onken, U., Behr, A, 1996, Chemische Prozeßkunde - Lehrbuch der technischen Chemie (Vol. 3), Stuttgart: Georg Thieme Verlag, 35-38.
Rebsdat, S., Mayer, D., 2005, Ethylene Oxide. In: Ullmann's Encyclodaedia of Industrial Chemistry (7th ed.), Ullmann, F. et al., Weinheim: Wiley-VCH Verlag.
Sprow, F.B, 1967, Evaluation of research expenditures using triangular distribution functions and monte carlo methods, Ind Eng Chem 59(7):35-38.
Wiesel, A., Polt, A., 2006, Conceptual steady state process design in times of value based management, Proceedings of PSE 9 and ESCAPE 16, 799-804

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Cutting Edge Biodiesel Production by Catalytic Reactive Absorption

Anton A. Kiss

AkzoNobel Research, Development & Innovation, Process & Product Technology, Arnhem, The Netherlands, tony.kiss@akzonobel.com

Abstract

This study takes previous work on separative reactors for biodiesel production to a new level by proposing a novel technology based on reactive absorption. This is a significant step forward as reactive absorption offers important advantages in addition to typical benefits of reactive distillation: reduced capital investment and operating costs due to the absence of the reboiler and condenser, and lower temperature profile in the column to avoid thermal degradation of the products. Computer aided process engineering tools such as AspenTech Aspen Plus are used for the process design and simulation of a plant producing 10 ktpy biodiesel from waste oil with high free fatty acids content.

Keywords: reactive absorption, green catalysts, sustainable fuels

1. Introduction

Biodiesel is an alternative fuel with similar propersties as petroleum diesel. Therefore, it can be used alone, or blended with conventional petrodiesel, in unmodified diesel-engine vehicles. Tipically, biodiesel is produced from green sources such as vegetable oils, animal fat or even waste cooking-oil from the food industry.[1,2] Therefore, biodiesel is a *green fuel* that has many advantages over conventional petrodiesel: it is safe, renewable, non-toxic and biodegradable, it contains insignificant amounts of sulfur and its increased lubricity extends the life of diesel engines. In addition, it has a high cetane number (above 60 compared to 40 for petrodiesel), a high flash point (>130°C) and it emits ~70% fewer hydrocarbons, ~80% less CO_2, and ~50% less particles.[1-3]

As a non-petroleum-based diesel fuel, biodiesel consists of short chain alkyl esters of fatty acids, currently produced by acid/base-catalyzed (trans-)esterification, followed by several neutralization and purification steps. Nevertheless, all the conventional methods suffer from problems associated with the use of homogeneous acid or base catalysts, leading to serious economical and environmental consequences, especially considering the recent growth of the overall biodiesel production scale. The increasing worldwide interest in biodiesel is illustrated by the exponential increase of the production, mostly in Western Europe, USA, and Asia (Figure 1).

This study presents a novel biodiesel technology based on reactive absorption that offers significant advantages compared to conventional methods: simple and robust process, high conversion and selectivity, elimination of conventional catalyst-related operations, no thermal degradation of products, no waste streams, as well as reduced capital investment and operating costs. The process design proposed in this work is based on experimental results and rigorous simulations performed using AspenTech AspenONE Engineering Suite as computer aided process engineering tool.

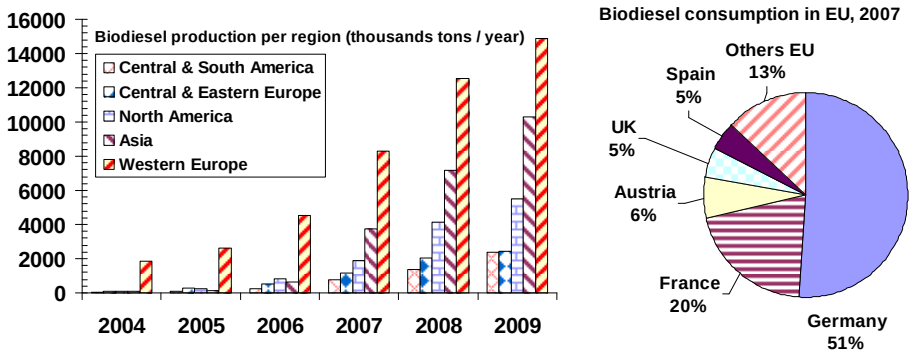


Figure 1. Biodiesel production per region (left), and biodiesel consumption in EU (right).

2. Problem statement

There are three basic methods to produce fatty esters from oils/fats: 1) base catalyzed trans-esterification, 2) acid catalyzed esterification, and 3) enzymatic catalysis.[1-3] The first method is the most frequently used. However, due to the escalating costs of fatty raw materials, the current trend is to use less expensive alternatives such as animal fat, waste cooking oil from catering premises, or waste vegetable oil (wvo).

The problem with waste oils is the very high content of free fatty acids (FFA) that lead to soap formation in a conventional base catalyzed process. Therefore, in order to avoid production loss and soap associated problems, the FFA's must be completely converted first to fatty esters by esterification. Moreover, the conventional biodiesel processes employ liquid catalysts, such as H_2SO_4, NaOH or methoxides.[3] The problem is that homogeneous catalysts require neutralization, washing, separation, recovery, and waste disposal operations with severe economical and environmental penalties.

To solve this problem we propose a novel fatty esterification process based on reactive absorption (RA) using solid acids as catalysts[4,5] and therefore eliminating the additional separation steps and the salt waste streams, thus simplifying the downstream processing. Table 1 presents an overview of the available solid acid catalysts for biodiesel production.[4-6] In this work we selected the metal oxides as acid catalysts, but the ion-exchange resins are also suitable due to the moderate temperatures used in the process.

Note that previous literature studies on separative reactors for biodiesel production are solely based on reactive distillation (RD).[6-9] The novel RA technology proposed in this work offers additional advantages compared to RD, as for example lower temperature profile in the reactive separation column to avoid the thermal degradation of the fatty esters products. As a consequence the process becomes much simpler and more robust, meaning reduced capital investment and operating costs due to the absence of a reboiler (no product vapors return to the column) and condenser (no reflux of water by-product). Moreover, the integrated RA unit is able to shift the chemical equilibrium to completion by continuous removal of products instead of using an excess of reactant. This novel approach based on RA is particularly suitable for treating waste oil or animal fat, including tri-glycerides with up to 100% free fatty acids (FFA) – this is in fact the worst case scenario considered in the work described here.

Table 1. Advantages and disadvantages of the acid catalysts tested for fatty acids esterification.

Catalyst type	Benefits	Drawbacks
Ion-exchange resins (Nafion, Amberlyst)	Very high activity Easy regeneration	Low thermal stability Possible leeching
TPA ($H_3PW_{12}O_{40}$)	Very high activity	Soluble in water
TPA-Cs ($Cs_{2.5}H_{0.5}PW_{12}O_{40}$)	Super acid sites	Low activity per weight
Zeolites (H-ZSM-5, Y and Beta)	Controlable acidity and hydrophobicity	Small pore size Low activity
Sulfated metal oxides (zirconia, titania, tin oxide)	High activity Thermally stable	Deactivates in water, but not in organic phase
Niobic oxide (Nb_2O_5)	Water tolerant	Average activity

3. Simulation methods

Based on the mixture of any particular starting oil/fat used in the biodiesel production process, there will be an associated blend of fatty acid esters in the final biodiesel product. Therefore, the simulation can be performed using one of the available simulation methods illustrated in Table 2: rigorous, shortcut or hybrid method. Note that each method has its key benefits but also specific drawbacks and the requirements can differ considerably. Although favored, the rigorous method is virtually not feasible in practice due to the amount of input data required. On the other hand, the shortcut method provides merely low-fidelity models with very limited practical applications. Therefore, for practical reasons, the hybrid approach gives the best results. Note that in the following simulations the experimentally determined kinetic parameters[5,6] were used but the fatty components were lumped into one fatty acid/ester compound, according to the following chemical reaction:

$$R\text{-}COOH + CH_3OH \leftrightarrow R\text{-}COO\text{-}CH_3 + H_2O$$

Table 2. Simulation methods for biodiesel production: requirements, benefits and drawbacks.

	Rigorous method	Shortcut method	Hybrid method
Requirements	Properties for all species. VLL data and BIP's for all pairs of components. Kinetic parameters for all reactions possible.	Properties for single fatty acid/ester/tri-glyceride. VLL data for the system ester/glycerol/alcohol. Asumed conversion (no kinetic parameters).	Single or reduced list of fatty acid/ester/TG. Short list of VLL data and BIP's for components. Reduced list of kinetic parameters, few reactions.
Benefits	Easy optimization of reaction and separation. High fidelity model. Usable for many plants. Easy comparison for various feedstocks.	Simple model. Fast simulations. Easy-to-build mass and energy balance. No data needed for all species present.	Optimization possible for reaction and separation. Certain ability to compare various feedstocks. Better model fidelity. Fast simulations for RTO.
Drawbacks	Slow simulations and convergence problems. Expensive measurements. Limited RTO and model based control usage.	No comparison possible for various feedstocks. Low-fidelity model. Less ability to use RTO.	More effort to build component list and get kinetic parameters. More work to find VLL data and regress BIP's.

4. Results and discussion

The physical properties of the components present in this process were determined experimentally,[5-7] or estimated using state-of-the-art contribution methods such as UNIFAC – Dortmund modified. Vapor pressure is one of the most important properties with a critical effect in modeling reactive separations. At ambient pressure the boiling points of fatty esters are relatively high (over 300 °C). High purity products are possible in a RD setup, but the high temperature in the reboiler – caused by the high boiling points, is in conflict with the thermo-stability of the biodiesel product. This problem can be avoided by working at lower pressure or allowing methanol in the bottom product.[5-7]

By using reactive adsorption, the drawbacks of reactive distillation can be completely avoided and biodiesel can be produced at moderate temperatures and ambient pressure. Moreover, the water by-product is not refluxed in the RA column hence the detrimental effect of water on the equilibrium reaction and the catalyst is completely avoided.

Figure 2 presents the flowsheet of a biodiesel production process based on a reactive absorption column (RAC) as the key unit. The process was rigorously simulated and optimized using AspenTech AspenONE. The production rate considered for the biodiesel plant designed in this work is 10 ktpy fatty esters. Note that the kinetic parameters used in the simulations were previosly reported in the open literature.[5,6]

The RAC is operated in the temperature range of 135–160 °C, at ambient pressure. Out of the 15 stages of the integrated unit, the reactive zone is located in the middle of the column (10 stages). The fatty acid is pre-heated then fed as hot liquid in the top of the column while a stoichiometric amount of alcohol is injected as vapor into the bottom of the column, thus creating a counter-current flow regime over the middle reactive zone.

Water by-product is removed as top vapor, then condensed and separated in a decanter from which only the fatty acids are recycled back to the column while water can be re-used. The fatty esters are delivered as high-purity bottom product of the RAC. The hot product is flashed first to remove the remaining methanol, then it is cooled down and stored. Further heat-integration is possible but this is beyond the scope of this study.

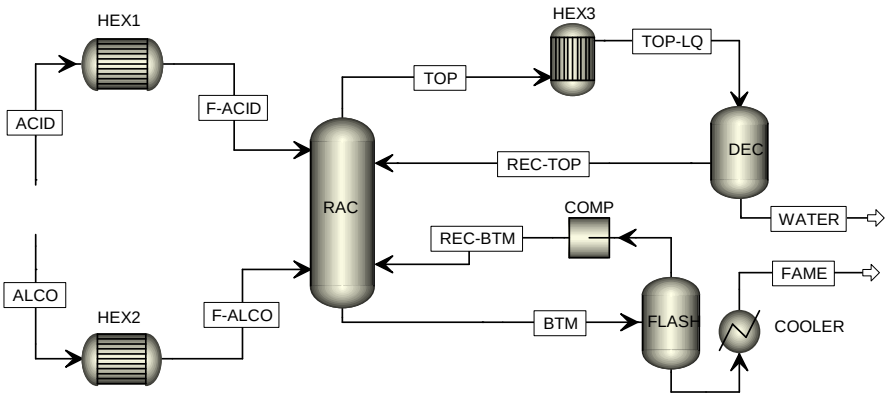


Figure 2. Flowsheet of biodiesel production by catalytic reactive absorption.

Table 3. Mass balance of a 10 ktpa biodiesel production process based on reactive-absorption.

	F-ACID	F-ALCO	BTM	REC-BTM	REC-TOP	TOP	WATER	FAME
Temperature C	160	65.4	136.2	246.2	51.8	162.1	51.8	30
Pressure bar	1.05	1.05	1.03	1.216	1	1	1	0.203
Vapor Frac	0	1	0	1	0	1	0	0
Mass Flow kg/hr	1166.755	188.306	1261.295	11.295	9.369	114.43	105.061	1250
Volume Flow cum/hr	1.492	157.565	1.417	8.949	0.011	213.042	0.109	1.259
Enthalpy Gcal/hr	-0.94	-0.279	-0.904	-0.013	-0.009	-0.337	-0.395	-0.957
Mass Flow kg/hr								
METHANOL	0	188.306	9.083	7.51	0.002	0.129	0.127	1.573
ACID	1166.755	0	0.112	0	9.13	9.147	0.017	0.112
WATER	0	0	0	0	0.237	105.154	104.917	0
ESTER-M	0	0	1252.1	3.785	0	0	0	1248.315
Mass Frac								
METHANOL	0	1	0.007	0.665	0	0.001	0.001	0.001
ACID	1	0	0	0	0.974	0.08	0	0
WATER	0	0	0	0	0.025	0.919	0.999	0
ESTER-M	0	0	0.993	0.335	0	0	0	0.999

The mass and energy balance is given in Table 3. High purity products are possible, the purity specifications exceeding 99.9%wt for the final biodiesel product (FAME stream). Water by-product is also recovered at high purity, hence this stream could be reused as industrial water on the same site. The energy usage is less than 135 kW/ton biodiesel. Note also that the total amount of the recycle streams (REC-TOP and REC-BTM) is not significant, representing only ~1.5% of the total biodiesel production rate.

Figure 3 shows the molar composition profiles in both liquid and vapor phase. The concentration of fatty acid and water increases from the bottom to the top of the column, while the concentration of fatty ester and methanol increases from the top to bottom. Therefore, in the top of the reactive absorption column there are vapors of water and liquid fatty acids, while in the bottom there are vapors of the methanol feed and liquid fatty esters product (biodiesel).

Figure 4 shows the temperature and reaction rate profiles in RAC, as well as the mass flowrates along the column. The temperature difference between the top section and the bottom part of the RA column is relatively low (~30 °C) compared to a RD column. However, the reaction rate profile is similar to the one in a RD column, exhibiting a maximum in the middle and thus keeping the best of both reactive-separation designs.

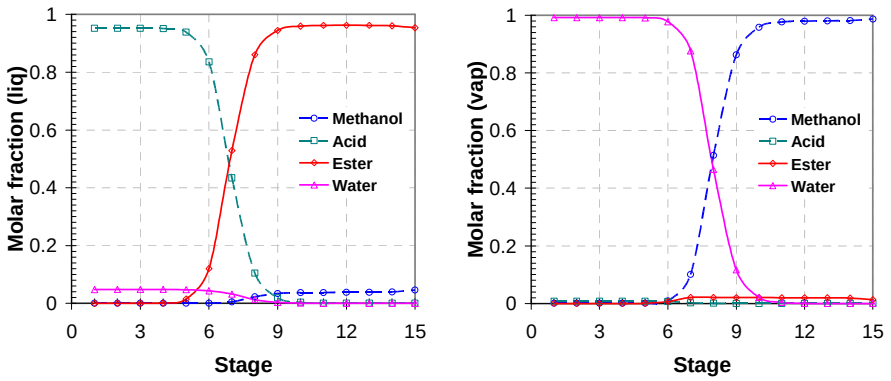


Figure 3. Profiles in RAC: liquid composition (left), vapor composition (right).

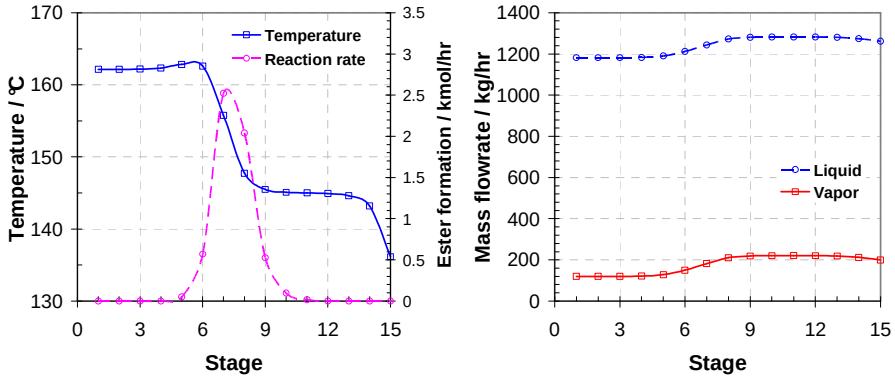


Figure 4. Temperature and reaction rate profiles (left), mass flowrates profiles in RAC (right).

5. Conclusions

This study uses computer aided engineering tools such as AspenTech Aspen Plus for the development of an innovative biodiesel process based on reactive absorption. This novel process radically improves the biodiesel production and dramatically reduces the number of downstream processing steps. The key benefits of this unique process are:

1. Simple and robust process with no catalyst-related waste salt streams, no soap formation, and sulfur-free biodiesel as solid catalysts do not leach into the product.
2. Elimination of conventional catalyst-related operations such as: handling of toxic chemicals and corrosive solutions, catalyst neutralization, separation and disposal of waste salts, waste water treatment, recovery and recycling of excess alcohol.
3. Reduced equipment costs, with up to ~60% savings on the total capital investment.
4. Low operating costs due to the integrated design with no reboiler or condenser.
5. Effective use of the reactor volume leading to significantly high unit productivity.
6. Efficient use of raw materials: no thermal degradation, down to stoichiometric reactants ratio, high conversion as equilibrium is shifted towards completion.
7. Multifunctional plant suitable for a large range of alcohols and fatty raw materials with very high FFA content, such as frying oils, animal tallow, waste vegetable oil.

References

1. M. Balat, H. Balata, 2008, Energy Conversion and Management, 49, 2727.
2. M.G. Kulkarni, A. K. Dalai, 2006, Industrial & Engineering Chemistry Research, 45, 2901.
3. K. Narasimharao, A. Lee, K. Wilson, 2007, J. Biobased Materials & Bioenergy, 1, 19.
4. T. Okuhara, 2002, Chemical Reviews, 102, 3641.
5. A. Kiss, A. C. Dimian, G. Rothenberg, 2006, Advanced Synthesis & Catalysis, 348, 75.
6. A. Kiss, G. Rothenberg, A. C. Dimian, F. Omota, 2006, Topics in Catalysis, 40, 141.
7. A. Kiss, A. C. Dimian, G. Rothenberg, 2008, Energy & Fuels, 22, 598.
8. S. Steinigeweg, J. Gmehling, 2003, Ind. Eng. Chem. Res., 42, 3612.
9. C. Dimian, F. Omota, A. Bliek, 2004, Chem. Eng. & Proc., 43, 411.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Study of silver and oxide hybrids of catalysts formaldehyde production by using NLP model

Anita Kovac Kralj

Faculty of Chemistry and Chemical Engineering, University of Maribor, Slovenia, EU

Abstract

Hybrid catalysts are engineered to contain elements of two or more catalysts. Hybrid catalysts are becoming increasingly important advantages for obtaining novel and desirable catalyst activities and properties. The nature of a hybrid catalyst can be changed with additional characters. The disadvantages can be divided and the advantages multiplied.

This study describes a novel technology for the synthesis of formaldehyde production by using hybrid catalysts and the optimization of them. The nature of the hybrid catalyst can be changed by the addition of positive properties. Both, 100 % oxide and 100 % silver formaldehyde production processes can be optimized during hybrid processes. The application of a nonlinear programming, (NLP) mathematical method can be used for optimizing the hybrids of catalysts consisting of percentages of oxide and silver catalysts. The NLP model contains equations for parametric optimization.

The optimized silver process modification was based on 80 % silver catalyst material and the addition of 20 % oxide catalyst. A 20 % change of catalyst can enhance the available heat and product production by 3.8 %.

Keywords: hybrid catalysts, formaldehyde, NLP model

1. Introduction

Formaldehyde is a chemical compound with the formula H_2CO. It is the simplest aldehyde - an organic compound containing a terminal carbonyl group: it consists of exactly one carbonyl. It was first synthesized by the Russian chemist Aleksander Butlerov, but was conclusively identified by August Wilhelm von Hofmann (Reuss et al., 2001). Formaldehyde exists in several forms aside from H_2CO: the cyclic trimer trioxane and the polymer paraformaldehyde. It exists in water as the hydrate $H_2C(OH)_2$.

Two important routes are prominent during industrial formaldehyde production (Fischer, 1992 and Twigg, 1996). The first synthesis route is performed over an electrolytic silver catalyst under lean-air conditions, and was been commercialized at the beginning of the 20^{th} century. The second route, industrialized since the late 1959's, is by oxidation in excess air over a ferric molybdate catalyst. Other routes have not as yet been commercialized, e. g. the sodium catalyzed dehydrogenation of methanol to anhydrous formaldehyde and the partial oxidation of methane over silica catalysts. In 2000, about 55% of the industrial production capacity in Western Europe was based on the silver-catalyzed route (Burridge, 1983), thus it plays an important role. Its advantages are the relatively low investment cost and stable production.

Formaldehyde production can be used in either of the two different kinds of process plants: silver process and oxide process.

This study presents a novel technology for the synthesis of formaldehyde production by using hybrid catalysts.

2. Hybrid catalyst processes

Hybrid catalyst processes can change the characteristic of basic process. It can enlarge the available heat or/and amount of product production.

The silver and oxide processes were analysed one by one, by using an Aspen Plus simulator (Aspen Plus User, 2004) and then compared graphically because of the simple comparisons between the processes (with the same feed flow-rate). The catalyst effect can be analysed by using Aspen Plus and the results are presented with linear equations (1-4). It is a very simple method for an effective approach when using different catalysts. The parameter X_{oxid} can denote the value of oxide catalyst amount. The pure silver process is presented as a 0 % oxide process amount with $X_{oxid} = 0$. The oxide process is presented as a 100 % oxide process amount with $X_{oxid} = 100$. There were different hybrid processes using different catalysts amount between the silver and oxide processes. The parameter X_{oxid} was calculated in fraction (not in percent) in all equations.

The right thermodynamic properties for oxygen and hydrocarbon mixtures during process simulation and optimisation can be estimated by the Wilson property model. In the case-studied formaldehyde plant, the flow rates of oxygen (O_2) and hydrocarbon mixtures can be simulated using an Aspen Plus simulator, but accurate estimates regarding the thermodynamic properties of oxygen and hydrocarbon mixtures need to be made. In this case study, the model is in relatively good agreement with experimental measurements in the existing formaldehyde production plant. So we can use these results for a possible theoretically - identifiable topic area, as key to hybrid catalyst themes. The most important parameters such as: inlet oxygen's mass flows, inlet methanol mass flows, conversion in the reactor, temperature in reactor and available heat, can be defined between the oxide and silver processes.

The processes used different mass flows of oxygen as raw materials. The silver process is using 1130 kg/h of oxygen. The oxide process uses 4189 kg/h of oxygen. Hybrid processes use different catalytic amounts and use different inlet oxygen mass flows between 1130 to 4189 kg/h. The mathematical approximation equation (1) of oxygen's mass flows (q_{ox}) for hybrid was:

$$q_{ox} = 3059\, X_{oxide} + 1130 \quad (1)$$

Conversion in the reactor was different; the oxide process had higher conversion. The oxide process had 99 % conversion and the silver process 80 % conversion. The hybrid processes, using different catalytic amounts, had different conversions between 80 % and 99 %. The mathematical approximation equation (2) of the conversion in reactor (C_{on}) for hybrid was:

$$C_{on} = 19\, X_{oxide} + 80 \quad (2)$$

The temperature in reactor was higher during the silver process. The silver process had an approximate temperature of about 600 ^{o}C. The oxide process had an approximate temperature of about 300 ^{o}C. The mathematical approximation equation (3) of temperature in the reactor (T_r) for hybrid was:

$$T_r = -\,300\, X_{oxide} + 600 \quad (3)$$

The oxide process had the higher available heat. The available heat can be estimated approximately by using linear equation (4). Almost exact estimation can be established by using the Aspen Plus simulator. The mathematical approximation equation (4) of total available heat (Φ_{av}) for hybrid was:

$$\Phi_{av} = 7238\, X_{oxide} + 2149 \quad (4)$$

We have to choose the best one from among the hybrid processes as the criteria energy analysis was used and for maximal additional product production. The energy criteria used was the thermodynamic method better known as "pinch analysis". Pinch analysis helps us to better analyse heat integration, available heat during a process using graphical representation as a grand composite curve (see Figures 1 and 2). It is a powerful technique for the synthesis of a utility system and the results from this method can be used to postulate a superstructure (Zhu et al., 1995). Pinch analysis is guiding heat and power integration using an extended grand-composite curve (Tjoe et al, 1986). Thermodynamic analysis does not guarantee a global optimum solution because it cannot be used simultaneously with the material balance, but it quickly proposes good ideas for the heat and power integration of complex processes (Yee et al., 1990).

The hybrid simulations were simulated by using an Aspen Plus simulator. The different hybrid processes were tested. Both the 100 % oxide and 100 % silver formaldehyde production processes can be modified during different hybrid processes. Comparisons between the grand composite curves of the oxide and silver hybrid processes were decisive for choosing the kind of hybrid. Figure 1 presents comparisons between the pure oxide process and hybrids with 80 % oxide and 20 % silver catalysts, and 60 % oxide and 40 % silver catalysts. The energy efficiency of oxide hybrid can be analysed using the grand composite curve (Fig. 1). The highest available heat areas have 80 %, then 100 %, and the last 60 % oxide catalyst.

Figure 2 presents comparisons between the pure silver process and hybrids with 80 % silver and 20 % oxide catalysts, and 90 % silver and 10 % oxide catalysts. The energy efficiency of silver hybrids can be analysed using the grand composite curve (Fig. 2), too. The highest available heat areas have 80 %, then 90 %, and the last 100 % oxide catalyst.

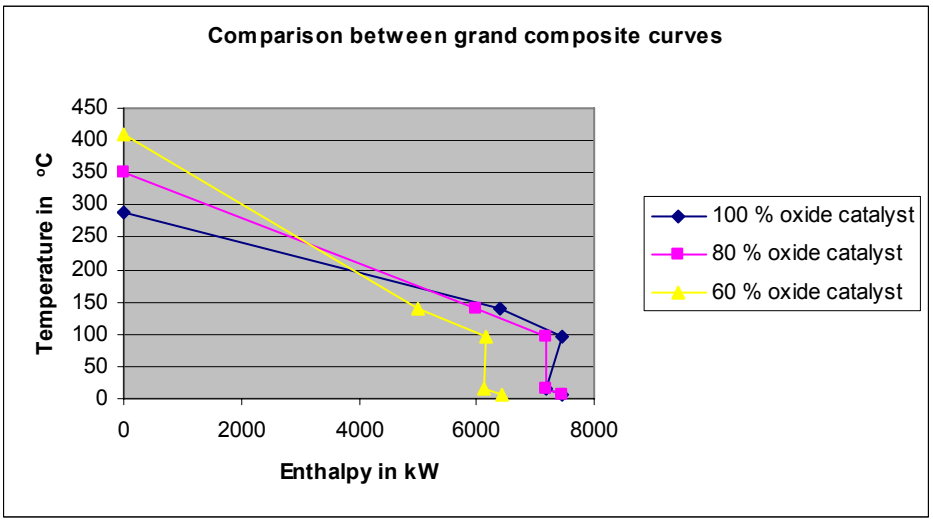


Figure 1: Comparison of grand composite curve of oxide hybrids processes.

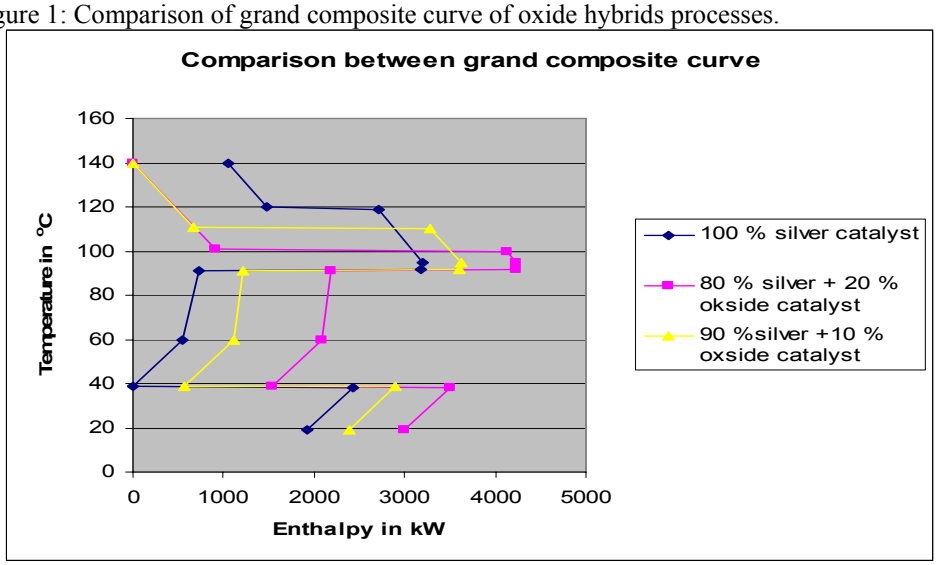


Figure 2: Comparison of grand composite curve of silver hybrids processes.

Energy efficiency (E_{ff}) as a fraction for both hybrids can be denoted by using mathematical polynomial of the 4th order (Eqs. 5−6):

$$E_{ff} = a\, X_{oxid}^{**4} + b\, X_{oxid}^{**3} + c\, X_{oxid}^{**2} + d\, X_{oxid} + e \quad (5)$$

$$E_{ff} = -33.405\, X_{oxid}^{**4} + 67.694\, X_{oxid}^{**3} - 42.015 X_{oxid}^{**2} + 8.0105\, X_{oxid} + 0.5137 \quad (6)$$

The mathematical approximation equation (4) of total available heat (Φ_{av}) for hybrid was not linearly constant; it was multiplied by Energy efficiency (E_{ff}) in the objective function (Eq. 7).

3. The NLP model

The hybrid process parameters are optimized using a nonlinear programming (NLP) model (Brooke et. al, 1992). Optimization could increase annual profit.

The parameters in the hybrid model were simultaneously optimized using the GAMS/MINOS software. This NLP can be solved using a large-scale reduced gradient method (e. g. MINOS). This model is non-convex, it does not guarantee a global optimization solution, but it quickly gives good results for non-trivial, complex processes. This NLP model used the simple linear equations from 1 to 4 and polynomial of the 4th order. The parameter X_{oxid} can denote the value of oxide catalyst mass. The pure silver process is presented as a 0 % oxide mass (X_{oxid} = 0). The pure oxide process is presented as a 100 % oxide mass (X_{oxid} = 1). The parameter X_{oxid} can variable between 0 and 1; it can be presented by the composition of a hybrid catalyst.

The objective function (Eq. 7) of the NLP model is to maximize the annual profit in EUR/a, V; it includes the income from formalin production ($C_f\, C_{on}\, q_M$), income from available heat ($C_{heat}\, \Phi_{av}\, E_{ff}$), and annual cost of raw material for oxygen ($C_{ox}\ q_{ox}$) and cost of oxide catalyst ($C_{ox\text{-}cat}\ X_{oxid}$; Table 1). The silver catalyst is cheaper than the oxide. The cost for average equipment in both process, C_{eq} was 4 MEUR/a.

Maximal additional annual profit is defined by eq. 7 :

$$V_{max} = C_f\, C_{on}/100\; q_M\, t_{op} + C_{heat}\; \Phi_{av}\, E_{ff} - C_{ox}\; q_{ox}\, t_{op} - C_{ox\text{-}cat}\, X_{oxid} - C_{eq} \quad (7)$$

Where q_M is the mass flow rate of methanol, C_{on} is the conversion of methanol in formaldehyde. Time of operation (t_{op}) is 8000 h/a.

Table 1: Cost items for the example processes.

Price of formaldehyde, C_f/(EUR/t): 150
Price of heat, C_{heat}**/(EUR/kW · a): 60
Cost of oxygen and including the equipment, C_{ox}/(EUR/t): 50
Additional cost of oxide catalyst, $C_{ox\text{-}cat}$/(EUR/a): 200 000
Cost of average equipment in both process C_{eq}/(MEUR/a): 4

** Swaney, 1992

The objective function of the optimized hybrid process was 0.85 MEUR/a. The parameter X_{oxid} was 0.2, which was denoted by 20 % composition of oxide and 80 % silver catalysts. Conversion was 83.8 % (0.838 in fraction). The available heat including efficiency was 2050 kW. It needed 1.74 t/h of oxygen.

The optimised silver process modification with 80 % silver and 20 % oxide catalysts - the silver hybrid process (Fig. 3) was very close to the silver process, having little difference and was simulated by using an Aspen Plus simulator. The inlet stream of air increased from 1130 kg/h of oxygen to 1742 (Eq. 1). The temperature in the reactor was lower by about 60 °C (Eq. 3) but with an additional reactor heat enthalpy of 1983 kW. This heat can be used for heating the column reboiler, so during this process steam is not needed for heating. The constant need for heat can be covered in the process. The silver hybrid process had the lower catalytic costs (McKetta, 1985). The production of formaldehyde was higher by 3.8 % for the static inlet value of methanol.

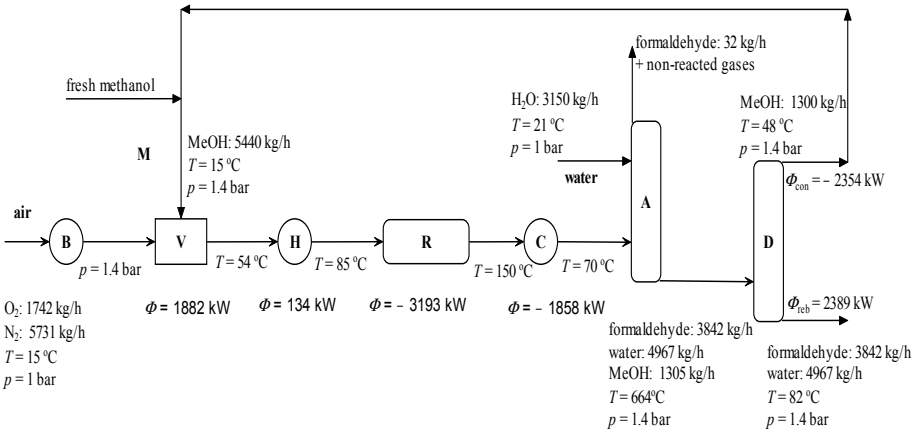


Figure 3: Optimized silver process modification.

4. Conclusion

Hybrid catalysts are used more frequently in enzyme catalysis than in chemical catalysis. The nature of a hybrid catalyst can be changed with additional positive

characteristics. Opportunities for positively increasing these characteristics depend very much on the hybrid catalysts used. An important characteristic of a hybrid catalyst is its effect on process efficiency. The basic characteristic of the silver and oxide formaldehyde processes can be maintained. The hybrid silver and oxide processes can be affected on the quality and/or quantity of available heat and product production. The criteria as to which hybrid can be used, depends on the quality and/or quantity of available heat and maximal product production. Pinch analysis is used for better energy analysis, which helps us to choose the better hybrid by using a graphical representation of the grand composite curve.

The results and methodology for possible theoretically identifiable hybrid catalytic themes can be analysed by using Aspen Plus and presented by linear equations. The most important parameters: such as inlet oxygen's mass flows, inlet methanol mass flows, conversion in the reactor, temperature in the reactor, and available heat, can be defined between hybrid, and the oxide and silver processes.

We can keep the basic characteristics of silver and oxide formaldehyde processes. The hybrid process has an affect on the quality and/or quantity of available heat and product.

Hybrid catalysts were optimized using the NLP algorithm. The NLP model contains equations of parametric optimization. In the optimization model a number of possible alternatives for hybrid catalysts can be included and computer based tools can find the best new catalyst composition. This procedure does not guarantee a global cost optimum, but it does lead to good, perhaps near-optimum, retrofit.

The optimized silver process modification was based on 80 % silver catalyst material and the addition 20 % oxide catalyst. The 20 % changing catalyst can enlarge the available heat and product production by 3.8 %.

References

Aspen Technology, 2004, ASPEN PLUS User Manual Release 12. Aspen Technology, USA.

A . Brooke, D. Kendrick and A. Meeraus, 1992, GAMS: *A User's Guide*, Palo Alto, Scientific Press.

E. Burridge, 1983, Eur. Chem. News 75, , p. 16.

B.A.J. Fischer, 1992, Formaldehyde, Encyclopaedia of Chemical Technology, 4th Edition, Vol. 11, Wiley, New York , p. 929.

J. J. McKetta, W. A. Cunningham, 1985, Encyclopedia of chemical processing and design, 23, Marcel Dekker, Inc., New York, p. 350-371.

G. Reuss, W. Disteldorf, A.O. Gamer, A. Hilt, 2001, Formaldehyde, Ullmann's Encyclopedia of Industrial Chemistry, 6th Edition, Vol. A11, VCH, Weinheim, p. 619.

J. M. Smith, 1981, Chemical engineering kinetics, McGraw-Hill Book , p. 350–355.

R . Swaney. 1989, Thermal integration of processes with heat engines and heat pumps. AIChE Journal 35/6, 1010.

T.N. Tjoe and B. Linnhoff, 1986, Using pinch tehnology for process retrofit. Chem. Engng 28, p. 47-60.

M.V. Twigg, 1996, Catalyst Handbook, 2nd Edition, Manson Publishing, London, pp. 490.

X.X. Zhu, B.K. O'Neill, J.R. Roach and R.M. Wood, 1995, A method for automated heat exchanger networks synthesis using block decomposition and nonlinear Optimization, Trans IChemE 73/A**,** p. 919 – 930.

T.F Yee. and I.E. Grossmann, 1990, Simultaneous optimization models for heat integration II. Heat exchanger network synthesis, Comput. chem. Engng 14/10, p. 1165 – 1184.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Design and Optimization of Thermally Coupled Distillation Sequences for Purification of Bioethanol

Carlo Edgar Torres-Ortega[a], Juan Gabriel Segovia-Hernández[a], Salvador Hernández[a], Héctor Hernández[a], Adrián Bonilla-Petriciolet[b], Rafael Maya-Yescas[c]

[a]*Departamento de Ingeniería Química, División de Ciencias Naturales y Exactas, Universidad de Guanajuato, Campus Guanajuato, Noria Alta S/N, Guanajuato, Gto, 36050, México*
[b]*Instituto Tecnológico de Aguascalientes, Departamento de Ingeniería Química, Av. Adolfo López Mateos #1801 Ote. Fracc. Bonagens, Aguascalientes, Ags., 2025, México.*
[c]*Facultad de Ingeniería Química, Universidad Michoacana de San Nicolás de Hidalgo, Ciudad Universitaria, Col. Felicitas del Río, 58060, Morelia, Michoacán, México.*

Abstract

An important problem in the bioethanol production process is the purification of ethanol from a dilute solution, i.e., approximately 10% ethanol in water. The key factor in the purification process is the formation of the ethanol-water binary homogeneous azeotrope, and an additional process is required to obtain high purity ethanol that can be used in motor vehicles. This study examines the design and optimization of three extractive distillation options (two with thermal coupling) for the purification of a representative mixture of ethanol and water. These extractive arrangements can produce ethanol as distillate with the required purity and energy savings, reduction in CO_2 emissions, high thermodynamic efficiencies and good control properties.

Keywords: extractive distillation, thermal coupling, energy savings.

1. Introduction

A large portion of the world economy has been based on processes highly dependent on petroleum; as a consequence, two major problems have arisen: draining of oil reserves, and increasing gas emissions, associated with global warming. Researchers in many areas are working on solutions to mitigate these problems. One option that offers a partial solution is to intensify production and use of biofuels, including biodiesel, bioethanol, biomass, etc. Of these, bioethanol is currently being used in unmodified combustion engines in a mixture of up to 20%. The use of this mixture is important because it enables improved oxidation of hydrocarbons and, as a result, reduction in both hydrocarbon and carbon dioxide emissions (Quintero et al., 2008).

Currently, most bioethanol is produced from sugar cane, with corn in second place. However, a great deal of current research is focused on industrial production from lignocellulosic material such as agricultural and forest residues. In the production process, four main steps can be identified: Treatment of raw material to obtain the cellulosic mass, saccharification to obtain sugars from the cellulosic mass, fermentation of the sugars, and recovery of ethanol. Independently from the raw material and/or the

process used, the product obtained from the fermentation step is a dilute solution of ethanol in water, from which ethanol is separated and purified to the desired concentration. In addition to research efforts in the saccharification and fermentation process, the separation step should also be viewed as a challenge, and studied accordingly, because of the energy it requires. Assuming that the fermentation process produces a dilute solution of ethanol in water (10% in moles of ethanol) requiring treatment in order to obtain high-purity ethanol for mixing with gasoline (Cardona and Sanchez, 2007; Wingren et al., 2008), the production of high-purity ethanol using distillation requires significant quantities of energy and mass separation agents such as ethylene glycol, NaCl, KI or $CaCl_2$.

The key factor in the purification process is the formation of the ethanol-water binary homogeneous azeotrope. This azeotrope is formed with 96 mass percent of ethanol in water, and an additional process is required to obtain high-purity ethanol to be used in motor vehicles. Several methods can be used: the first is dehydration using a salt, e.g., NaCl, KI, CaCl2, while a second method involves the use of ethylene glycol as an entrainer. Other option is the use of membranes for dehydration by pervaporation (Jiang et al., 2008). This study examines the design and optimization of three extractive distillation options for the purification of a representative mixture of ethanol and water using ethylene glycol (Figure 1), where two of the options use thermally coupled extractive distillation sequences (thermally coupled extractive distillation sequence, Figure 1b; and extractive Petlyuk column, Figure 1c). The study is complemented by a thermodynamic efficiency analysis, CO_2 emissions calculations and determination of control properties at zero frequency.

As previously mentioned, bioethanol can be used in a mixture with gasoline in current combustion engines (Goldemberg, 2007; Goldemberg and Guardabassi, 2009), and it is the most important biofuel produced at present, with worldwide output of about 32 million tons in 2006, of which 90% was from only two countries, Brazil and the USA (Dimian and Bildea, 2008). If bioethanol were to replace gasoline, emissions of greenhouse gases would be reduced by more than 85%, taking the complete fuel cycle into consideration (Bergeron, 1996).

2. Design of Purification Options

Figure 1 shows the three alternatives for the purification of the mixture. A dilute feed of 100 lb-mol/h ethanol in water (10% in moles of ethanol in water) as saturated liquid at 1 atm is introduced into a conventional distillation column that removes the binary homogeneous azeotrope as distillate. This study focuses on the separation stage for ethanol with a high mass fraction (0.995). The bottoms product of the first distillation column is almost pure water. This conventional distillation column is needed in all three distillation options. The first alternative (Figura 1a) uses an extractive conventional distillation column with ethylene glycol as entrainer; the distillate of the column is ethanol with a mass fraction of 0.995. The bottoms product of the extractive distillation column is a ternary mixture of ethanol, water and ethylene glycol. This mixture is fed to a third distillation column in order to recover the entrainer as bottoms product, where the distillate is a mixture of ethanol and water that can be returned to the first distillation column, where the azeotrope is formed. The second and third options (figures 1b and 1c), in the extractive stage of the separation, use a distillation column coupled to a side rectifier and a Petlyuk column, respectively.

Design and optimization methods for thermally coupled extractive distillation are reported in Gutiérrez-Guerra et al., 2009, and briefly described below: To overcome the complexity of the simultaneous solution of the tray arrangement and energy consumption in a formal optimization algorithm, we decoupled the design problem into two stages: (i) tray configuration; (ii) optimal energy consumption. The first stage of the approach begins with the development of preliminary designs for the complex systems, starting from the design aspects of conventional distillation columns. After the tray arrangement for the complex extractive sequence has been obtained, an optimization procedure is used to minimize the heat duty supplied to the reboilers of the arrangement, taking into account the constraints imposed by the required purity of the product streams. Next, the degrees of freedom that remain after design specifications and tray arrangement are used to determine operating conditions that provide minimum energy consumption. The optimization strategy can be summarized as follows: (a) A base design for the complex scheme is obtained. (b) Values for the extractant stream stage and interconnecting flows are assumed. (c) A rigorous model for the simulation of the complex scheme with the proposed tray arrangement is solved (in this study, Aspen Plus™ was used for this purpose). If product compositions are obtained, then the design is kept; otherwise, appropriate adjustments must be made. (d) One value of interconnecting flow is changed, returning to step (c) until a local minimum in energy consumption for the assumed value of the extractant stream stage is identified. (e) The value of the extractant stream stage is modified, returning to step (c) until the energy consumption is minimum (optimization criterion). This result implies that an optimum value has been identified for the design of the complex scheme.

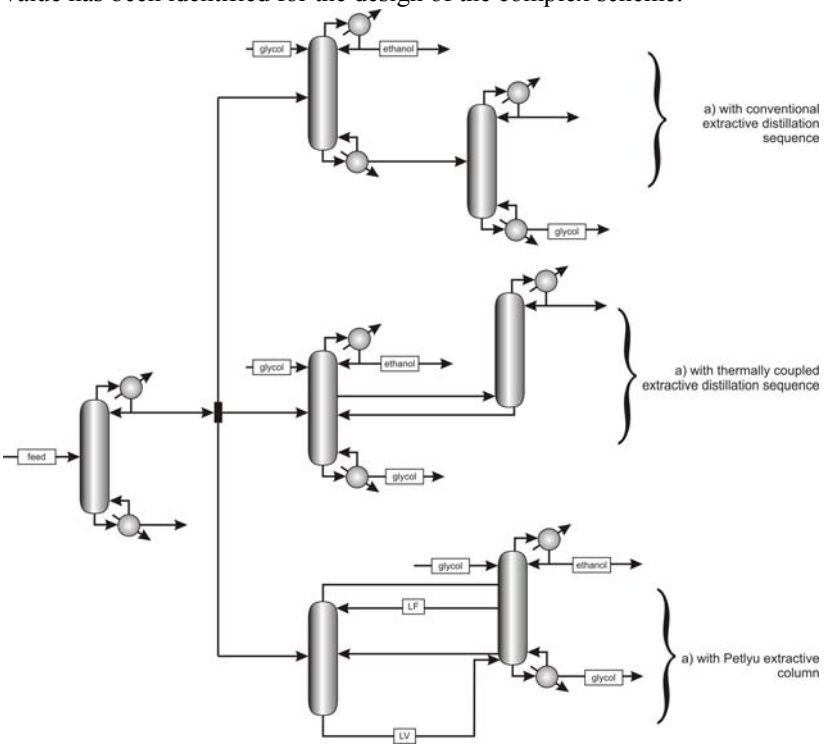


Figure 1. Purification options studied.

Thermodynamic properties of the liquid and vapor phases were calculated using the NRTL model and the Redlich-Kwong equation, respectively. Proper modeling of thermodynamic properties is very important since, in the first stage of the separation procedure, a binary distillation column is required to obtain the binary ethanol-water azeotrope. In this regard, the NRTL model can predict the formation of the binary azeotrope. In all distillation columns, the design pressure for each separation was chosen to ensure the use of cooling water in the condensers. A pressure drop of 10 psi was assumed for each distillation column of the separation sequences.

3. Results

The resulting designs and their performance with respect to optimum solvent to feed ratio (E/F), optimum energy consumption (Q), CO_2 emissions (Gadalla et al., 2005), thermodynamic efficiency (η) and control properties [minimum singular value (σ_*) and condition number (γ) at zero frequency] are discussed. As mentioned previously, the first option (Figure 1a) uses an extractive conventional distillation column with ethylene glycol as entrainer. This option has total energy consumption (including three columns) of 487.63 kW (Table 1) and it is considered a basis for the analysis.

Table 1. Results for the cases of study.

Option	E/F Optimum	Total Energy Consumption (kW)	CO_2 Emissions (Ton/h)	η (%)	σ_*	γ
With conventional extractive sequence	1.5	487.63	0.105	8.13	0.0296	852
With thermally coupled extractive distillation sequence	1.4	497.32	0.111	7.97	0.125	57.5
With extractive Petlyuk column	1.5	362.30	0.0798	12.5	0.987	45.8

Typical optimization curves for the thermally coupled extractive distillation sequence are shown in Figure 2, where the optimal value for the extractant stage and interconnecting flowrate can be identified in order to guarantee minimum energy consumption. The optimization curves show an interesting effect of the search variables on energy consumption. The design is sensitive, in terms of its energy consumption, to changes in interconnecting flowrates and extractant stage. An implication of this observation has to do with operational considerations (the presence of recycle streams can contribute to good dynamic behavior). For all alternatives, the results can be summarized as follows: (i) the energy savings achieved by the option using a Petlyuk

extractive column are 25.7% over the option using a conventional extractive arrangement; (ii) the least favorable option is the sequence using a thermally coupled extractive distillation sequence, as it shows the highest energy consumption and lowest thermodynamic efficiency; (iii) the second law efficiency (η) of the option using a Petlyuk extractive column is higher than that of the corresponding conventional extractive distillation option; (iv) the reduction in global CO_2 emissions, in the arrangement with the Petlyuk structure, is considerable: 24% over the option using the conventional extractive sequence. The inefficiency of conventional sequences (associated with increased CO_2 emissions) has been reported as a consequence of remixing (Hernández, et al., 2006).

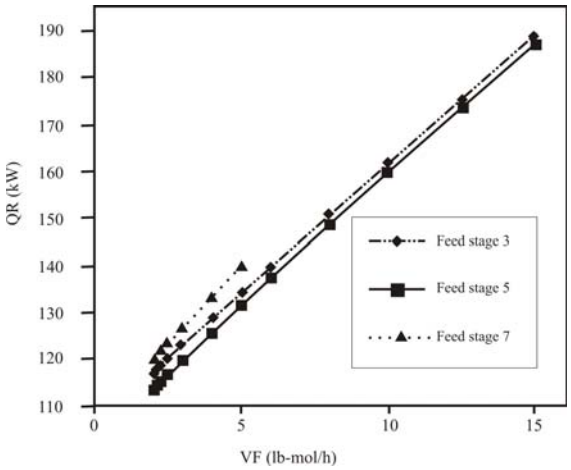


Figure 2. Search for minimum energy consumption in the thermally coupled extractive distillation sequence design.

Therefore, proper optimization of the Petlyuk extractive sequence avoids such a remixing problem. The methodology used generates designs where the effect of the remixing is eliminated. In addition, it is important to analyze the composition profiles in the extractive distillation systems. The option using the conventional extractive column can produce ethanol with a high mass fraction (0.995), with the bottoms product composed of ethylene glycol, water and ethanol. The options using complex extractive columns produce ethanol with the same mass fraction, but it is important to highlight that these complex columns separate the entrainer as the bottoms product and the side rectifier or side stream removes a mixture of ethanol and water.

A final test was carried out on the control properties of the options studied, using the singular value decomposition (SVD) technique at zero frequency. The singular values of the transfer function matrix of a process at zero frequency are the squared roots of the eigenvalues of the product of the transfer function matrix by its transpose and conjugated. Two parameters of interest are the minimum singular value and the ratio of maximum to minimum singular values, or condition number. The systems with higher σ_* values and lower γ values are expected to show the best dynamic performance under feedback control. Table 1 shows the SVD results for the three options analyzed. The arrangement using an extractive Petlyuk column presents the highest minimum singular value; therefore, it can be expected that this complex system will exhibit better closed-loop dynamic behavior than the option using an extractive conventional sequence. Condition number results show that the sequence with the Petlyuk structure offers the

best value. As a result, it can be expected that the complex distillation system is better conditioned to the effect of disturbances than the conventional arrangement.
According to the results, the Petlyuk extractive arrangement is the best scheme option for separation of the binary homogeneous azeotrope of ethanol and water as regards energy consumption, CO_2 emissions, thermodynamic efficiency and control properties. In addition, this complex distillation sequence can reduce capital costs since it can be implemented in a single distillation column using a dividing wall.

4. Conclusions

In this work, the separation of a typical mixture of ethanol and water from a fermentation process was studied, considering an extractive conventional distillation column and two thermally coupled extractive distillation sequences using ethylene glycol as entrainer. The results show that the option using an extractive Petlyuk column can produce energy savings of 25.7% over the option with an extractive conventional distillation arrangement. Moreover, the structure with the Petlyuk scheme shows the best results with regard to CO_2 emissions (associated with energy consumption), thermodynamic efficiency and control properties. In general, the results are important because they indicate that the use of complex extractive distillation systems in the bioethanol purification process is feasible as regards economic, thermodynamic and control aspects.

5. Acknowledgements

We acknowledge the financial support provided by CONACyT, CONCyTEG, Universidad de Guanajuato and Instituto Tecnológico de Aguascalientes (Mexico).

6. References

P. Bergeron, 2006, Environmental Impacts of Bioethanol, Handbook on Bioethanol: Production and Utilization, 89-92.

C. A. Cardona, O. J. Sanchez, 2007, Fuel ethanol production: Process design trends and integration opportunities, Bioresource Technology, 98, 2415-2457.

A.C. Dimian, C.S. Bildea, 2008, Bioethanol Manufacturing, Chemical Process Design, 429-459.

M.A. Gadalla, Z.Olujic, P.J. Jansens, M. Jobson, R. Smith, 2005, Reducing CO2 Emissions and Energy Consumption of Heat-Integrated Distillation Systems, Environ. Sci. Technol., 39, 6860 – 6870.

S. Hernández, J. G. Segovia-Hernández, V. Rico-Ramirez, 2006, Thermodynamically equivalent distillation schemes to the Petlyuk column for ternary mixtures, Energy, 31, 1840-1847.

J. Goldemberg, 2007, Ethanol for a Sustainable Energy Future, Science, 315, 808-810.

J. Goldemberg, P. Guardabassi, 2009, Are Biofuels a Feasible Option?, Energy Policy, 37, 1, 10-14.

R. Gutiérrez-Guerra, J.G. Segovia-Hernández, S. Hernández, 2009, Reducing Energy Consumption and CO_2 Emissions in Extractive Distillation, Chemical Engineering Research and Design, in press.

L.Y. Jiang, T.S. Chung, R. Rajagopalan, 2008, Dehydration of alcohols by pervaporation through polyimide Matrimid® asymmetric hollowfibers with various modifications, Chem. Eng. Sci, 63, 204-216.

J.A.Quintero, M.I. Montoya, O.J. Sánchez, O. H. Giraldo, O.H., C.A.Cardona, 2008, Fuel ethanol production from sugarcane and corn: Comparative analysis for a Colombian case, Energy, 33, 385-399.

A.Wingren, M.Galbe, G. Zacchi, 2008, Energy considerations for a SSF-based softwood ethanol plant, Bioresource Technology, 99, 2121-2131.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

The Influence of Tangent Pinch Points on the Performance of Batch Rectifications: Development of a Conceptual Model for Ternary Mixtures

Karina A.Torres,[a] José Espinosa,[b]
[a,b] *INGAR – CONICET - UNL, Avellaneda 3657, S3002 GJC Santa Fe, Argentina*

Abstract

This contribution explores the influence of tangent pinch points on the performance of batch distillations of highly non ideal ternary mixtures and its incorporation into a conceptual modeling framework under the assumption of a batch rectifier with infinite number of stages.
Two algorithms are developed to determine the maximum feasible distillate composition on the line of preferred separation and its corresponding limiting reflux ratio. The first one is based on a region elimination method (REM) and relates the occurrence of a saddle-node bifurcation to the appearance of a maximum in the curve reflux vs. molar liquid fraction of any of the components of the mixture. This method takes into account the bifurcation analysis of reversible distillation profiles. The second method is analytical and consists in solving an equation system that represents two facts: the liquid compositions of the bifurcation point, together with its vapor in equilibrium and its distillate must be aligned, and a turning point occurs. The resolution of the system is simplified because of the fact that in batch distillation of ternary mixtures operating at reflux below or equal to the limiting one, the bifurcation point coincides with the instantaneous still composition in the reboiler. Then, the dependence of feasible distillate mole fractions on reflux ratios above the limiting one is calculated by solving a non linear equation system, which incorporates the tangency condition and provides the unknown compositions of tangent pinches. A numerical method, based on the improved memory method, is implemented here to solve this non linear system. The conceptual model replaces the one based on the assumption of the invariance of a controlling saddle pinch, only valid for ideal systems and non ideal mixtures without tangent pinch points, and emphasizes the variation of the composition of the tangential pinch points for reflux above the limiting one. Results obtained from the conceptual model for instantaneous column performance are in excellent agreement with rigorous simulation in Hysys and highlight the effectiveness of the proposed algorithm. Two highly non ideal ternary mixtures are studied along the contribution: the systems methanol/2-propanol/water and acetone/chloroform/benzene.

Keywords: batch distillation, saddle-node bifurcation, reversible profiles, tangent pinches.

1. Quantitative determination of the trajectory of feasible distillate compositions in nonideal ternary mixtures

1.1. Quantitative determination of the maximum feasible distillate composition on the line of preferred separation

Calculation of the maximum feasible distillate composition on the equilibrium vector or preferred separation line for a given instantaneous still composition is the key ingredient

of a conceptual model which attempts to incorporate the influence of tangent pinches on the performance of batch distillations. Duessel[1] was the first to introduce the idea of solving the eigenvalue problem of the Jacobian of the equilibrium function in the instantaneous still composition x_B to obtain a linearization of the adiabatic profile in the neighborhood of the still composition and hence, an estimation of the instantaneous minimum reflux ratio. Note that for an instantaneous still composition x_B, the feed to the column is a vapor stream whose composition $y^*_{x_B}$ is in phase equilibrium with the composition of the liquid in the still and the liquid leaving the lower column end x_N is located on the linear approximation of the adiabatic profile departing from x_B. The mass balance envelope used in the conceptual model is shown in Figure 1(a).

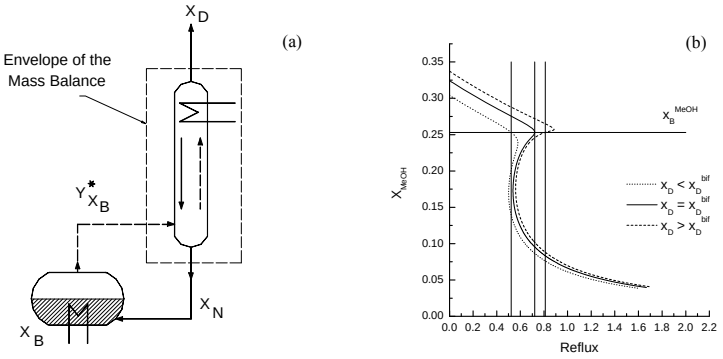


Figure 1. (a) Envelope of the mass balance used in the conceptual modeling of batch distillations; (b) System methanol/2-propanol/water. Pinch point curve x_{MeOH} *vs.* r for three distillate compositions on the preferred line with direction $\left(y^*_{x_B} - x_B\right)$.

On the other hand, Espinosa and Salomone[2] adapted Duessel's approach to handle the operation at constant reflux ratio and they identified three different situations: i) for reflux ratios between zero and $R^* = \left(x^*_D - y^*_{x_B}\right) \Big/ \left(y^*_{x_B} - x_B\right)$, the distillate compositions are aligned with the equilibrium vector, starting from $y^*_{x_B}$ ($r = 0$) up to the point x^*_D (where the composition of the heaviest component becomes zero). A pinch at the rectifier lower end, with identical composition to that of the mixture in the still, controls the separation; ii) for reflux ratios between R^* and $R^{**} = \left(x^{**}_D - y^*_{x_B}\right) \Big/ \left(y^*_{x_B} - x_{N,\max}\right)$, the distillate compositions are located in the binary edge corresponding to the lighter components between x^*_D and x^{**}_D. $x_{N,\max}$ is the intersection of vectors $\left(x^{**}_D - y^*_{x_B}\right)$ and the linearization of the adiabatic profile $x_N = x_B + \lambda v$. There is now a binary saddle pinch x^{II}_P, with invariant composition; iii) for $r \geq R^{**}$ the distillate composition equals x^{**}_D, the composition of the more volatile species with a pinch at the top of the column controlling the separation.

In this contribution, we will focus on the incorporation of tangent pinches into the conceptual modeling of batch distillations, which affects the maximum distillate composition achievable on the preferred line. Similarly to ideal mixtures, the lower pinch controls the separation for values of the reflux ratio within an interval with

bounds $\left[0; r^{bif}\right]$. At $r = r^{bif}$, the lever arm rule determines the maximum feasible distillate composition x_D^{bif} for which the corresponding pinch point curve shows a saddle-node bifurcation at the instantaneous still composition x_B. The operation at reflux ratios above the bifurcation value is controlled by tangent pinches of unknown compositions.

1.1.2. Quantitative determination of profiles of reversible distillation

In order to calculate the reversible profiles or pinch point curves, the following system of linear equations must be solved:

$$\frac{d\vec{f}_c(T,\vec{x})}{dT} = \vec{0} = \frac{\partial \vec{f}_c}{\partial T} + \left(\frac{\partial f_{c,i}}{\partial x_j}\right)^{\Sigma}\left(\frac{d\vec{x}}{dT}\right)^{\Sigma} \qquad (1)$$

where both $\frac{\partial \vec{f}_c}{\partial T}$ y $\left(\frac{\partial f_{c,i}}{\partial x_j}\right)^{\Sigma}$ can be analytically calculated. The supraindex Σ in Eq. (1) indicates that the mole fraction summation over all components is accounted for. The numerical integration of $\left(dx/dT\right)^{\Sigma}$ in Eq. (1), (beginning from an initial point such as a pure component or azeotrope) allows the calculation of the entire reversible profile for a given distillate $\mathbf{p}$.

1.1.3. Region Elimination Method (REM)

As it was mentioned before, the maximum feasible distillate mole fraction on the line of preferred separation is characterized by a saddle-node bifurcation of the corresponding reversible profile at the instantaneous still composition, and it is related to the appearance of a maximum in the curve r vs. the composition of any of the components of the mixture. This fact can be appreciated in Figure 1(b) and it is used to determine the limiting distillate composition x_D^* and its corresponding reflux ratio $r^* = r^{bif}$. Given a distillate $\mathbf{p}$ belonging to the line of preferred separation, the main idea of this method is to first calculate two points x_B^+ and x_B^- of the reversible profile around x_B by numerically integrating the equation system (1) for given values of the temperature increments ΔT_1 and $\Delta T_2 = -\Delta T_1$. Then, reflux ratios are obtained from the lever arm rule for the three compositions and finally, values of the derivatives $\Delta r_1 / \Delta x_k^1 = \left(r_{x_B} - r_{x_B}^-\right) / \left(x_{B,1} - x_{B,1}^-\right)$ and $\Delta r_2 / \Delta x_k^2 = \left(r_{x_B} - r_{x_B}^+\right) / \left(x_{B,1} - x_{B,1}^+\right)$ on both sides of the still composition are estimated. If the signs of the derivatives differ each other, a saddle-node bifurcation at instantaneous still composition is found with the corresponding distillate mole fraction as the maximum feasible one. Otherwise, another distillate composition must be picked up (according to the bisection method) until convergence. The initial search interval is $\left[y_{x_B}^*, z_D\right]$, where z_D is the intersection of the equilibrium vector with either the binary axis corresponding to the more volatile components (zeotropic systems) or the pinch distillation boundary, PDB (azeotropic mixtures with unstable distillation boundaries).

1.1.4. Analytical Method

The following equation system, adapted from the work of Fidkowski et al.[3], represents the tangent pinch situation in a given column section (e.g., for the rectifying section):

$$\begin{cases} -y(x_k) + \dfrac{r}{r+1} x_k + \dfrac{1}{r+1} p_k = 0 & (2) \\ \det\left(Y - \dfrac{r}{r+1} I \right)\Big|_{x_k} = 0 & (3) \end{cases}$$

where $Y = \left[\dfrac{\partial y_i}{\partial x_j} \right]^{\Sigma}_{\hat{x}_{i \neq j}}$ is the Jacobian matrix of the equilibrium function and x_k, $y(x_k)$ and p_k are the *k -th* components of the bifurcation point, its vapor in equilibrium and the distillate, respectively. This method uses the fact that in batch distillation of ternary mixtures operating at a reflux ratio $r \leq r^{bif}$, the bifurcation point coincides with the instantaneous still composition x_B. Therefore, a simple way to obtain the maximum feasible composition in the line of preferred separation consists of the following steps: i) calculate the elements of the Jacobian matrix of the equilibrium function from their analytical expressions[4,5], ii) solve the eigenvalue problem given by equation (3) at x_B to obtain the bifurcation value of the reflux ratio $r = r^{bif} = r^*$, iii) calculate the maximum feasible distillate composition $p = x_D^* = x_D^{bif}$ through the lever arm rule, equation (2). It is interesting to note that two values of the reflux ratio are obtained from solving equation (3). For systems where the reversible profile either resembles a pinch point curve without a tangent pinch point or a reversible profile showing a non-controlling tangent pinch, the reflux ratio values will take either negatives values or positive ones, giving rise in the last case to distillate compositions outside the composition simplex.

1.2. Quantitative determination of the curve of feasible distillate compositions for reflux ratios greater than r^{bif}

For reflux ratios above r^{bif}, the instantaneous operation is controlled by unknown values of tangent pinches x_p, lower end liquid compositions $x_N \neq x_B$, and distillate mole fractions x_D, that can be calculated by solving the equation system 4-8:

$$-y^*_{x_B} + \frac{r}{r+1} x_N + \frac{x_D}{r+1} = 0 \qquad (4)$$

$$-y^*_{x_p} + \frac{r}{r+1} x_p + \frac{x_D}{r+1} = 0 \qquad (5)$$

$$y^*_{x_p} = \frac{\gamma_i(x_p, T) p_i^0(T)}{p} x_p \qquad (6)$$

$$x_N = x_B + \lambda v, \qquad x_N \in [x_B, x_{N,\max}] \qquad (7)$$

$$\det\left(Y - \frac{r}{r+1} I \right)\Big|_{x_p} = 0 \qquad (8)$$

It's easy to check that consistency and solution unicity of this system are guaranteed. The numerical resolution of 4-8 is based on the improved memory method developed by Shacham[6] for the solution of a non-linear equation and is presented below:

Step 1. Given a value for the reflux ratio r, select values for $x_{N,1}$ and $x_{N,2}$, belonging

to the eigenvector within the interval $\left[x_B, x_{N,\max}\right]$. Therefore, equation (7) is obeyed.

Step 2. Solve equation (4) for $x_{D,1}$ and $x_{D,2}$ from the estimated values $x_{N,1}$ and $x_{N,2}$.

Step 3. For each distillate composition from Step 2 calculate, by integrating equation (1), the pinch point curve and Δr (the difference between the reflux ratio corresponding to the first i-th point of the pinch curve for which $r(i+1)-r(i)<0$, and the operating reflux ratio r). The selected point corresponds to a saddle-node bifurcation of the reversible profile (see Figure 1(b)) and equations (5) and (6) are satisfied (the error function Δr replaces the tangency condition, equation (8)). However, the bifurcation reflux differs in general from the operation reflux r.

Step 4. If $\Delta r(x_{N,1}).\Delta r(x_{N,2})<0$, initialize the improved memory method by taking $\left[x_{N,1}, x_{N,2}\right]$ as the initial search interval until determining the composition at the lower end column x_N for which the error function $\Delta r \cong 0$. Otherwise, go to step 1, select a new value for x_N and repeat steps 2 and 3 until an appropriate search interval is found.

From the analysis of the algorithm above it is clear that the method searches for a distillate composition whose corresponding pinch point at the saddle-node bifurcation [x_p , $y^*_{x_p}$, r^{bif} and T_p] has a reflux ratio r^{bif} such that $r^{bif}=r$ (Figure 2(a)). In other words, it searches for the root of the error function $\Delta r=\Delta r(x_N)$.

The results from the conceptual model were compared to rigorous simulations from Hysys for different values of $r \geq r^{bif}$ and all of them were in excellent agreement.

1.3. Conclusions

Figure 2(b) shows the curve of feasible distillate mole fractions corresponding to an instantaneous still composition x_B for the mixture acetone/chloroform/benzene at some reflux ratios $0 \leq r \leq r^{**}$. The products are situated on the "preferred line" for reflux ratios $r \leq r^{bif}$ and along a curved trajectory if $r > r^{bif}$. The system 4-8 replaces the model based on the assumption of the invariance of a controlling saddle pinch, valid for ideal systems and nonideal mixtures without tangent pinch points. Figure 2(c) emphasizes the variation of the composition of the tangential pinch points for $r \geq r^{bif}$.

2. Two ternary mixtures as examples of application

Two non ideal ternary mixtures are studied: the systems MeOH/IPA/W (unstable distillation boundary) and A/C/B (stable distillation boundary). The Wilson-Ideal model is used to calculate the liquid-vapor equilibrium. All the necessary parameters (also the analytical derivatives) are obtained from codes implemented in Matlab.

		$\mathbf{x_D^*}$	$\mathbf{r^{bif}}$
Methanol/2-propanol/Water	REM	[0.454636 0.358644 0.186720]	0.746
	Analytical Method	[0.454392 0.358727 0.186881]	0.744
Acetone/Chloroform/ Benzene	REM	[0.875536 0.095881 0.028583]	2.89
	Analytical Method	[0.875721 0.095777 0.028502]	2.87

Table 1**.** Maximum feasible distillate composition on the preferred line and limiting reflux ratio from the region elimination algorithm and the analytical method for two nonideal mixtures.

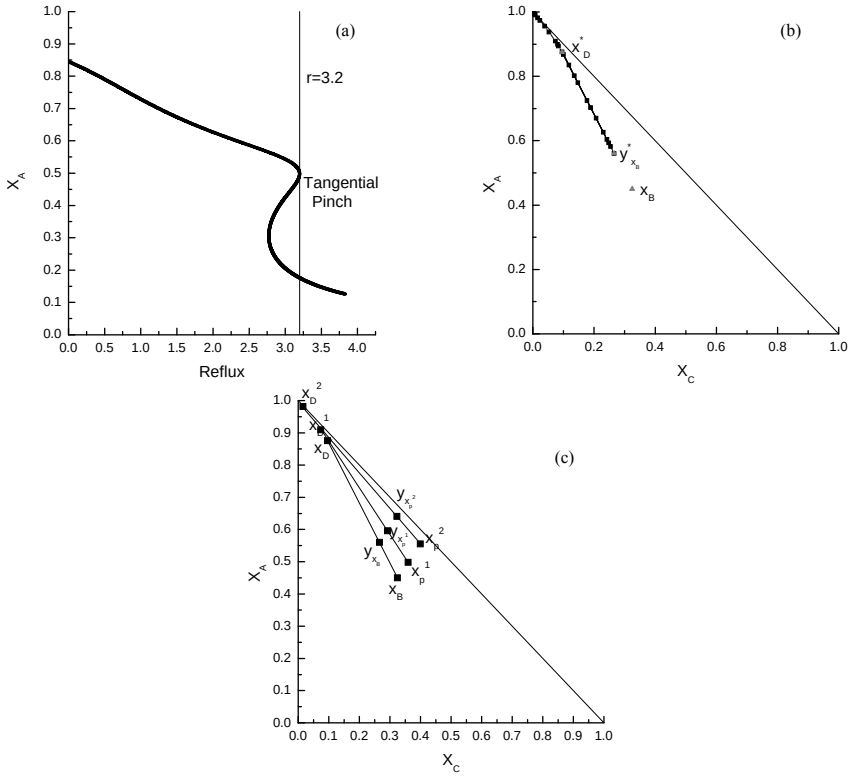


Figure 2. System ACB. (a) Pinch point curve for x_D =[0.9044 0.0774 0.0182] at $r = 3.2$; (b) Trajectory of feasible distillate composition corresponding to a still composition x_B for reflux ratios between 0 and r^{**}; (c) Three different tangential pinches for reflux ratios above the bifurcation one.

References

1. R. Duessel, 1996, Zerlegung azeotroper Gemische durch Batch-Rektifikation, *Ph. D. Thesis*, Technical University of Munich, Germany.
2. J. Espinosa; E. Salomone, 1999, Minimum Reflux for Batch Distillations of Ideal and Nonideal Mixtures at Constant Reflux, *Ind. Eng. Chem. Res.*, 38, 2732-2746.
3. Z. T. Fidkowski; M. F. Malone; M. F. Doherty, 1991, Nonideal Multicomponent Distillation: Use of Bifurcation Theory for Design. *AIChE J.*, *37* (12), 1761–1779.
4. P. Poellmann; E. Blass, 1994, Best Products of Homogeneous Azeotropic Distillations, GAS separation & purification, 8 (4), 194-228.
5. R. Taylor; H. A. Kooijman, 1991, Composition Derivatives of Activity Coefficient Models, *Chem. Eng. Commun*, 102, 87-106.
6. M. Shacham, 1989, An Improved Memory Method for the Solution of a Nonlinear Equation, *Chem. Eng. Sci.*, 44 (7), 1495-1501.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Development of Novel Property Prediction Methods for Phase Equilibrium Calculations Based on the Molecular Structure

Mordechai Shacham[a], Georgi St. Cholakov[b], Roumiana P. Stateva[c] and Neima Brauner[d]

[a]*Ben-Gurion University of the Negev, Chem. Eng. Dept, Beer-Sheva 84105, Israel*
[b]*Univ. Chem. Technol. Metallurgy, Dept. Org. Synthesis, Fuels, Sofia 1756, Bulgaria*
[c]*Bulgarian Academy of Sciences, Inst. Chem. Engineering, Sofia 1113, Bulgaria*
[c]*Tel-Aviv University, Faculty of Engineering, Tel-Aviv 69978, Israel*

Abstract

A new method for applying Quantitative Structure Property Relationship type methods for prediction of properties needed for phase equilibrium calculation is presented. The use of the new method is demonstrated by predicting vapor pressure for members of the 1-alkene homologous series. It is shown that the new method enables the prediction of vapor pressure within experimental uncertainty, depending on the level of similarity between the predictive compounds and the target compound.

Keywords: Property prediction, Phase equilibria, Molecular structure, QSPR.

1. Introduction

Phase equilibria calculations are of key importance for process and product design. In many cases the required data for these calculations are not available and have to be estimated. Up-to now, most published methods that relate molecular structure and properties have been devoted to constant properties of pure compounds, which are necessary but not sufficient for phase equilibria calculations of mixtures at different compositions, temperatures and pressures (Poling et al, 2001).

We have investigated the use of a new method, which is a combination of the Short-Cut Quantitative Structure-Structure-Property Relationship (SC-QS2PR, Cholakov et al, 2007) and the Targeted Quantitative Structure Relationship (TQSPR, Brauner et al. 2006) for prediction of the phase equilibrium related properties.

In the next section the properties needed for phase equilibrium computation are reviewed and categorized. In the 3rd section the proposed combination of the SC-QS2PR and TQSPR methods is described. The use of this new method is demonstrated by predicting vapor pressure (saturation temperature) for members of the 1-alkene homologous series.

2. Properties Needed for Phase Equilibrium Calculations

Vapor – liquid equilibrium ratio K (K-value) is commonly used for phase equilibrium computations. At low and moderate pressures the $\gamma - \Phi$ model is typically used. Using this model the K – value for the i-th component is given by:

$$K_i = \frac{\gamma_{iL} f^{0}{}_{iL}}{\phi_{i_V}} \tag{1}$$

where γ_{iL} and f^{o}_{iL} denote the activity coefficient and the standard-state fugacity of the i -th component in the liquid phase and ϕ_{iV} is the fugacity coefficients of the i -th component in the vapor phase. The pure component saturation pressure P_i^s is essential for calculating the standard-state fugacity of a pure liquid at a specified temperature and pressure. For ideal systems at low pressures Eq. (1) is reduced to:

$$K_i = P_i^s / P\,, \tag{2}$$

where P is the total pressure. Thus, for ideal systems the prediction of P_i^s is sufficient for phase equilibrium calculations. At high pressures the gas and liquid fugacity coefficients are usually modeled by equations of states. The equation of state form of the phase equilibrium ratio is usually written as:

$$K_i = \phi_{iL} / \phi_{iV}\,. \tag{3}$$

where ϕ_{iL} is the partial fugacity coefficient of a species in the liquid mixture and ϕ_{iV} is the partial fugacity coefficient of a species in the vapor mixture.

To calculate the fugacity coefficients, either the Soave-Redlich-Kwong or Peng-Robinson EoS are usually employed. In the EoS, the properties of the pure compounds required are the critical temperature (T_c) and pressure (P_c), and the acentric factor (ω). The application of EoS to mixtures requires the use of mixing rules for the mixture energy parameter a_{mix}, which accounts for interactions between the species in the mixture, and for the co-volume parameter b_{mix}, which accounts for the excluded volume of the species of the mixture:

$$a_{\text{mix}} = \sum_i \sum_j x_i x_j a_{ij}\,; \qquad b_{\text{mix}} = \sum_i x_i b_{ii} \tag{4}$$

The cross coefficient a_{ij} is related to the corresponding pure-component parameters by the following combining rule:

$$a_{ij} = \left(a_{ii} a_{jj}\right)^{0.5} \left(1 - k_{ij}\right) \tag{5}$$

Thus, for phase-equilibrium calculations some of the following properties are usually needed:

1. Pure component constant properties: T_c, P_c, and the ω.
2. Binary interaction parameters for the calculation of ϕ_{iL} and ϕ_{iV} (using equation of state), and the activity coefficients (e.g., using the UNIFAC model).
3. Pure component saturation pressures at various temperatures.

3. The Combination of the SC-QS2PR and TQSPR methods

The SC-QS2PR and TQSPR methods are described in detail by Cholakov et al., 2007 and Brauner et al., 2006. A brief description of the new combined method follows. It is assumed that an unknown *target property* of *a target compound* (y_t) can be represented

as linear combination of known property values of two *predictive* compounds y_{p1} and y_{p2}, provided that the predictive compounds are *similar* to the target compound:

$$y_t = \beta_1 y_{p1} + \beta_2 y_{p2}. \quad (6)$$

The coefficients β_1 and β_2 are calculated by solving the structure-structure relationship:

$$\begin{aligned} \beta_1 + \beta_2 &= 1 \\ \beta_1 \rho_{p1} + \beta_2 \rho_{p2} &= \rho_t \end{aligned} \quad (7)$$

where ρ_{p1} and ρ_{p2} are molecular descriptor values of the predictive compounds and ρ_t is the (known) molecular descriptor value of the target compound. Solving Eq. (7) for β_1 and β_2 and introducing these values into Eq. (6) yields:

$$y_t = \frac{\rho_{p2} - \rho_t}{\rho_{p2} - \rho_{p1}} y_{p1} + \frac{\rho_t - \rho_{p1}}{\rho_{p2} - \rho_{p1}} y_{p2}. \quad (8)$$

For obtaining reliable prediction via the property-property correlation (Eq. 8), the predictive compounds must belong to a *similarity group* of the target compound and the molecular descriptor ρ must be collinear with the property y for the members of the similarity group.

If the target compound belongs to a homologous series, the natural similarity group comprises members of the same homologous series for which experimental values of the target property are available. In other cases the identification of the similarity group can be carried out using methods described by Brauner et al. (2006). To identify a descriptor which is collinear with the property y for the members of the similarity group, a stepwise regression program is employed on a dataset which contains the property y as the dependent variable and various descriptors (associated with the members of the similarity group) as independent variable.

The prediction of pure component constant properties by applying the proposed method is straightforward (Cholakov et al, 2007). The prediction of binary interaction parameters is discussed elsewhere. Following is an example of application of the proposed method for prediction of vapor pressure of 1-alkenes.

4. Predicting the Vapor-Pressure (Saturation Temperature) of 1-alkenes Using the Proposed method

4.1. Databases and Programs Used in the Study

A database containing 1280 descriptors for 324 compounds was used for this study. The descriptors were calculated with the Dragon, version 5.4. software (DRAGON is copyrighted by TALETE srl, http://www.talete.mi.it) using minimized molecular models. The SROV (Shacham and Brauner, 2003) stepwise regression program was used for the selection of descriptors collinear with a property. Experimental normal boiling temperature data and Antoine equation constants were taken from the NIST (2005) database.

4.2. Deriving the Saturation Temperature (T_s) Prediction Equation

All the correlations which represent the vapor pressure variation as function of temperature (such as the Antoine, Riedel and Wagner) can be used for prediction of vapor pressure for the target compound. We will be using the Antoine equation for this demonstration. The Antoine equation is:

$$\log(P_i^s) = A_{pi} - \frac{B_{pi}}{C_{pi} + T}. \qquad (9)$$

where A_{pi}, B_{pi} and C_{pi} are the Antoine equation coefficients for the i^{th} predictive compound, and P_i^s (bar) is the saturation pressure at temperature T (K). Introducing the $\log(P_i^s)$ values from Eq. (9) into Eq. (8) yields $\log(P_t^s)$, where P_t^s is the saturation pressure (bar) of the target compound at temperature T (K). This method can be viewed as an enhancement of the "two reference fluid" estimation method (described, for example, by Poling et al., 2001) where the same two predictive compounds often used for all the target compounds and the acentric factor is employed, instead of the molecular descriptor in Eq. 8. Using predictive compounds which are similar to the target is, of course, preferable to the use of the same predictive compounds for all targets, also the use of molecular descriptors is preferable to the use of acentric factor (which requires experimental data of T_c and P_i^s value at $T = 0.7\ T_c$).

An alternative form of the Antoine equation, which is explicit in the saturation temperature, T_i^s, at pressure P is the following:

$$T_i^s = \frac{B_{pi}}{A_{pi} + \log(P)} - C_{pi}. \qquad (10)$$

Introducing the T_i^s values from Eq (10) into Eq. (8) yields the value of the saturation temperature of the target compound at pressure P. Recalling that the normal boiling temperature (T_b) is the saturation temperature at atmospheric pressure suggests the use of a descriptor collinear with T_b in the structure-structure correlation (Eq. 7). It is assumed that for all the members of the similarity group (and the target compound) the values of a descriptor collinear with T_b are available. For few (at least two) members of the similarity group experimental data and/or models for vapor pressure need also to be available.

4.3. Prediction of T_t^s for the 1-alkene Series

In Table 1 seventeen members of the 1-alkene series are shown with the vapor pressure data and equations that are available in the NIST (2005), and values of the descriptor Vev1 (a 2D eigenvalue-based descriptor: eigenvector coefficient sum from van der Waals weighted distance matrix , calculated by the Dragon program). The data shown includes the number of carbon atoms (n_C), T_b, the recommended temperature range of applicability of the Antoine equation for vapor pressure calculation the Antoine equation coefficients: A, B and C. Up to 1-hexadecene experimental T_b values are available as well as Antoine equations parameters. For higher n_C, only few T_b values and no Antoine-equation parameters are available. However, values of the VEv1 descriptor can be calculated, irrespective of n_C.

In Figure 1 a plot of T_b versus the values of the descriptor VEv1 for the 1-alkene series is shown. Observe that the T_b data can be represented by a straight line, T_b = 1103.7 Vev1-12.647, with a linear correlation coefficient R^2 = 0.9981. This equation can be used for predicting T_b for the compounds for which no experimental values are available.

Let us select 1-decene and 1-undecene as predictive compounds and 1 – dodecene as target compound. Reliable prediction of P_t^s or T_t^s can be obtained only if the predictions are carried out in the common temperature applicability range of the Antoine equations for the predictive compounds (for the case of P_t^s), or the common pressure applicability range of the same equations (for the case of T_t^s). In Figure 2 the vapor pressure vs. temperature of the predictive and target compounds (obtained by the Antoine-equation parameters of Table 1) are shown. Looking at the temperature range of validity of the high and low pressure-range parameters of the predictive compounds shows a rather limited overlap between those ranges. However, there is a complete overlap between the pressure ranges of validity. This provides an additional justification for preferring the prediction of $T_t^s(P)$ over the prediction of $P_t^s(T)$.

Introducing ρ_{p1} = 0.4093; ρ_{p2} = 0.4663; and ρ_t = 0. 0.448 into the structure-structure correlation (Eq. 7) yields the parameter values: β_1 = 0.48814 and β_2 = 0.51186. T_t^s was predicted using the SC-QS2PR method with these parameter values and compared to values obtained by the Antoine equation (the parameters are given in Table 1). The difference between the SC-QS2PR prediction and the Antoine equation calculation is the prediction error. The highest absolute prediction error obtained was 0.2 K (0.04 %) at P = 0.9 bar. Considering that the uncertainty in the reported T_b value of 1-dodecene is 0.6 K, the prediction uncertainty is below the experimental uncertainty in this case.

The use of the two immediate neighbors of 1-dodecene in the homologous series as predictive compounds can explain the very low prediction error in this case. For a longer range interpolation we have selected 1-octene (n_C = 8) and 1-tridecene (n_C = 13) as predictive compounds, keeping 1-undecene (n_C = 11) as target compound. This yielded an SC-QS2PR with β_1 = 0.37188 and β_2 = 0.62812. The highest prediction error in this case is 1.33 K (0.3%) at P = 0.9 bar. Thus, although there is some deterioration in the prediction accuracy when using a longer range interpolation, the precision is still very good.

For testing the extrapolation capabilities of the SC-QS2PR method, we have selected 1-octene (n_C = 8) and 1-undecene (n_C = 11) as predictive compounds, and 1-tetradecene (n_C = 14) as target compound. This yielded an SC-QS2PR with β_1 = -0.87029 and β_2 = 1.8703. The highest prediction uncertainty in this case was -3.63 K (0.7%) at P = 0.9 bar. Thus, extrapolation yields mixed-sign parameter values and higher prediction uncertainty than interpolation.

Dearden (2003) provides prediction uncertainty values for current state of the art QSPRs for predicting T_b. For the Alkenes group a QSPR which includes 5 topological descriptors is mentioned. The average prediction uncertainty using this QSPR reported us 2.3 %. Thus the prediction uncertainties with the proposed method are lower than those of state of the art QSPRs even for the case of extrapolation.

5. Conclusions

It has been demonstrated that the new combined SC-QS2PR and TQSPR methods enables the prediction of vapour pressure within experimental uncertainty, depending on the level of similarity between the predictive compounds and the target compound. In cases of longer range interpolation or extrapolation there is some degradation of the precision of the prediction, but in the cases tested (for n-Alkane, 1-alkene, alkanoic

acid, n-alcohol and alkyl-benzene homologous series) the prediction uncertainty did not exceed 1%.

The new method has been applied successfully for prediction of pure component properties and binary interaction parameters. The results of these studies are reported elsewhere.

References

N. Brauner, R. P. Stateva, G. St. Cholakov and M, Shacham, 2006, Ind. Eng. Chem. Res., 45, 8430-8437

G. St. Cholakov, R. P. Stateva, M. Shacham and N. Brauner, 2007, AIChE J , 53(1), 150-159

J. C. Dearden, 2003, Quantitative structure–property relationships for prediction of boiling point, vapor pressure, and melting point, Environ. Toxicol. Chem., 22, 1696–1709.

NIST Chemistry WebBook, NIST Standard Reference Database Number 69, June 2005, Eds. P.J. Linstrom and W.G. Mallard, National Institute of Standards and Technology, Gaithersburg MD, 20899 (http://webbook.nist.gov).

B. E. Poling, J. M. Prausnitz and J.P. O'Connel, 2001, *Properties of Gases and Liquids*, 5th Ed., McGraw-Hill, New York.

M. Shacham and N. Brauner, 2003, The SROV program for data analysis and regression model identification, Comput. Chem. Eng., 27, 701–714.

Table 1. Vapor Pressure Related Data from the NIST (2005) Database and Values of the Descriptor Vev1 for the 1-alkene Homologues Series.

No.	Component Name	No. of C atoms	Tb (K) Experimental	Descriptor Vev1	Temp. Range Low	Temp. Range High	Antoine Constants A	Antoine Constants B	Antoine Constants C
1	1-butene	4	266.8	0.2599	195.6	269.4	4.24696	1099.207	-8.256
2	1-pentene	5	304	0.2902	285.98	303.87	3.91058	1014.294	-43.367
3	1-hexene	6	337	0.3177	289.04	337.46	3.99063	1152.971	-47.301
4	1-heptene	7	367	0.3429	273.19	361.89	4.21811	1400.674	-34.193
5	1-octene	8	395	0.3663	318.04	395.37	4.05752	1353.486	-60.386
6	1-nonene	9	419	0.3884	339.76	421.01	4.07879	1435.359	-67.615
7	1-decene	10	440	0.4093	359.92	444.7	4.08526	1501.872	-75.572
8	1-undecene	11	466	0.4291	379.02	466.89	4.09152	1562.469	-83.407
9	1-dodecene	12	486.2	0.4480	396.8	487.62	4.10012	1619.862	-90.879
10	1-tridecene	13	506	0.4663	415.7	505.81	4.11053	1674.741	-97.936
11	1-tetradecene	14	524	0.4838	431.61	524.9	4.14495	1745.001	-102.675
12	1-pentadecene	15	541.4	0.5007	446.76	541.42	4.14045	1781.974	-110.568
13	1-hexadecene	16	557.6	0.5170	461.3	557.92	4.16208	1837.811	-115.859
14	1-heptadecene	17	573.2	0.5329	-	-	-	-	-
15	1-oktadecene	18	-	0.5483	-	-	-	-	-
16	1-nonadecene	19	-	0.5633	-	-	-	-	-
17	1-eicosene	20	614.2	0.5778	-	-	-	-	-

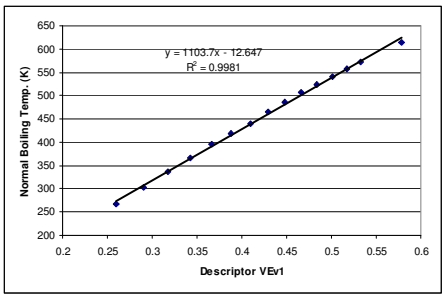


Figure 1. Plot of T_b versus the descriptor Vev1 for the 1-alkene series

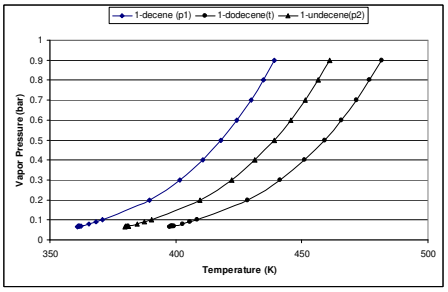


Figure 2. Vapor pressure versus temperature of the predictive and target compounds

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Simultaneous synthesis of the downstream process and the reactor concept for the Oxidative Coupling of Methane (OCM)

Steffen Stuenkel, Hamid R. Godini, Jens-Uwe Repke, Guenter Wozny

Berlin Institute of Technology, Department of Process Engineering, Chair of Process Operation and Dynamics, Straße des 17. Juni 135, 10623 Berlin, Germany
E-mail: Steffen.Stuenkel@tu-berlin.de

Abstract

The Oxidative Coupling of Methane (OCM) is a promising alternative for the oil based production of olefins. The aim is the catalytically conversion of methane containing natural gas to ethylene, which builds up a base for olefins and further synthesis. The overall yield is still limited to 30%. Beside new catalysts and reactors, new concepts for integrated downstream processes are necessary to overcome the exist limitations [3]. To meet this target as fast as possible, the downstream process, the reactor and the catalyst are designed simultaneously. This novel strategy causes particular interactions between the downstream process, the reactor and the catalyst design. Design specifications for each part of the plant are established as targets such as carbon dioxide concentration.
In order to modify and investigate the different approaches in our mini plant easily, the whole process is designed and built up modular. The entire process is divided into three units: reaction, purification and separation unit. On each unit interfaces of each unit the targets and the process specification are defined by laboratory screenings. Due to the novel process design strategies, the downstream impose specification to the reaction unit and the catalyst too e.g. maximum carbon dioxide concentration for the reaction product. For this purpose different reactor concepts like membrane reactors, fluidized bed and fixed bed reactors have to be investigated along with different integrated downstream concepts. The most promising reactor concepts are discussed in this article regarding their implementation and their influence on the separation process.
The downstream concept should also be design appropriately for the OCM Process. Therefore a novel hybrid purification process, a combination of a membrane and an amine unit for the carbon dioxide removal could be implemented and is presented in this article. The modification results of this novel hybrid process are presented along with the simulation results for the hybrid process and they are discussed regarding the whole OCM Process. Thus the results have to be validated in our mini plant.

Keywords: Downstream OCM, membrane, hybrid process, miniplant.

1. Introduction

Oxidative coupling of methane (OCM) is a novel technology for the conversion of natural gas to ethylene, reaching widespread attraction among various research groups in the last two decades. OCM is a set of catalyzed surface induced gas phase reactions and the overall yield of the desired products is still limited to reach 30%. New and efficient downstream processes are needed for sustainable exploitation of resources. The novel strategy of simultaneous design for each part of the whole OCM process causes particular interactions between the downstream process, the reactor and the catalyst.

Various processes for the OCM with integrated downstream concepts have been proposed 0. The generalized sequence of sections in the OCM process is shown in Figure 1. All processes emphasize on the importance of the product separation under high pressure and recycling the methane, which have major impact on the process economics. Evaluating the alternatives in process synthesis regarding to separation efficiency, energy consumption, operating and investment cost are essential for industrial application of the OCM Process, but rarely have been reported.

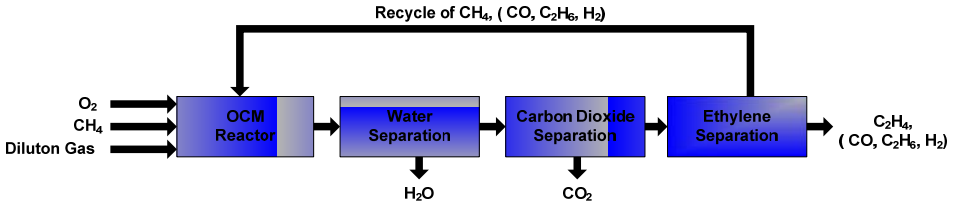


Figure 1: Process Flow Diagram of the OCM Process

2. Process synthesis

The OCM Process has not been industrially applied yet. Apart from general information on the reaction kinetics, the heat and mass transfer efficiency, fundamental studies of the process possibility, catalyst life time, effects of recycles and efficiencies of each unit are crucial for process implementation. These can only be achieved by investigation of the real process. The mini plant technique is a well known technique for the process synthesis, to obtain the required fundamental information experimentally. According to the simultaneous process design and due to economical reasons, the downstream imposes some requirements to the reaction unit and the catalyst, especially to the yield, the C_2 and CO_2 selectivity, and the methane conversation rate and dilution concentration.

2.1.1. Novel design Strategy - Simultaneous design and construction

For the process synthesis, the whole OCM Process is divided into three general unit operations: the reaction unit, the purification unit and the separation unit. All of them are linked and investigated simultaneously. Therefore a design case for each unit has been defined according to the composition, which is presented in table 1.

Table 1: Defined feed gas concentration in [vol%] for the process units

Unit	CH_4	O_2	C_2H_4	C_2H_6	CO_2	H_2O	Inert gas
Reaction	60 – 70	20	-	-	-	-	20 – 10
Purification	45	-	10	10	25	-	20 -10
Separation	60	-	13	13	-	-	14

Table 2: Process conditions for each unit

Unit	Pressure range	Temperature range
Reaction	1 to 30 bar	30 to 900 °C
Purification	1 to 35 bar	30 to 100 °C
Separation	Up to 35 bar	Down to -100 °C

The range of the process and operating conditions for the units is presented in table 2. At this step, those conditions are defined by literature study, since they have to be determined and evaluated carefully.

3. The reactor concept

The operating parameters in the reactor along with the reactor concept affect not only the performance of this unit but also sequentially affect the efficiency and economy of the separation process and subsequently the whole OCM Process.

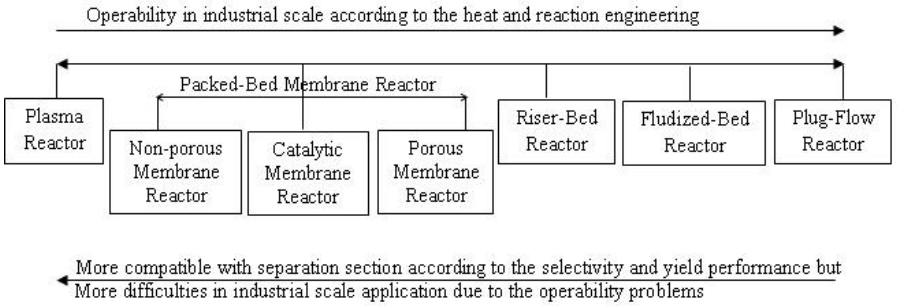


Figure 2: Comparison of different reactor concepts

Figure 2 addresses the major reactor concepts which have been reported for the OCM Process [4]. Some of these structures are very efficient in points of selectivity and yield but they have some difficulties in operation and their performance is not stable which is very crucial for an industrial-scale reactor. The more practical reactor concepts sever from low yield or selectivity and consequently affect the process economy. Therefore most of these alternatives can not be simply applied in an industrial scale yet. The potential of operability is also pointed out in Figure 2. In conventional reactors, like the plug flow or the fluidized bed reactor, either the yield is to low or the selectivity is not acceptable, and this reduces the process performance drastically. On the other side, the application of advanced reactor concepts such as membrane reactors reduce the separation effort and increase the process economy, but they have practical implementation difficulties. The dense membrane reactor uses an oxygen-permeable dense membrane. This membrane is non-permeable for other components and shows a very high selectivity. Due to the low permeation rate, the scale-up performance for this kind of reactor is not validated yet. The plasma reactor is even harder to be implemented as an industrial scale reactor regarding the energy consumption and providing a sustainable plasma media. The riser-bed reactor struggles with mass transfer limitations, back-mixing effects and low selectivity due to the undesired gas phase reactions. It has to be noticed that membrane reactors are preferred regarding the yield and selectivity, but the operability in general has to be improved in advanced reactor concepts. Whereas the fluidized bed reactor are preferred regard to the operability due to the better heat transfer performance, but the selectivity for this reactor concept should be improved. In case of operating condition, especially the composition of reactor product stream should be considered as an important factor. For example the type and quantity of the inert gas for diluting the reaction media, strongly affect the separation section. This becomes more important when the hybrid separation system such as the proposed one in this research, are used to remove the undesired components and purify the products. Furthermore the reactor can be operated under pressure up to 30 bar, which lowers the energy demand for the product separation.

4. The downstream concept for the OCM Process

The downstream section of the OCM Process consists of a phase separation unit, a carbon dioxide removal unit and a product separation unit, as addressed by diffrent researchers 0. Concerning the simultaneous design and improvement in the miniplant, the state of the art separation processes are taken as a base case for the downstream units. The purification unit consists of an amine based absorption process for the carbon dioxide separation. The separation unit consists in a cryogenic distillation for the product separation. Due to the high energy consumption of the cryogenic distillation, the pressure is increased up to 35 bar to increase the boiling point of the hydrocarbons. Regarding the idea of using carbon dioxide as an inert gas for the dilution in the OCM reactor, the carbon dioxide removal step becomes even more important in the downstream process. Therefore, the purification section as the first downstream unit is picked out and the carbon dioxide removal is investigated particularly.

4.1. The purification unit

The specification of this unit is to remove the whole carbon dioxide from the product stream as given in table 1. Such a requirement is not unusual in the process industry, but attracts wide interest nowadays. Different approaches are known for those separation units like: absorption processes, adsorption processes, cryogenic distillation or membrane processes. These processes are based on different physical and chemical principles:

- Absorption: physical or chemical absorption in liquids, caused by the gas solubility in the liquid or in combination with a superimposed chemical reaction.
- Adsorption: physical or chemical adsorption on a particle surface of the sorbent.
- Cryogenic separation: caused by the different condensation points of the gas species
- Membranes: selective solubility and diffusion or molecular sieve effects caused by different molecule dimensions and Knudsen diffusion.

Here again, there is a trade off between the operability and efficiency of the separation process. A well developed and industrial applied technique is the absorption process with chemical or physical absorption liquids. A modern technique is the membrane separation, which has low selectivity for carbon dioxide yet and is mainly implemented as stand alone units for biogas or natural gas cleaning. The adsorption is available only for small gas streams and is hard to handle, concerning the sorbent regeneration. The range of the operation conditions for the cryogenic distillation are up to 60 bar for the pressure and from -20 °C to 20 °C for the temperature, which is economically applicable only in combination with liquefaction and storage of the carbon dioxide. The membrane and the absorption techniques can be an alternative for the purification unit of the OCM miniplant.

4.1.1. Absorption processes

The absorption technique for the carbon dioxide separation is well developed and it is available in industrial scale. As physical absorption processes the UOP Selexol® or the Lurgi Rectisol® Process are known, which using dimethyl ether and methanol as sorbent respectively. Those physical absorption processes causes high product losses of more than 30 vol%, due to a nearly similar solubility of the product and the carbon dioxide in the absorbent. Therefore chemical solvents like Monoethanolamine (MEA), Diethanolamine (DEA) and Methyldiethanolamine (MDEA) or a mixture of them are more applicable for the purpose in the OCM miniplant. Those chemicals are used in amine sweetening processes like the aMDEA® Process in different concentration ranges 0.

Rigorous simulations for the absorption process of the miniplant were carried out in Aspen Plus® 0. As sorbent 15 wt% MEA and 30 wt% MDEA solution were compared. Table 3 summarizes the basic engineering details for the column design.

Table 3: Technical and hydrodynamic operation conditions of the absorption process

Packing height [m]	Column diameter [m]	F-factor [$Pa^{0,5}$]	Gas feed kg/h]	Packing section [-]	Packing capacity [m^2/m^3]	Maximum liquid load [m^3/m^2h]	Liquid stream [kg/h]	Top pressure [bar]
5	0.04	0.8	21	50	450	55	70	35

The carbon dioxide concentration can be reduced to 15 vol% with 25 kW energy demand using MEA solution, whereas with MDEA the carbon dioxide concentration can be reduced down to 7 vol% with only 5 kW energy input for regeneration. This separation efficiency obviously is caused by the different solvent concentration, but they are corrosion limited 0. Neither MEA nor MDEA as solvent can remove the carbon dioxide totally in a standalone absorption process under the given conditions.

4.1.2. The one stage Membrane unit

The advantages of a membrane unit are the easy operation and having a short start up and shut down time due to their small size 0. Those units are very flexible in operation, regard to their modular design. Furthermore no additives are necessary for the operation. Therefore no regeneration is required, which is the most energy consuming part of absorption processes. For vapor/gas membrane separation different kinds of materials are available:

- Polymeric membranes: rubbery or glassy polymers, with different solubility and diffusion properties for carbon dioxide and hydrocarbons.
- Molecular sieves: different pore size, separation by the different molecule diameters.

Glassy and rubbery polymeric membrane preferred for the carbon dioxide separation and hydrocarbon recovery 0. In the investigated membrane unit a carbon dioxide selective membrane is applied 0. The membrane unit is calculated using geometry and permeability data for a GKSS flat sheet membrane module 0.
The carbon dioxide concentration could reduce with a one stage membrane unit of an area of 0.5 m² down to 14 vol %. The product losses in this unit are in the range of 30 vol%, similar to those of the physical absorption and for the purpose in the OCM miniplant not applicable.

4.1.3. The two stage membrane process

By the application of a two stage membrane unit the product losses could reduced drastically, whereas the carbon dioxide reduction is of the same range as for one stage membrane system: down to 14 vol%.

Table 4: Technical requirements of the two stage membrane process

Membrane surface for the 1st Stage	Membrane surface for the 2nd stage	Pressure second stage	Product losses
1 m²	0.5 m²	6 bar	10 vol%

Table 5: Energy balance of the hybrid process

Compressor power	Cooling power	Heating power	Pumping power
1.5 kW	3.5 kW	5 kW	1 kW

The dimension and process conditions of the two stage process are shown in Table 4. Neither with the one stage membrane unit nor with the two stage membrane unit the carbon dioxide can be removed totally. Furthermore membrane units can easily extended and therefore the downstream process can be adjusted easier to changed process and operation conditions.

5. Conclusion – The hybrid separation process

Concerning the mentioned facts for the reactor concepts, the main remaining alternatives for the OCM reactor are the membrane and the fluidized bed reactor. The limitations for those concepts must be improved and have to overcome for industrial application. Regarding the required purity of the product stream and the high energy demand of energy for the absorption, the use of a membrane section for the pre - separation of carbon dioxide could apply successfully. This hybrid membrane amine process, figure 3, can remove the carbon dioxide in the product stream totally. With membrane units the downstream is more flexible regarding the process and operation conditions. A two stage membrane system shows the lowest hydrocarbon losses and was combined with an absorption column to form a hybrid process. The overall energy consumption is listed in table 5. The ethylene loss could reduced to 10 vol% with an over all energy consumption of 11 kW. In a next step the whole process has to be economically optimized and the simulation results have to be validated in the miniplant by experiments.

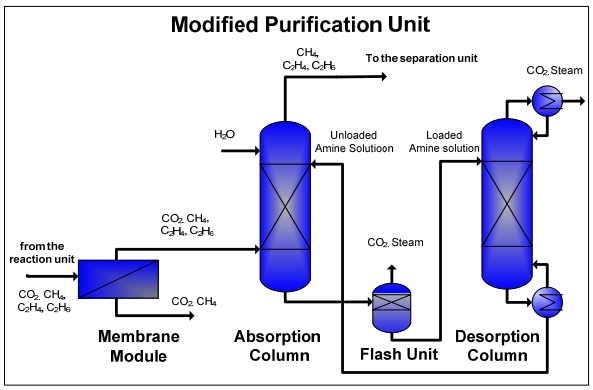


Figure 3: Hybrid separation process for the purification

6. Acknowledgements

The authors acknowledge the support from the Cluster of Excellence "Unifying Concepts in Catalysis, coordinated by the Berlin Institute of Technology and funded by the *German Research Foundation (DFG)*.

References

Kent R. Hall, "A new gas to liquids or gas to ethylene technology", Catalysis today, 2005, 106, 243

Hugill et. al., "Feasibility study on the co-generation of ethylene and electricity through oxidative coupling of methane", Applied Thermal Engineering 25 (2005) 1259–1271

M. Driess et.al., "Unifying Concepts in Catalysis" Evaluation Presentation of the Cluster of Excellence for the German Research Foundation , Bad Honef, 6. Juni 2007

E.E. Wolf (Ed), Methane Conversion by Oxidative Processes, Fundamental and Engineering Aspects. Van Nostrand Reinhold, New York, 1992

Kohl A. L., Nielsen R., „Gas Purification", Gulf Pub Co, 5th edition, 1997

Stuenkel et. al. "Modelling and simulation of a hydrid separation process for the Carbon Dioxide removal of the Oxidative Coupling of Methane (OCM) Process", accepted on the 19 th European Symposium on Comupter Aided Process Engineering, Poland

Deibele, Dohrn "Miniplant-Technik", WILEY-VCH Verlag, Weinheim 2006

Brinkmann et al. "Membranverfahren in der Erdgasaufbereitung", Chemie Ingenieur Technik, 2003, Vol. 75, Iss. 11, pP 1607 – 1611

GKSS-Forschungszentrum Geesthacht in der Helmolz-Gesellschaft, Germany

Baker, R.W., Membranes for vapor/gas separation, Membrane Technology and research, Inc, 2006

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Rigorous Design of Complex Liquid-Liquid Multi-Staged Extractors Combining Mathematical Programming and Process Simulators

José A. Caballero[b], Ignacio E. Grossmann[b]

a Department of Chemical Engineering, Universidad de Alicante, Apartado de Correos 99, 03080 Alicante, Spain
b Department of Chemical Engineering, Carnegie Mellon University, Pittsburgh, PA. 15213 USA.

Abstract

This work addresses the design of liquid-liquid multistage extractors using a mixed approach that combines mathematical programming with explicit equations, and the rigorous models that are available in commercial Chemical Process Simulators. A superstructure that has embedded all the potential configurations and interconnections is proposed. Based on this superstructure representation the problem is formulated as an optimization problem using generalized disjunctive programming (GDP) to minimize the total cost of the process, subject to design specifications. The method determines the optimal number of equilibrium stages and flow rates needed to obtain the specified product separation and recovery, and accounts for the possibility of side feed streams and product extractions.
A commercial chemical process simulator (HYSYSTM or UnisimDesingTM) that is dynamically connected with the external optimizer that also deals with explicit equations performs the complex liquid-liquid equilibrium calculations. A case study involving a quaternary mixture is presented in order to illustrate the robustness and efficiency of the model.

Keywords: Liquid-Liquid extraction, MINLP, Process Simulators, Process Synthesis, Disjunctive Programming

1. Introduction

Liquid-liquid extraction is a competitive separation technology mainly used when distillation is difficult or impossible to apply, i.e. very low relative volatilities, thermo sensitive products or when huge amount of heat is needed [1] The current methods for design of extractors can be separated into three categories [2]. Graphical Methods, simulation based methods and superstructure optimization based methods. Graphical methods address the design problem using triangular diagrams, where the mass balances are represented by operation lines and the tie lines connect the composition of phases in equilibrium, and therefore are restricted to ternary systems [3,4].
Simulation based methods include approaches that apply numerical methods to solve simulated process formulations using geometric principles employed in graphical methods. Marcilla et al [5] have proposed an analytical extension of the classical Ponchon-Savarit method for distillation columns that allows the direct determination of the number of stages for a given separation. However, simulation based approaches are

usually 'blind' to the final user because they are generally included in a chemical process simulator in which the number of stages is fixed and the design must be done by a trial and error approach. It is worth remark, however, that the process simulators are very robust and reliable

The most complex and flexible approach rely on a superstructure that embeds all the possible alternatives treating the existence or no existence of the different stages with binary variables and taking into account the non linear equations for the equilibrium relations. The design problem is then formulated as an MINLP problem or a Disjunctive Programming Problem (GDP) [6]. The major difficulty with this last approach is related with the non-linear and non-convex nature of the liquid-liquid phase equilibrium that in most situations makes difficult even to find a feasible solution.

Commercial chemical process simulators (Hysys™, AspenPlus™, Prosim™, UnisimDesign™) include unit operations that deal with the vapor-liquid-liquid equilibrium in a fast a reliable way using tailored numerical methods –of course proved that the thermodynamics correctly describes the system- that take explicitly into account aspects like trivial solutions, existence or not of phase separation, etc.

An ideal situation would be taking advantage of the best of both approaches, the numerical reliability robustness included in chemical process simulators and the flexibility and optimization capabilities of a superstructure base optimization. In this line Papadopoulos and Linke [2] presented a superstructure based optimization using stochastic (Simulated annealing, genetic algorithms and ant colony) methods. These methods have the advantage of being very easy of implementing, do not require any special structure in the model and usually obtain near optimal global solutions. However, they usually require a large number of function evaluations, which remarkably increases with the complexity of the problem and usually includes a set of adjustable parameters that affect the performance of the algorithm.

In this work, we propose a mathematical programming approach based on a superstructure that includes implicit equations, in particular each stage of the liquid-liquid extractor that are calculated in a process simulator, and explicit equations (mass balances between stages, purity constraints, recycles, etc). The problem then is reformulated as an MINLP problem –in which some equations are implicit- or a disjunctive programming problem [7].

2. Problem statement

The design problem can be stated as follows: Given a set of feed streams with known composition and a specified desired separation of solutes in the product stream. The problem consists of determining the optimal number of stages, the feed streams locations, the solvent flow rate, and the existence and locations of intermediate solvent streams ,side feed product streams, fractional by-passes of the initial feed stream and the kind of contact (countercurrent, cross flow, mixed flow). The objective is to minimize the total annualized cost.

3. General extractor

The first step to solve the problem consists of generating a superstructure that takes into account all the possibilities commented above embedded in it. All process simulators include the LL extractor as a standard unit operation. However, it is not possible in a straightforward way, changing the number of equilibrium stages. Although, it would be possible to follow an approach similar to that proposed by Caballero et al [8] in distillation systems, in which a special Master problem is derived –in the context of

disjunctive programming- maintaining the distillation column as whole in the process simulator. In this case, it is better explicitly considers the extractor as an inter-related succession of Liquid-Liquid equilibrium stages, a chemical process simulator solves each equilibrium stage.

Figure 1 shows a possible superstructure, (the number of RL streams can be increased, but we include only one for the sake of clarity). The extract stream (normally organic phase) and the raffinate stream (normally aqueous phase) are in contact throughout all the stages. E_1 and R_n represent the final products. The stages are in equilibrium, although experimental efficiency factors could be easily included in order to be more realistic. Different arrangements of equilibrium stages can be used to perform extraction operations, such as single contact, repeated contact with a fresh solvent at each stage (multistage cross flow) countercurrent and countercurrent contact with solvent reflux [9]

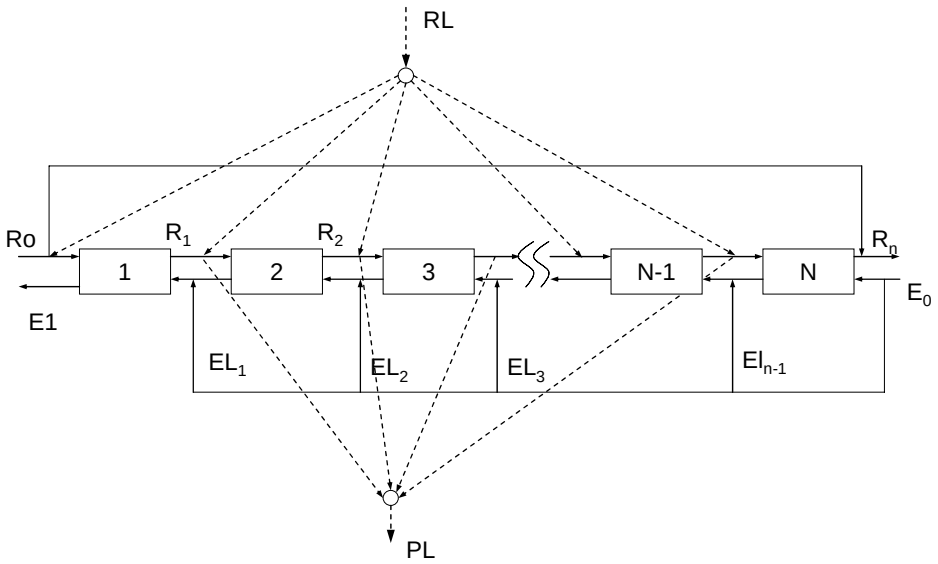


Figure 1. Superstructure of a multi-staged liquid-liquid extractor

In order to achieve a specific separation an upper bound to the number of stages in the cascade must be specified. A value of ten is usually large enough in most of the liquid-liquid multi-staged extractors. If a solution involves ten stages, one can increase the upper bound until the number of calculated staged is strictly lower than the upper bound. Due to the non-linear nature of these models it is convenient maintain this upper bound as low as possible.

The superstructure can be optimized using a big M reformulation forcing the flows entering to a given stage to take a zero value. However, this approach produces numerical problems in the process simulator (we have used UnisimDesign™ or HYSYS™). Yeomans and Grossmann [10] proposed a better approach in the context of distillation columns design, in which, for non-existing stages, the equations considered are simple input-output relations without mass or energy transference. Therefore, the only difference between existing or non-exiting stages is the application of equilibrium equations in the process simulators. This approach also simplifies the model formulation if a pure disjunctive algorithm is used. See Figure 2

4. Disjunctive Model

Conceptually the problem can be represented as follows

$$\min: Z = Total\ Cost$$

$$s.a.\quad Mass\ balances\ in\ all\ mixing\ and\ splitting\ points$$

$$Purity\ and\ recovery\ constraints$$

$$\begin{bmatrix} Y_s \\ \left(R_{i,s}, E_{i,s}\right) = \Omega(EI_{i,s}, RI_{i,s}) \\ Cost\ Stage_s = CS \end{bmatrix} \veebar \begin{bmatrix} \neg Y_s \\ R_{i,s} = RI_{i,s} \\ E_{i,s} = EI_{i,s} \\ Cost\ Stage_s = 0 \end{bmatrix} \quad \begin{matrix} s \in Stages \\ i \in Components \end{matrix}$$

$$Y_s \Rightarrow Y_{s-1}\ \forall s \in Stages \mid s > 1$$

Where Ys is a Boolean variable that takes the value "True" if the stage s exists, and the value False otherwise. Figure 2 shows graphically the disjunction representation.

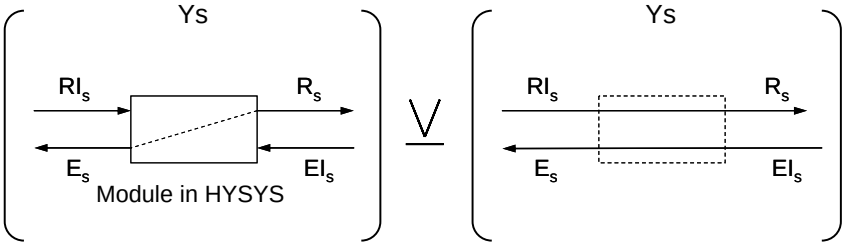


Figure 2. Disjunction for stage existence

The operator Ω refers to the calculations made by the process simulator: outputs (R, E) in terms of the inputs (RI, EI). The mass balances between the input of a given stage, and the output of previous (next) stage must be explicitly included.
Logical relationships between Boolean variables are necessary to avoid degenerate solutions (same number of stages but different locations).
The total cost is the cost of new solvent plus the cost of the stages.

$$Total\ Cost = \sum_{s \in Stages} Cost\ Stage_s + Cost\ Solvent \cdot E_0$$

5. Case study

Consider two water streams that contain acetone, and acetic acid.
The objective is designing a multistage extractor with minimum cost and determining the optimal number of stages and the optimal position of the two aqueous feeds. We want to obtain a final combined water stream with a molar fraction of water greater than 0.98, molar fraction of acetone and acetic acid lower than 0.005 and 0.05 respectively

using chloroform as organic solvent and a recovery of acetone in the organic phase greater than 90 % in molar basis. All relevant data to the problem and the optimal solution found is in Table 1.
The optimal solution found (see Figure 3) shows that the aqueous stream R02 splits in two. The 77% of this stream is introduced in stage 2 and the remaining 13% in stage 3.
The problem was solved in 374 seconds of CPU time using a proprietary implementation of the LP-NLP-BB algorithm for solving MINLP (or Disjunctive) mathematical programming problems. The problem was solved using HYSYS and UnisimDesign as chemical process simulators with identical results.

Table 1. Data and results for the case study

Feed streams

R01 = 3.5348 kmol/h [0.75, 0.12, 0.13] mol fraction Water, Acetone, Acetic Acid

R02 = 3.2810 kmol/h [0.808, 0.1158, 0.0762] mol fraction Water, Acetone AcH.

Thermodynamics: UNIQUAC, HYSYS default parameters

Optimal Number of Stages: 5

Feed Stages: R01 stage 1; R02: split 87 % to stage 2 and 13 % to stage 3

Organic Solvent: Chloroform. Optimal Flow Rate = 9.3236 kmol/h

Objective function = 516.1822 u.m. (stage cost = 10 u.m.; Solvent cost = 50 u.m/kmol·h)

	Aqueous Phase			
Stage	**Water** (mol fraction)	**Acetone** (mol fraction)	**Acetic Acid** (mol fraction)	**Chloroform** (mol fraction)
1	0.8917	0.0068	0.1001	0.0013
2	0.9209	0.0029	0.0755	0.0008
3	0.9404	0.0004	0.0586	0.0005
4	0.9601	0.0000	0.0396	0.0003
5	0.9800	0.0000	0.0198	0.0002
	Organic Phase			
1	0.0035	0.0751	0.0558	0.8657
2	0.0017	0.0397	0.0428	0.9158
3	0.0006	0.0073	0.0337	0.9584
4	0.0003	0.0007	0.0230	0.9760
5	0.0001	0.0000	0.0000	0.9999

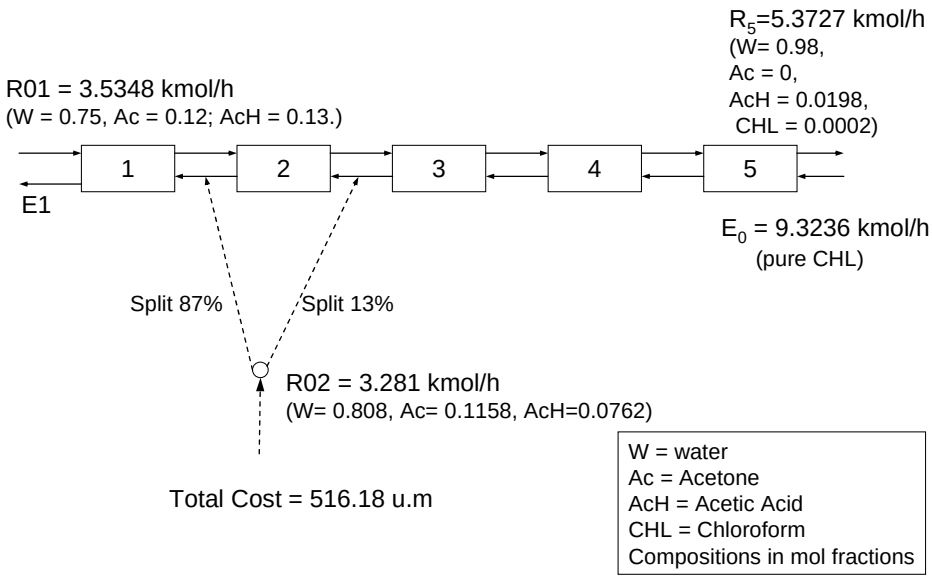


Figure 3. Optimal solution for the case of study

Acknowledgements

The authors gratefully acknowledge financial support from Spanish ‘‘Ministerio de Ciencia y Tecnología’’ under project CTQ2005-05456/

References

[1] Thornton, J.D. 1992, Science and Practice of Liquid-Liquid Extraction. Oxford University Press. New York.
[2] Papadopoulos, A. I.; Linke, P. 2004. On the synthesis and optimization of liquid-liquid extraction processes using stochastic search methods. Comput. & Chem. Engng ,28, 2391-2406.
[3] Minotti, M.; Doherty, M.F.; Malone, M.F. 1996. A geometric method for the design of liquid-liquid extraction. Ind. Eng. Chem. Res., 35(8), 2672.
[4] Minotti, M.; Doherty, M.F.; Malone, M.F. 1998. Economic trade offs for extraction systems. Chemical Engineering Research and Design. 76 (A3), 371.
[5] Marcilla, A.; Gomez, A.; Reyes, J. 1999. New method for quaternary systems liquid-liquid extraction tray to tray design. Ind. Eng. Chem. Res. 38, 3083.
[6] Grossmann, I. 2002. Review of nonlinear mixed integer and disjunctive programming techniques. Optimization and Engineering 3, 227-252.
[7] Caballero, J.A.; Odjo, A.; Grossmann, I.E. 2007. Flowsheet optimization with complex cost and size functions using process simulators. AIChE Journal, 53(9), 2351-2366.
[8] Caballero, J.A.; Milan-Yañez, D.; Grossmann, I.E. 2005. Rigorous design of distillation columns. Integration of disjunctive programming and process simulators. Ind. Eng. Chem. Res. 44(17), 6760-6775.
[9] Wankat, P. C. 1988. Equilibrium stage separation. Ed Elsevier. New York.
[10] Yeomans, H.; Grossmann, I.E.; Disjunctive Programming models for the optimal design of distillation columns and separation sequences. 2000. Ind. Eng. Chem. Res. 39(6), 1637-1648.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Simulation and Optimization of H_2 Production by Autothermal Reforming of Glycerol

Giovanilton F. Silva,[a] Andrea L. O. Fereira,[a] Samuel J. M. Cartaxo,[a] Fabiano A. N. Fernandes[a]

[a] *Department of Chemical Engineering, UFC-Universidade Federa do Ceará Campus do Pici, Bloco 709, Pici, 60455-760 – Ceará – Brazil*

Abstract

This case study focused on optimization of hydrogen production for fuel cell applications. In this case study, glycerol was chosen as a raw material and with autothermal reforming as a process of produce hydrogen. Using a commercial dynamic flow sheeting software, HYSYS 3.1, the process of hydrogen production was successfully simulated. In this research, fuel processor consists of an autothermal reactor, three water gas shift reactors and a preferential oxidation reactor was successfully developed. The purpose of this case study is to identify the effect of various operating parameters such as air-to-fuel (A/F) ratio and steam-to-fuel (S/F) ratio to get the optimum hydrogen production while made carbon monoxide lower than 10 %. From the results, an optimum A/F and S/F ratio are 5.5 and 3.5, respectively to produce hydrogen – 34.7 % (v/v), CO_2 – 60% (v/v), and CO – 0.02% (v/v). Under these optimum conditions, 83.6% of fuel processor efficiency was achieved.
Keywords: glycerol, hydrogen, autothermal reforming, HYSYS.

1. Introduction

Hydrogen will play an important role as an energy carrier of the future. Hydrogen will be used as fuel in almost every application where fossil fuels are being used today, plus the advantages of hydrogen to compare with other fossil fuels is hydrogen fuel will not emission harmful or hazardous gas. There are three categories that been analyzed for input or raw material for hydrogen production. The categories that had been studied was natural gas (consist methane, ethane, propane and butane), alcohol (methanol, ethanol and glycerol) and naphtha (kerosene or fuel jet, gasoline and diesel).

An increase in biodiesel production would decrease the world market price of glycerol. Glycerol is a waste by-product obtained during the production of biodiesel. Biodiesel is one of the alternative fuels used to meet our energy requirements and also carbon dioxide emission is much lesser when compared to regular diesel fuel. Biodiesel and glycerol are produced from the transesterification of vegetable oils and fats with alcohol in the presence of a catalyst. About 10 wt% of vegetable oil is converted into glycerol during the transesterification process. Although glycerol is used in medicines, cosmetics, and sweetening agents, world demand is limited. As such, when mass production of the biodiesel is realized, novel processes that utilize glycerol must be developed.

When biodiesel is produced in large quantity, it is important to find useful applications for the resulting large quantity of glycerol in the world market. Tyson (2003) reported that glycerol markets are limited; an increase in biodiesel production may cause glycerol

prices to decline from $1/L to $0.7/L by 2010. The money invested in purifying the glycerol would also be high (Prakash, 1998). Also, Tyson, 2003 reported that net biodiesel production costs can be reduced from US$0.63/litre of B100 to US$0.38/litre of B100 by adding value to the glycerol by-product.

Glycerol is a potential feedstock, for hydrogen production because one mole of glycerol can produce up to four moles of hydrogen. Hydrogen (H2) is mostly used in refinery hydrotreating operations, ammonia production and fuel cells (Rapagna et al., 1998). When glycerol is cracked at high temperature to produce hydrogen, it is possible to get carbon monoxide as one of the gaseous products. Studies on the degradation of glycerol have been also presented in previous papers, (Herai et al., Bühler et al. and Sadashiv et al.).

One possibility is to use glycerol as a source of hydrogen, and, in this regard, steam reforming of glycerol would be a suitable reaction. In the autothermal reforming of glycerol, synthesis gas that contains both carbon monoxide (CO) and hydrogen (H2) is produced.

Steam-reforming is endothermic and partial oxidation is exothermic. It is possible to reaction glycerol and other hydrocarbon fuels with a mixture of steam and oxygen or steam and air and carry out both reactions simultaneously. The exothermic oxidation supplies the energy for the endothermic reforming. Careful control of the oxygen content of the entering mixture is essential in these processes for maintaining proper reaction temperatures. The products of these reactions are carbon monoxide, carbon dioxide, and hydrogen. The CO requires high- and low temperature water gas shifts to oxidize it and provide additional hydrogen (Brown, 2001).

Autothermal reforming is a combination of steam reforming and partial oxidation and some other reaction that occurred depend on the conversion of raw material, catalyst used, ratio of raw material and the temperature provide during the process (Iwasaki et al., 2005). Lenz et al. (2005) described that autothermal reforming is known as the simultaneous conversion of hydrocarbons with steam and oxygen. The endothermic steam reforming reaction is generally given by four reactions:

Partial oxidation:

$$C_3H_8O_3 + 2O_2 \rightarrow 3CO + 4H_2O \quad (1)$$

Parallel Steam reforming of glycerol:

$$C_3H_8O_3 + 3H_2O \rightarrow 3CO_2 + 7H_2 \quad (2)$$

Series Water-gas shift reaction:

$$CO + H_2O \rightarrow CO_2 + H_2 \quad (3)$$

Overall reaction:

$$2C_3H_8O_3 + H_2O + 2O_2 \rightarrow 5CO_2 + 9H_2 + CO \quad (4)$$

2. Simulation and Optimization

2.1 Problem definition

The hydrogen production from glycerol for fuel cell was simulated using HYSYS software as a Figure 1 shows it. Typically, the simulation process takes the following stages:

i. Preparation Stage
   a) Selecting the thermodynamic model
   b) Define chemical components

ii. Building Stage
   a) Adding and define streams

b) Adding and define unit operations
i. Auto-thermal reforming reactor
ii. Water gas shift reactor
1. High temperature shift reactor
2. Medium temperature shift reactor
3. Low temperature shift reactor
c) Connecting streams to unit operations
d) Add auxiliary unit
i. Heater
ii. Cooler
iii. Heat exchanger
iii. Execution
a) Starting integration
b) Optimization the whole plant

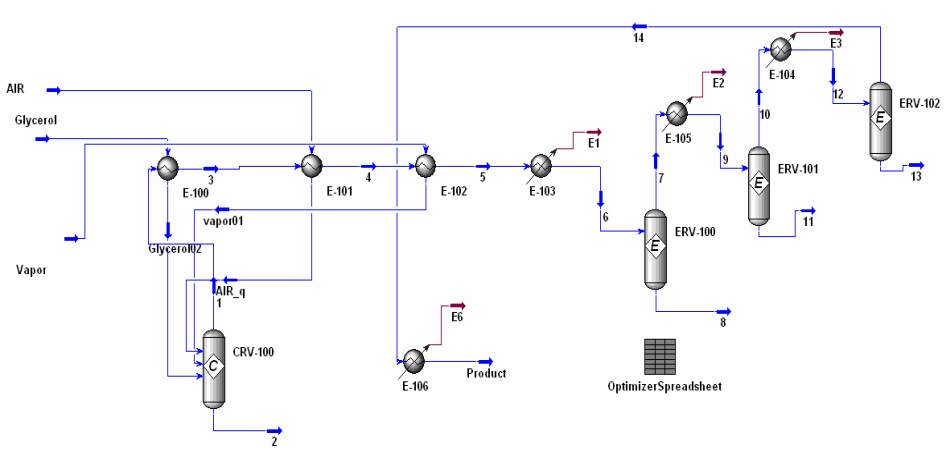


Figure 1 – HYSYS process flow diagram (PFD) of Reforming Glycerol plant.

The system considered in this study were simulated with the same basic data, show within Table 1. The Peng-Robinson Equation of State (EOS) is used to model the thermodynamics of hydrogen production for both steady-state.

Table 1 – Steady state operating conditions.

The parameter of simulation	Value
Feed flow rate of glycerol (kmol/h)	100
Reformer temperature (°C)	500
Reformer pressure (250 kPa)	250
Vapor pressure (kPa)	500
Air temperature (°C)	25

The system fuel processor efficiency can be calculated by:

$$\eta = \frac{n_{H_2} LHV_{H_2} \quad + \quad n_{CO} LHV_{CO}}{n_{C_3H_8O_3} LHV_{C_3H_8O_3}} \tag{5}$$

The lower heating value (LHV) of hydrogen, CO and glycerol are shown in Table 2.

Table 2 - Lower heating value (LHV) for hydrogen, CO and glycerol.

Component	LHV (kJ/kmol)
Glycerol	1470
Hidrogen	241.83
CO	282.00

2.2 Optimization

Optimization of the autothermal reforming of glycerol was conducted searching for the operating conditions (air-to-fuel (A/F) ratio and steam-to-fuel (S/F) ratio) that result in the highest production hydrogen. HYSYS contains a multi-variable Steady State Optimizer. The Flowsheet has been built and a converged solution has been obtained, it can use the Optimizer to find the operating conditions which minimize or maximize an Objective Function. The Objective function is given by optimization of Eq. (5). For maximize of objective functions was used the method SQP - Sequential Quadratic Programming.

CO should decrease because fuel cells (FC) do not tolerate excessive amounts of CO. The FC does not tolerate more than in the order of 50ppm CO; the lower the CO concentration, the higher the efficiency of the cell. H_2/glycerol ratio should be increased because all glycerol would be reacted to product (H_2).

3. Results and Discussions

Optimization for ATR was done by varying the air molar flow rate to get the best flow rate of air to be introduced into the ATR. Two case studies were developed in order to do this optimization. The first case study was developed to monitor the temperature at the ATR vapour stream after varying the air molar flow rate from 350 kmol/h to 800 kmol/h. The second case study was developed to monitor the molar flow rate of carbon monoxide and hydrogen after varying air molar flow rate within the range that was chosen from first case study. The result is shown in Figures 2 and 3.

From Figure 2, the temperature out of ATR is over 800 °C only after the molar flow rate of air greater or equal 600 kmol/h. With that air molar flow rate range, the hydrogen and CO molar flow rate was monitored. From figure 3, the flow rate of hydrogen produced by the reactor reforming of is decreasing when of air molar flow rate greater than 550 kmol/h. Then it began constant after 750 kmol/h.

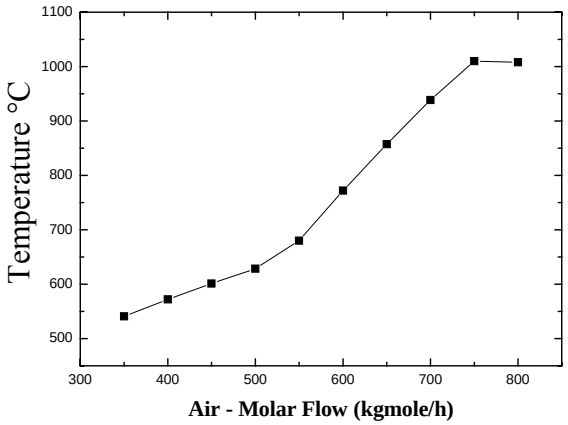


Figure 2 – Temperature of reactor of reforming for varies air Feed molar flow

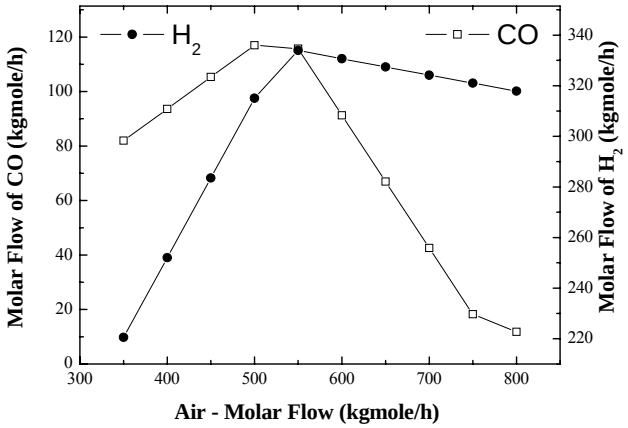


Figure 3 - Molar flow of CO and H2 effluent for varies air feed molar flow.

In Water Gas Shift Optimization , one case study was developed to optimized value of feed water molar flow to reduce concentration of CO through water gas shift reaction. Figure 3 shows the result of case study where the concentration of H2 and CO after water gas shift reactors was monitored.

Water molar flow rate was optimized from 2000 to 8000 kg/h. As we can see from Figure 5.8, the H_2 show an increasing slope and the increasing is a bit slower at 7500 kg/h. The optimum water molar flow rate was taken when hydrogen at its higher molar flow rate. So, the value of water molar flow rate that was chosen was 5500 kg/h. At this point, H_2 produced the greatest flow rate and CO reduced the lowest flow rate.

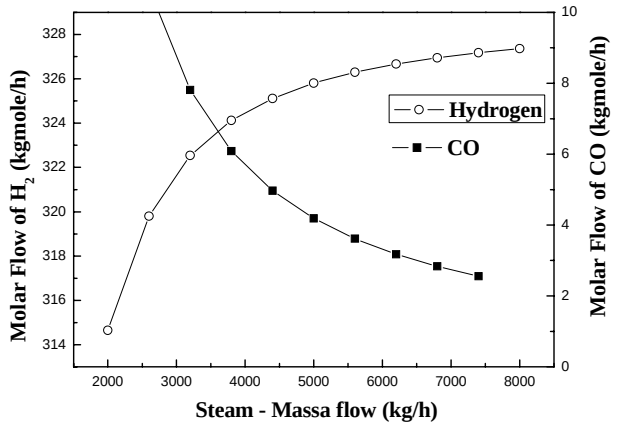


Figure – 3 Molar flow of CO and H2 effluent for varies water feed molar flow.

In the following, the results obtained for an autothermal reforming of glycerol system are presented. With the developed system models which are implemented in the HYSYS 3.1 process simulator, effluents from all reactors are simulated. In this model, the air to fuel ratio is set to 5.5 and the steam to fuel ratio is set to 3.5 In these conditions, 90% glycerol is converted to produce 34% hydrogen, 20.07% CO2 and 0.01% CO. Also, under these conditions, oxygen is 100% consumed. As we know, WGS reaction will convert CO into CO2 and hydrogen with the existence of steam.

Therefore, the percentage of CO is decreasing, while the percentage of CO2 and hydrogen is increasing respectively. In the same time, the percentage of steam is decreasing.

4. CONCLUSIONS

Basically, for 100 kmol/h of glycerol was inserted to the process and it produced about 327.37 kmol/h hydrogen. For the first reactor that is at the reforming reactor, hydrogen that produced after the optimization was 262.65 kmol/h. Optimization had been done for every reactor where by for the ATR reactor, we got 550 kmol/h of air for the highest hydrogen production and the lowest CO besides temperature around 672.9 °C. A number of important observations were noted based on the analysis of conclusions:

1. The simulation of hydrogen production plant model using autothermal reforming of methanol had been successfully developed using HYSYS 3.1.
2. The optimum A/F and S/F ratios are 5.5 and 3.5, respectively to produce hydrogen – 34.7 % (v/v), CO_2 – 60% (v/v), and CO – 0.02% (v/v).
3. With optimum parameters above, 83.6% of fuel processor efficiency was achieved.

References

Brown, L.F. A Comparative Study of Fuels for On-Board Hydrogen Production for Fuel-Cell-Powered Automobiles. International Journal of Hydrogen Energy. 26:381-397, 2001.

Bühler, W. E. Dinjus, H.J. Ederer, A. Kruse, C. Mas, Ionic reactions and pyrolysis of glycerol as competing reaction pathways in near- and supercritical water, Journal of Supercritical Fluids 22 37–53, 2002.

Hirai, T., N.O. Ikenaga, T. Miyake, and T. Suzuki. Production of hydrogen by steam reforming of glycerin on ruthenium catalyst. Energy and Fuels 9: 1761-1762, 2005.

Lenz, B. and Aicher, T.. Catalytic Autothermal Reforming of Jet Fuel. Journal of Power Sources. 149:44-52, 2005.

Prakash, C.B., "A Critical Review of Biodiesel as a Transportation Fuel in Canada", Report to Transportation System Branch, Air Pollution Prevention Directorate, R gasification Biomass to produce Hydrogen Rich Gas", Int. J. Hydrogen Energy 23, 551-557, 1998.

Rapagna, S., N. Jand and U.P. Foscolo, "Catalytic Gasification Biomass to produce Hydrogen Rich Gas", Int. J. Hydrogen Energy 23, 551-557 1998.

Sadashiv M. Swami and Martin A. Abraham, Integrated Catalytic Process for Conversion of Biomass to Hydrogen, Energy & Fuels, 20, 2616-2622, 2006.

Tyson K. S., "Biodiesel R & D", Montana Biodiesel Workshop, October 8, 2003.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Design of heat-integrated distillation processes using shortcut methods and rigorous optimization

Korbinian Kraemer[a], Andreas Harwardt[a], Wolfgang Marquardt[a]

[a]*AVT-Process Systems Engineering, RWTH Aachen University, 52056 Aachen, Germany*

Abstract

The stepwise optimization based procedure for the design of heat-integrated distillation processes as presented by Harwardt et al. (2009) is extended to include complex column designs such as the divided wall column. In the first design step, the optimal flowsheet structure of simple columns is identified using a superstructure formulation and shortcut models based on rigorous thermodynamics. The column pressure is variable in this early design stage to allow for heat integration between the column and the network of heat exchangers. In the second design step, a rigorous MINLP optimization of the most promising flowsheet structure is performed for a thorough assessment of the cost savings potential of heat integration and to determine the optimal column tray numbers and feed stage location. The rigorous optimization can be solved with good robustness, efficiency and reliability due to a continuous reformulation of the MINLP. In a third design step, it is investigated via rigorous optimization whether the energy and capital costs can be further reduced by a fully heat-integrated complex column system such as a divided wall column setup. The design procedure is illustrated by a case study considering the complete separation of a quaternary azeotropic mixture.

Keywords: rectification body method, column sequencing, heat integration, rigorous optimization, dividing wall column.

1. Introduction

The rising energy cost and the promotion of sustainable processes in chemical industry require the design of energy efficient processes. The energy requirement of a distillation train can be significantly influenced by the structure of the column sequence. Further reduction of the energy requirement is possible by heat-integrated column systems. Solving these design problems within one optimization problem, where all important design variables are optimized at the same time is possible, but very time consuming and usually performed without rigorous thermodynamic models. The systematic study of alternatives was first introduced with the concept of a separation task by Hendry and Hughes (1972). Later, Sargent and Gaminibandara (1976) introduced superstructure formulations to deal with the model complexity. Evaluation of the separation cost is performed using the Underwood shortcut method, which restricts the investigation to ideal or near ideal mixtures. In the area of column sequencing with heat integration, only limited work has been published so far. Caballero and Grossmann (2006) performed a single MINLP optimization for the entire superstructure based on the Underwood method. Dünnebier and Pantelides (1999) have investigated the rigorous optimization of thermally coupled column designs but also only considered ideal mixtures. As a consequence of the lack of powerful optimization-based design tools, the design process in industry still relies on tedious simulation studies.

Kossack et al. (2009) proposed a systematic synthesis framework for the optimal design of distillation processes composed of a sequential combination of both shortcut evaluation and rigorous optimization. With this combination of synthesis methods of increasing level of detail, separation processes for homogeneous azeotropic multicomponent mixtures can be efficiently evaluated on the basis of rigorous thermodynamics and the optimal distillation flowsheet, the optimal process operating point and the optimal column specifications can be determined.
Accordingly, in the first step of this work the optimal simple column flowsheet structure is identified using superstructure formulations and shortcut methods, also considering the possibility of heat integration between the reboilers and condensers of the distillation sequence by simultaneously optimizing column pressures in this early design stage. In the second design step, a rigorous MINLP optimization of the most promising flowsheet structure is performed for a thorough assessment of the cost savings potential by heat integration. The aim is to determine the process and column specifications that yield the lowest total annualized cost, taking into account capital and operating costs. In the third step, it is investigated whether the energy and capital costs can be further reduced by a fully heat-integrated complex column system such as a divided wall column setup.

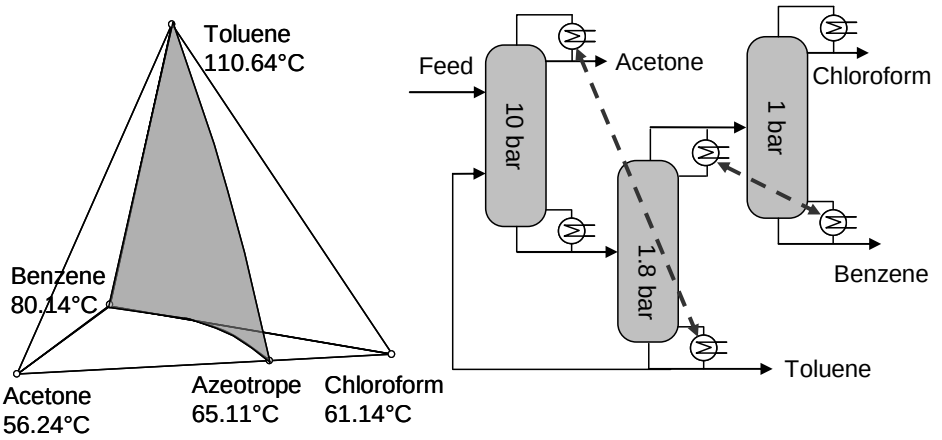


Figure 1: Distillation boundary of the example mixture at 1 bar and optimal flowsheet structure with pressure levels and heat exchange between the columns.

2. Investigation of mixture topology

In the case study, an equimolar feed (10 mol/s, 3129 kg/h) of acetone, chloroform, benzene and toluene is to be separated into its pure components in a multicolumn process. The mixture exhibits a maximum boiling binary azeotrope between acetone and chloroform (see Figure 1). The distillation boundary starts at the azeotrope and runs into the benzene/toluene/chloroform plane as already observed by Brüggemann (2005). Due to this property, acetone can be separated completely in a first distillation column for the given feed, when a recycle of toluene, benzene or a combination thereof is added. An algorithmically accessible formulation of the distillation boundary is given at minimum reflux by the so-called Pinch Distillation Boundary (PDB). The properties and the detection of the PDB have been described in detail by Brüggemann (2005). With the help of the PDB, the minimum recycle flowrate, which allows a complete

separation of acetone in the first column, can be calculated: The minimum required amounts are 5.43 mol/s for a toluene recycle, 4.54 mol/s for a benzene recycle and 4.82 mol/s for an equimolar recycle of toluene and benzene. In the following step, distillation network structures are generated and evaluated for these three alternative recycles.

3. Identification of the optimal flowsheet structure

Multicomponent mixtures can be separated by a large number of alternative distillation sequences. The identification of the optimal sequence structure of simple columns is described in detail by Harwardt et al. (2008, 2009). A brief summary is given here.
A superstructure of alternative flowsheet variants, which is given by a State Task Network representation (Sargent and Gaminibandara, 1976) of unique separation tasks, is evaluated with MILP techniques based on the objective of minimum energy requirement. The minimum energy requirements of the single columns are determined by the rectification body method (RBM, Bausa et al., 1998), a modern, pinch point based shortcut method. The RBM relies on rigorous thermodynamics and is applicable to azeotropic multicomponent mixtures, since it is a non-graphical, algorithmic method. The column pressures are variable within discretized steps in the superstructure optimization, such that the heat sinks (reboilers) and the heat sources (condensers) are combined for a further reduction of the energy demand. The optimization can be performed in just 215 seconds using the CPLEX MILP solver within the optimization platform GAMS on a 1.8 GHz, 1 GB RAM PC.
A structure with a toluene recycle was found to be optimal (c.f. Figure 1). It consists of an acetone column, where the lightest boiling pure component is separated, an indirect split in the toluene column, and the separation of the remaining binary mixture in the chloroform/benzene column. The pressure levels and the heat exchanger connections determined for optimal energy integration are indicated in Figure 1. The total minimum energy requirement is found to be 475 kW, where the equivalent sequence would require 1152 kW if no heat integration were considered.

4. Rigorous optimization

A rigorous process optimization problem is formulated next for the selected promising flowsheet structure and solved by minimizing a total cost function. The rigorous optimization step provides information about the optimal number of column trays and the optimal feed tray locations. Since these design variables are discrete variables while the energy duties, flowrates and compositions are continuous variables, a mixed-integer nonlinear optimization problem (MINLP) has to be solved. Considering the large scale and complexity of the three-column process and the nonlinearity of the underlying nonideal thermodynamics, it is obvious that this MINLP problem is particularly hard to solve. In this work, the robust and efficient solution of the MINLP problem is achieved by a favorable initialization strategy based on the results of the preceding shortcut evaluation and a reformulation as a purely continuous problem. A detailed description of the column model formulation as well as the solution algorithm can be found elsewhere (Kraemer et al., 2009). Here, a brief summary is given.
The column pressure levels and the heat exchanger connections that have been determined in the shortcut evaluation for optimal process heat integration are fixed during this rigorous optimization step. For the continuous reformulation of the MINLP problem, the discrete variables are replaced by continuous decision variables and special integrality constraints are added to the resulting NLP problem in order to force the

continuous decision variables to discrete values. For better convergence to good local optimal solutions, the integrality constraints are relaxed at first and are then successively tightened in a series of a few NLP solution steps until an integer solution is reached. Due to the excellent initialization provided by the shortcut evaluation and the continuous reformulation of the MINLP, the rigorous optimization of the large-scale example problem can be solved with excellent robustness, efficiency and reliability.

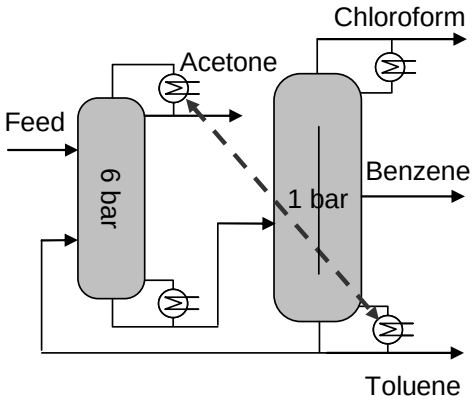
The recycle flowrate is defined as an optimization variable and the product purities are set to 99% in the rigorous optimization step. Contrary to the shortcut evaluation, the fresh feed is fed to the process as a boiling liquid at 1 bar. After being pumped to 10 bar it is preheated in an additional heat exchanger by the hot acetone product stream. The distillate stream of the toluene column that is fed into the chloroform/benzene column is not condensed but transferred as saturated vapor, thus effectively reducing the required area of the heat exchanger linking these two columns. Nonlinear cost models by Douglas (1988), updated by the M&S index, are employed for the calculation of the capital cost. A depreciation time of 5 years is considered. While low pressure steam at 3 bar (8 €/t) suffices as additional hot utility for the chloroform/benzene column, high pressure steam at 20 bar (10 €/t) is required for the acetone and the toluene columns. Maximum tray number specifications of 80, 40 and 60 for the acetone, toluene and chloroform/benzene column, respectively, yield a problem size of about 4500 continuous variables, including 440 decision variables. The optimization problem was formulated within the optimization platform GAMS 22.7 and solved by the SQP-based solver SNOPT on a PC with a 3 GHz Dual-Core CPU in 276 seconds.

Table 1: Costs, column configurations and operating point for all columns of the heat-integrated process of simple columns. Results for the same flowsheet without heat integration are shown for comparison.

	process without heat integration			**heat-integrated process**		
	acetone column	**toluene column**	**chl./benzene column**	**acetone column**	**toluene column**	**chl./benzene column**
capital cost (€/a)	110089	76819	83340	116925	76712	82355
operating cost (€/a)	55781	57696	33887	82258	12046	6345
TAC (€/a)		417612			376641	
reboiler duty (kW)	453	511	266	539	75+297	0+320
cond. duty (kW)	462	339	417	0+297	102+320	479
number of trays	51	25	40	70	28	34
feed tray	14	13	17	17	14	15
recycle feed tray	8			8		
diameter (cm)	68.6	72.5	64.3	49.7	70.7	76.5
recycle	8.27 mol/s [0, 0, 0.01, 0.99]			5.86 mol/s [0, 0, 0, 0.01, 0.99]		

The optimization results are displayed in Table 1 together with the results for the same flowsheet without heat integration, where all columns are operated at 1 bar and supplied with low pressure steam. The heat-integrated process yields a total annualized cost (TAC) of about 41000 €/a lower than the process without heat integration, which corresponds to a cost reduction of about 11%. Reboiler energy savings of 50% are observed. Note that the operating cost of the acetone column is higher for the heat-integrated process due to the expensive 20 bar hot utility. However, this cost increase is more than compensated by the energy savings for the remaining columns. The cost optimum for the chloroform/benzene column of the heat-integrated process is found at a

lower tray number when compared to the conventional process since no hot utility is required for this column. Furthermore, the heat-integrated process has a significantly lower recycle flowrate to avoid heat losses in the high pressure acetone column, where the expensive hot utility is used.



	acetone column	DWC
capital cost (€/a)	124978	167735
operating cost (€/a)	66697	13912
TAC (€/a)	373322	
reboiler duty (kW)	518	56+367
cond. duty (kW)	0+367	597
number of trays	66	61
dividing wall between trays		14-54
feed tray	16	27
recycle feed tray	8	
side stream tray		29
diameter (cm)	52.6	85.4
recycle	6.41 mol/s	

Figure 2: Flowsheet with a divided wall column (DWC) and additional heat integration between the acetone column and the DWC. Table 2: Costs, column configurations and operating point for both columns of the DWC process.

The evaluation of the energy savings potential of fully heat-integrated column designs such as divided wall columns (DWC) with shortcut methods is not a simple task, since equivalent systems of simple columns have to be constructed. In addition, the capital cost savings of integrated designs cannot be determined this way. We have however extended the rigorous optimization model to include a DWC, which could be solved robustly and efficiently with initial values from the shortcut evaluation of simple column flowsheets. Again, we reformulate the MINLP problem as a NLP problem.
The flowsheet in Figure 1 is modified to incorporate a DWC as shown in Figure 2. In addition to the energy savings of the DWC itself, heat integration is considered between the condenser of the acetone column and the reboiler of the DWC. Since the DWC operates at 1 bar, the pressure of the acetone column could be lowered to 6 bar and 10 bar steam (9 €/t) was used as hot utility for this column. As suggested by Dünnebier and Pantelides (1999) a surcharge factor of 15% is added to the capital cost of the DWC to account for the increased installation cost of the complex DWC. Despite the increased complexity of the DWC process (about 590 decision variables), the optimization could be solved in 585 seconds due to the favorable reformulation as a continuous problem.
Table 2 lists the optimization results for the DWC flowsheet (cf. Figure 2) with additional heat integration between the acetone column and the divided wall column. The energy requirement of the fully heat-integrated DWC process (574 kW) is only slightly lower than the energy requirement of the heat-integrated simple column process (614 kW). The total annualized cost of the DWC process is slightly lower as well, where the main cost saving can be allocated to the lower pressure of the acetone column.

5. Conclusions

We have presented a stepwise optimization-based design procedure for the rapid synthesis of cost optimal, heat-integrated distillation processes. Powerful shortcut methods based on nonideal thermodynamics and rigorous optimization are used in sequential design steps to obtain the optimal flowsheet structure and the optimal column specifications with outstanding robustness and computational efficiency. The possibility of heat integration is considered from the conceptual design step on such that energy savings can be optimally exploited. The benefit of our approach lies in the ability to deal with nonideal mixture behavior, the detailed column design of simple and complex columns by rigorous optimization, very fast computation times and comparably low engineering effort. With this combination of enumeration, MILP optimization and rigorous optimization, the cost of an optimal distillation process of heat-integrated simple columns for the azeotropic four component example could be determined robustly in less than 10 minutes computational time. Energy savings of 50% can be achieved. Finally, it has been studied whether the energy requirement and the total cost for the separation of the example mixture can be further lowered by applying a divided wall column (DWC) in a fully heat-integrated process. A rigorous optimization of the DWC process yields only marginal energy savings and total cost savings compared to the heat-integrated simple column process.

6. Acknowledgement

Korbinian Kraemer would like to acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG) within project MA 1188/26-1 and the Max-Buchner-Forschungsstiftung. Andreas Harwardt would like to acknowledge financial support by the Cluster of Excellence "Tailor-Made Fuels from Biomasse", which is funded by the Excellence Initiative by the German federal and state governments.

References

J. Bausa , R. Watzdorf, W. Marquardt, 1998, Shortcut methods for nonideal multicomponent distillation: 1. Simple columns, AIChE J., 44(10), 2181-2198

S. Brüggemann, 2005, Rapid screening of conceptual design alternatives for distillation processes, PhD thesis Lehrstuhl für Prozesstechnik, RWTH Aachen University, Fortschrittberichte VDI, Reihe 3, Nr. 841, VDI Verlag, Düsseldorf

J.A. Caballero, I.E. Grossmann, 2006, Structural considerations and modeling in the synthesis of heat-integrated-thermally coupled distillation sequences, Ind. Eng. Chem. Res., 45, 8454-8474

J.M. Douglas, 1988, Conceptual Design of Chemical Processes, McGraw-Hill, New York

G. Dünnebier, C.C. Pantelides, 1999, Optimal design of thermally coupled distillation columns, Ind. Eng. Chem. Res., 38, 162-176

A. Harwardt, S. Kossack, W. Marquardt, 2008: Optimal column sequencing for multicomponent mixtures, In: Bertrand Braunschweig and Xavier Joulia (Eds.): ESCAPE 18, 91-96

A. Harwardt, K. Kraemer, W. Marquardt, 2009, Design of heat-integrated distillation processes using shortcut methods and rigorous optimization, submitted to the 8th World Congress of Chemical Engineering, Montreal, Canada

J.E. Hendry, R.R. Hughes, 1972, Generating separation process flowsheets, Chem. Eng. Prog., 68(6), 71-76

S. Kossack, K. Kraemer, R. Gani, W. Marquardt, 2008, A systematic synthesis framework for extractive distillation processes. Chem. Eng. Res. Des., 86, 781

K. Kraemer, S. Kossack, W. Marquardt, 2009, Efficient optimization-based design of distillation processes for homogeneous azeotropic mixtures, submitted to Ind. Eng. Chem. Res.

R.W.H. Sargent, K. Gaminibandara, 1976, Optimum design of plate distillation columns, In L.C.W. Dixon: Optimization in Action, Academic Press, London

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling and Simulation of Nanoparticles Formation Process: A Diffusive Approach

Luciane S. Ferreira and Jorge O. Trierweiler

GIMSCOP - DEQUI - UFRGS. R. Luiz Englert, s/n°. Porto Alegre. RS. 90040-040. Brazil. Emails: {luciane,jorge}@enq.ufrgs.br

Abstract

In this work the modeling and simulation of nanoparticle formation according to the technique of nanoprecipitation was done. In this method, the particle is formed due to the further diffusion of solvent into the aqueous phase, resulting in the aggregation of the associated polymer chains. In order to predict the characteristics of the nanoparticle and also to improve the process, it was developed a mathematical model that considers: (a) the type of polymer; (b) interaction between solvent and polymer; and, (c) dynamics of solvent diffusion. The diffusivity between polymer-solvent was modeled according to the Vrentas & Duda Free Volume Theory, including the Sanchez-Lacombe equation-of-state. The model was written in terms of Partial Differential Equation, and solved with MAPLE® for a given initial size. Additionally, it is a moving boundary problem because the diffusion of the solvent out of the droplet leads to its size reduction. In this work only the diffusion behavior of one droplet was considered. The dynamic simulation shows both the evolution of the solvent inside the droplet and the variation of size in time.

Keywords: nanoprecipitation, modeling, simulation, diffusion, moving boundary.

1. Introduction

Polymeric nanoparticles are of special interest from the pharmaceutical point of view. First, they are more stable in the gastrointestinal tract than other colloidal carriers, such as liposomes, and can protect encapsulated drugs from gastrointestinal environment. Second, the use of various polymeric materials enable the modulation of physicochemical characteristics (e.g. hydrophobicity, zeta potential), drug release properties, and biological behavior (e.g. targeting, bioadhesion, improved cellular uptake) of nanoparticles. Finally, the particle surface can be modified by adsorption or chemical grafting of certain molecules such as polyethylene glycol (PEG), poloxamers, and bioactive molecules (lectins, invasins, etc.).

In this work, in order to model the nanoparticle formation only one droplet in a high diluted medium will be considered. We suppose that outside the droplet the solvent concentration is null because of the dilution and agitation. It was developed a mathematical model that takes in account: (a) the type of polymer; (b) interaction between solvent and polymer; and, (c) solvent diffusion process. After the description of the nanocapsules preparation, the mathematical model is explained. Then, the numerical simulation and its results are discussed.

2. Nanocapsules Preparation

The nanocapsules that are modeled in this work are usually prepared according to the method of nanoprecipitation (Fessi *et al.*, 1989). This method is based on the spontaneous emulsification of the organic internal phase, in which the polymer is dissolved, into the external aqueous phase. In order to allow the better understanding of the paper by the reader, the nanoprecipitation is then explained.

Nanocapsules of poly(ε-caprolactone) containing 3-benzophenon (solar protection factor) are prepared according to the following procedure: 100 mg of poly(ε-caprolactone) polymer (PCL), 76.6 mg of sorbitan monostearate, 333 mg of Mygliol 810 (caprylic/capric triglyceride) and 30 mg of Benzophenon-3 are first dissolved in acetone (27 ml). The resulting organic solution is poured in 53 ml of water containing 76.6 mg of polysorbate 80. The aqueous phase immediately turns milky with bluish opalescence as a result of the formation of nanocapsules, the wall of which is mainly constituted by PCL, and the oily core by the benzophenon-mygliol solution.

3. Mathematical Modeling

It was considered that the nanoprecipitation produces perfect spherical particles and also that each nanoparticle is originated from one droplet formed immediately after the mixing of organic phase and aqueous phases.

The major model assumptions are:

(a) There is a negligible relative velocity between the droplet and the water, the external mass transfer is approximated by diffusion. This assumption can be done based on the order of the Stokes number, which is related to the particle velocity and is defined as (Crowe, 2005, Rielly and Marquis, 2001):

$$St \equiv \frac{\tau_v}{\tau_F} = \frac{{d_p}^2 \rho_p V}{18 \mu L} \tag{1}$$

where τ_F is the characteristic time of the flow field, τ_v is the particle relaxation time, d_p is the particle diameter, ρ_p is the particle density, V is the fluid velocity, μ is the fluid viscosity, and L the characteristic dimension of the obstacle. For small Stokes numbers the particles follow the fluid motion, but for large St, the particles follow different trajectories from the fluid elements. The Stokes number was calculated for fluid velocities between 1×10^{-3} and 1 $m.s^{-1}$ (based on CFD simulations for a stirred reactor that are not showed here), and particle diameters from 100 to 4000 nm. As the Stoke number is in all cases less than 10^{-5}, the assumption of negligible relative velocity between the droplets and the external phase can be considered valid;

(b) the diffusion is one-dimensional along the radial direction; and,

(c) the diffusivity varies with time and concentration only.

Based on these assumptions, the mass balance equation for the solvent written in spherical coordinates is:

$$\frac{\partial c_1(r,t)}{\partial t} = \frac{D}{r^2}\left(\frac{\partial}{\partial r}\left(r^2\left(\frac{\partial c_1(r,t)}{\partial r}\right)\right)\right) \tag{2}$$

where: c_1 is the concentration of solvent; r is the particle radius; D is the diffusivity; and t is the time. Note that it is a moving boundary problem, since the size of the droplet reduces because of the diffusion.

As the dimensions of dependent and independent variables are not the same, a variable normalization was done including 3 new variables τ, rh, and c_{1h} defined as,

$$t_0 = \frac{r_0^2}{D},\ \{r = rh \cdot r_0,\ t = \tau \cdot t_0,\ c_1(r,t) = c_{1h}(rh,\tau) \cdot \rho_1\} \qquad (3)$$

where: r_0 is the initial droplet radius; D is the diffusivity polymer/solvent; and ρ_l is the solvent density.
After the normalization, eq. (2) is then rewritten as:

$$\frac{1}{t_0} \cdot \frac{\partial c_{1h}(rh,\tau)}{\partial \tau} = \frac{D}{r_0^{\,2} \cdot rh^2}\left(\frac{\partial}{\partial rh}\left(rh^2\left(\frac{\partial c_{1h}(rh,\tau)}{\partial rh}\right)\right)\right). \qquad (4)$$

3.1. Initial and Boundary Conditions
Assuming that the initial solvent concentration inside the droplet is uniform, the initial condition is,

$$c_{1h}(rh,0) = c_{1h0} \qquad 0 \le rh \le r_0. \qquad (5)$$

Where c_{1h0} is evaluated according to the experimental conditions.
The boundary condition at the center of the droplet (rh=0) arises from the symmetry,

$$\left.\frac{dc_{1h}}{dt}\right|_{rh=0} = 0 \qquad t \ge 0. \qquad (6)$$

Additionally, the boundary condition at the interface can written as

$$\left.\frac{dc_{1h}}{dt}\right|_{\mathrm{rh=R}} = \frac{D_{S-W}}{4\pi R^2}\left(c_{1h}(R) - c_{1h}(\infty)\right) \qquad (7)$$

in which R is the actual radius of the droplet, $D_{S\text{-}W}$ is the diffusivity of the solvent in the external phase, $c_{1h}(R)$ is the concentration of solvent at the interface and $c_{1h}(\infty)$ is the bulk concentration.

3.2. Boundary Movement
The boundary movement is calculated based on the assumption that both the mass of polymer, oil and drug remain constant during the diffusion process. The volume of the droplet is defined as

$$V_D = V_2 + V_1 + V_{oil} + V_{drug} \equiv \frac{4}{3}\pi R^3 \qquad (8)$$

where V_D is volume of the droplet, and V_1 , V_2, V_{oil} , V_{drug} are the volume of solvent, polymer, oil, and drug, respectively.
Per definition the volume of polymer and solvent are

$$V_2 = m_2/\rho_2 \qquad V_1 = (c_1/\rho_1)V_D \qquad (9)$$

with m_2 as the mass of polymer and ρ_2 as polymer density.
Substituting eq. (9) in eq. (8)and isolating for R, then the radius can be calculated as

$$R(t) = \sqrt[3]{\left[3 \cdot \left(\frac{m_2}{\rho_2} + V_{oil} + V_{drug}\right) \cdot \left(1 - \frac{c_1(t)}{\rho_1}\right)^{-1}\right] \Big/ 4\pi}. \qquad (10)$$

3.3. Model Parameters
The two main parameters of this model are the diffusivity solvent/polymer and the diffusivity solvent/water.
The experimental data presented by Wild (2003) was adjusted as a polynomial curve to describe the diffusivity of acetone in water.

$$D_{S\text{-}W} = 1\text{x}10^{-9} * (-4.737 w_S + 15.92\ w_S - 14.71\ w_S + 4.738) \qquad (11)$$

where w_s is the molar fraction of water in the external phase.

The diffusivity between polymer and solvent was modeled according to the Free Volume Theory (Vrentas and Duda, 1976, Vrentas and Duda, 1977a, Vrentas and Duda, 1977b, Vrentas and Duda, 1979). Those authors applied the Flory–Huggins thermodynamic model in their free volume diffusion theory to describe the polymer solvent enthalpic and entropic interactions. For the estimation of solvent diffusion coefficient in polymer solution systems, free-volume parameters for the both polymer and solvent must be available. The free volume (FV) diffusion model developed by Vrentas & Duda describes the solvent self-diffusion coefficient (D_1) and the polymer/solvent binary mutual diffusion coefficient (D) as given by eqs. (12) and (13), respectively.

$$D_1 = D_0 \exp\left(\frac{-E}{RT}\right) \cdot \exp\left(\frac{-\left(w_1\hat{V}_1^* + w_2\xi\hat{V}_2^*\right)}{w_1\left(\frac{K_{11}}{\gamma}\right)\left(K_{21} - T_{g1} + T\right) + w_2\left(\frac{K_{12}}{\gamma}\right)\left(K_{22} - T_{g2} + T\right)}\right) \tag{12}$$

$$D = D_1\left(1-\phi_1\right)^2\left(1-2\chi\phi_1\right) \tag{13}$$

In eq.(12) the first exponential term can be considered as the energy factor, and the second exponential term is the free-volume factor. There are 13 independent parameters to be evaluated in eq.(13). Some of them can be grouped reducing this number to the following variables: K_{11}/γ, $K_{21}-T_{g1}$, K_{12}/γ, $K_{22}-T_{g2}$, D_0, E, $\hat{V}_1^*$, ξ, $\hat{V}_2^*$, and χ, that must be determined to estimate mutual diffusivities. All of these parameters have physical significance, and therefore one must be able to evaluate every parameter from sources other than diffusion studies. The guidelines to calculate each one of these parameters, clarified by Zielinski and Duda (1992), were used in this work. Additionally the modification proposed by Wang *et al.* (2007), according to the Sanchez-Lacombe equation-of-state (SL EOS), was also taken in account. The process was considered to be isotermic and isobaric (T=298K, P=1bar and E=0). All parameters are listed in Table 1.

Table 1: Model Parameters.

$\hat{V}_1^*$	0.9695 cm^3/mol	$D_0 \times 10^4$	14.3 cm^2/s	T^*	668 K
$\hat{V}_2^*$	0.8181 cm^3/mol	K_{21}-T_{g1}	-12.12	P^*	4035 bar
δ_1	18.29 $J^{1/2}/cm^{3/2}$	(K_{11}/γ)	0.983 x 10^{-3}	T_{g2}	213 K
δ_2	20.85 $J^{1/2}/cm^{3/2}$	ρ^*	1.1427 g/cm^3		

4. Numerical Simulation

The model was simulated in MAPLE®, which is capable of finding solutions for higher order PDE and PDE systems. Based on the initial size, eq. (4) is solved. After that, the new radius can be thus calculated. If the relative difference (ΔDiam) between the new and old radius is less than $1x10^{-5}$, the process ends; if not, the iteration process goes on as illustrated in Fig. 1. As only one particle is simulated here, it was considered that the acetone concentration at the external phase is equal to zero.

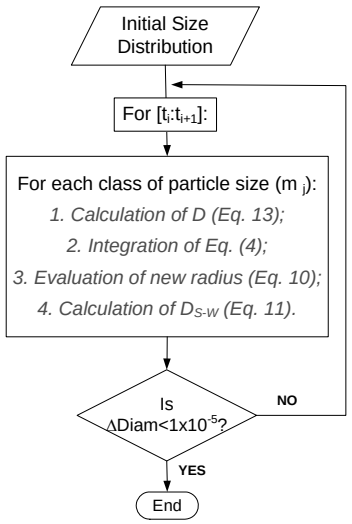


Fig. 1. Scheme of solution implemented in MAPLE®.

5. Results and Discussion

The model predicts that in about 15 ms the particle reaches the final diameter, as shown in Figure 2. This result is in qualitative agreement with what is observed experimentally, that is, as the organic phase is mixed in the aqueous phase, the suspension becomes *immediately* (at least for the human eyes) opaque, as a result of the nanoparticle formation. The evolution of particle size follows a first order kinetic profile, without any detectable delay time. Throughout the diffusion process, gradients of oil, acetone and polymer arise into the droplet leading to the reduction in size and the formation of the nanoparticle, as can be seen in Fig. 3. At the beginning, a slow reduction of acetone concentration at the interface takes place, mainly because of the initial droplet big size. As the acetone diffuses out of the droplet, the diameter reduces and then, the solvent approaches the interface at a faster rate. All these factors allow, for the last 2/3 of the precipitation, a further faster reduction of the solvent concentration at the interface. The order of magnitude of the diffusion time is in agreement with that found by Moinard-Chécot *et al.* (2008). They tried to measure the duration of the solvent diffusion step with a stopped-flow apparatus. In this experiment, only the signal corresponding to the final state could be observed, i.e., the diffusion step is less than 20 ms (the acquisition time of the apparatus).

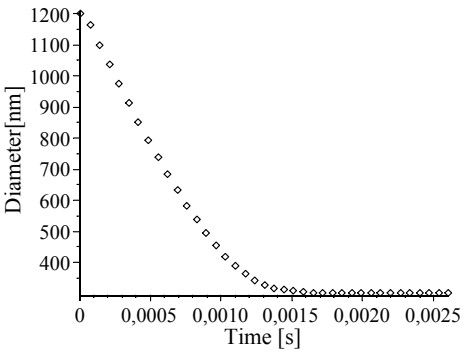


Fig. 2. Evolution of particle size.

6. Conclusions

This work shows that it is possible to obtain satisfactory results using a simplified PDE model for the nanoprecipitation. The adopted approach considers several variables that

have influence on diffusion, like type of polymer, solvent and non-solvent; affinity among them; and, polymer solubility.

As a result, it is then possible to evaluate the particle size during the nanoprecipitation, which takes no more than 20 ms.

The next step in this work is to consider an initial distribution of droplets, and apply the same methodology in order to predict the final distribution of particles. Of course, in this case, the assumption of zero concentration of acetone in the external phase is no more valid. Also, the proposed model could be combined with CFD simulator in order to improve the predictions. From the simulation, one could try to evaluate more accurately the initial droplet distribution. This methodology could thus be applied to study the influence of several configurations/sizes of reactors and mixers in the final properties of the nanoparticles.

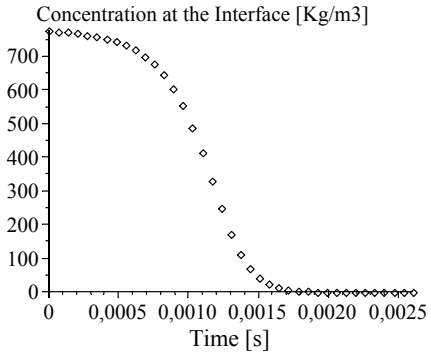


Fig. 3. Concentration at the Interface.

Acknowledgments: Prof. Wolfgang Marquardt (PT/AVT – RWTH Aachen) and CNPq/DAAD.

References

Crowe, C. T. (2005) *Multiphase Flow Handbook*, CRC Taylor & Francis.

Fessi, H., Puisieux, F., Devissaguet, J. P., Ammoury, N. and Benita, S. (1989) Nanocapsule formation by interfacial polymer deposition following solvent displacement. *International Journal of Pharmaceutics,* 55**,** R1-R4.

Moinard-Chécot, D., Chevalier, Y., Briançon, S., Beney, L. and Fessi, H. (2008) Mechanism of nanocapsules formation by the emulsion-diffusion process. *Journal of Colloid and Interface Science,* 317**,** 458-468.

Rielly, C. D. and Marquis, A. J. (2001) A particle's eye view of crystallizer fluid mechanics. *Chemical Engineering Science,* 56**,** 2475-2493.

Vrentas, J. S. and Duda, J. L. (1976) Diffusion of Small Molecules in Amorphous Polymers. *Macromolecules,* 9**,** 785-790.

Vrentas, J. S. and Duda, J. L. (1977a) Diffusion in Polymer-Solvent Systems. I. Reexamination of the Free-Volume Theory. *J. Polymer Sci., Polymer Phys.,* 15**,** 403 - 416.

Vrentas, J. S. and Duda, J. L. (1977b) Diffusion in Polymer-Solvent Systems. II. A Predictive Theory for the Dependence of Diffusion Coefficients on Temperature, Concentration, and Molecular Weight. *J. Polymer Sci., Polymer Phys.,* 15**,** 417 - 439.

Vrentas, J. S. and Duda, J. L. (1979) Molecular Diffusion in Polymer Solutions. *AIChE J.,* 25**,** 1 - 24.

Wang, B.-G., Lv, H.-L. and Yang, J.-C. (2007) Estimation of solvent diffusion coefficient in amorphous polymers using the Sanchez-Lacombe equation-of-state. *Chemical Engineering Science,* 62**,** 775-782.

Wild, A. (2003) Multicomponent Diffusion in Liquids. TU München.

Zielinski, J. M. and Duda, J. L. (1992) Predicting polymer/solvent diffusion coefficients using free-volume theory. *AIChE Journal,* 38**,** 405-415.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Software for the calculation of the exergetic efficiency in distillation columns

Cauê T. O. G. Costa[1,*], Eduardo M. Queiroz[1] and Fernando L. P. Pessoa[1]
[1] *Escola de Química – Universidade Federal do Rio de Janeiro, Av. Athos da Silveira Ramos, n. 149, Edifício Centro de Tecnologia, Bloco E, Sala 209,CEP 21941-909, Rio de Janeiro - RJ, Brazil.*
[*] *E-mail address: caue.costa@gmail.com*

Abstract

Many processes in chemical industries consume a lot of energy and often operate inefficiently. This implies high operation costs, which can be reduced through improvements in the use of energy in the process.
Distillation is an important unit operation that is widely used in industry and requires an excessive energy demand.
The exergetic analysis, which is based on the second law of thermodynamics, is an alternative to minimize this demand. Exergy or availability is the measure of the maximum amount of stream energy that can be converted into shaft work if the stream is taken to the reference state. This property may be applied to measure the quality of energy employed in a process. Therefore, the purpose of this study is to develop a program to calculate the exergetic efficiency and suggest, if necessary, possible operational changes that lead to the reduction of energy losses by irreversibilities in distillation columns.
The preliminary conclusion is that exergy analysis is vital for evaluating inefficiencies in industrial processes and thus the exergetic analysis can be used as an important tool not only for process synthesis but also for process optimization activities.

Keywords: exergy, distillation column, exergetic efficiency.

1. Introduction

Among the common utilized in industry processes, most attention was given to distillation, an important unit operation that is widely used and requires an excessive energy demand (Pessoa 2005). To evaluate the energy consumption, efficiencies based in the produced work-supplied energy ratio are often used. But not all energy can be turned into work, according to the second law of thermodynamics (Szargut et al 1988). Thus, it is important measure the availability in industrial plants. The exergetic or availability analysis can be used as alternative to minimize the irreversibility, resulting in a better consume of energy. Exergy is defined as the amount of work obtainable when some matter is brought to a state of thermodynamic equilibrium with the common components of the natural surroundings by means of reversible process, involving interactions only with the above-mentioned components of nature. This state property may be applied to measure the quality of energy employed in a process (Szargut et al 1988).
The purpose of this work is to develop a software capable of calculate the exergetic efficiency in distillation columns, working with a simulator process, and to suggests improvements for a better exergetic utilization.

2. Bibliographic Review

The exergetic analysis has been largely applied in processes of power generation, although not much applied in distillation processes (Maia 2001). This concept was introduced by Rant and is directly related with the second law of thermodynamic.
Smith et al, presenting at chapter 5 a detailed description about the second law and entropic balance. The generated entropy is proportional to the destructed exergy in a process, as shown in the law of Gouy-Stodola:

$$\delta B = T_0 \dot{S}_p$$

where δB is the exergy losses, T_0 is the reference state temperature and $\dot{S}_P$ is the entropy produced.
Szargut et al defined the concept of exergy as in the preview section and deduces the law of Gouy-Stodola abovementioned. They also show various forms of exergy calculation to different thermal, chemical and metallurgic processes; the physical, chemical and radioactive flow exergy calculation. Beyond, they show a data base for the standard exergy to dry air and sea water components, and reference reactions for calculation of the standard exergy to these reference environments.
Røsjorde et al studied the entropy production in diabatic and adiabatic separations of propylene/propane, optimizing lower entropy production, and evaluating and discussing the effects of changing operational conditions and the geometry of the column. The main conclusions are that for less generation of entropy it was necessary increase the areas of heat exchange in the reboiler and condenser and the number of trays.
Araujo et al studied the case of purification of EDC. In his work, showed ways of calculating the exergetic efficiency in different approaches and identified that most of the exergtic lost in the proposed separations occurs in the condenser.
Koeijer et al evaluted the production of entropy and exergetic losses on the experimental column and identified that the column loses less exergy when there is heat exchange in the trays.
Maia explained the concept of the reversible column and quasi-reversible column proposed by Zemp (1994). And also optimizing separations networks through the exergetic analysis for 3, 4 and 5 components systems.
Rivero et al simulated and evaluated the exergy losses in Tertiary Amyl Methyl Esther production comparing the system with the adiabatic and diabatic and measuring the Exergy Improvement Potential.

3. Algorithm development

Exergy is not conservative, opposed to energy (Szargut 1988). This thermodynamic property flows through the system boundaries by three different ways: mass flows, heat and shaft work (Rivero 2001). As previously described, to the system availability balance evaluation of the studied system, it is necessary computing the exergetic inputs and outputs of the system, the holdup factor and also the exergetic losses.
Figure 1 shows a separation system in the process simulator UniSim® Design from Honeywell. It describes a system made by a simple distillation column operating in steady-state; the exergetic inputs are given by the feed (F), and the heat added in the reboiler (E-200). The exergetic outputs are the products (D and B), and also the heat removed in the condenser (E-100). Since the described system is in steady-state, the

holdup term is null. For a more rigorous approach, it is also necessary to consider the heat lost by the column for the surroundings, which can be calculated in a system energy balance.

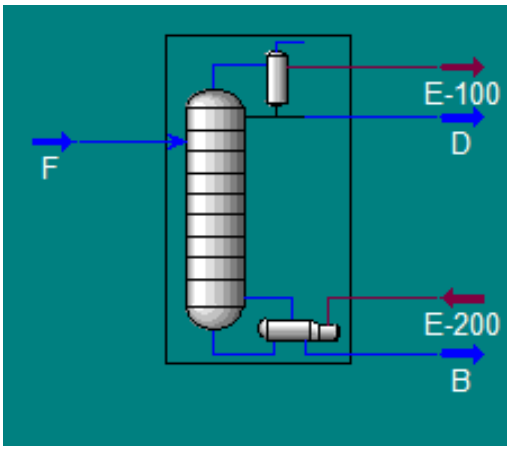


Figure 1 – Distillation column in HYSYS

To measure the exergy of stream, the procedure described by Maia, physical exergy is calculated by H – T_0*S, where H is the flow enthalpy, S is the entropy and T_0 is the reference temperature. The chemical exergy was neglected.
The calorific contribution is given by the involved heat multiplied by the efficiency of Carnot engine for the system. It is easily justified by the fact that the efficiency of this engine is the maximum thermodynamically allowed.
Thus, the exergetic balance is given by the equation:

$$\delta B = B_{feed} - B_{dist} - B_{bot} + \dot{Q}_{reb}\left(1 - \frac{T_0}{T_{reb}}\right) - \dot{Q}_{cond}\left(1 - \frac{T_0}{T_{cond}}\right) - \dot{Q}_{col}\left(1 - \frac{T_0}{T_{col}}\right)$$

where δB is the lost work, B is the exergy of a stream in kJ/h, $\dot{Q}$ is the heat of equipment or column, $\dot{Q}_{col}$ is the heat lost to environment and T_{col} is the mean temperature of the column.
The minimum work of separation W_{min} was computed and is given for the equation:

$$W_{\min} = \left(B_{dist} + B_{bot}\right) - \left(B_{feed}\right)$$

For the calculated the exergetic efficiency, used the following equation:

$$\eta = \frac{W_{\min}}{W_{\min} + \delta B}$$

The algorithm was developed in MATLAB® from The MathWorks, for use with the UniSim® Design process simulator. The flow rate, temperature, pressure, composition, enthalpy and entropy of each flow are the data input, and also the involved heat in the process heat exchangers.

4. Results and Discussion

The model was tested initially with a quaternary separation proposed by Yeomans et al. The column was feed by a current saturated liquid at 12 bar discharge of 100kmol/h with 30% propane, 30% n-butane, 20% n-pentane and 20% n-hexane.
The separation target was to recovery 98% of light and heavy key component of the feed in the top and bottom product, respectively. The model chosen was the thermodynamic equation of state of Peng-Robinson.
Columns were tested, 25 and 13 stages, with food in the trays 11 and 6. In both simulations, the values of the exergetic balance were very close to the law of Gouy-Stodola, resulting in an absolute difference in the order of 10^{-9} and relative difference of 10^{-14}. Table 1 shows the results for these simulations:

Trays/Feed	Exergy Losses	Minimum Work	Exergetic efficiency
25/11	53120 KJ/h	140930 KJ/h	72.63%
13/6	55623 KJ/h	139800 KJ/h	71.54%

Table 1: Results of first simulation of system with propane, n-butane, n-pentane and n-hexane.

It is important to point out that the lost exergy in both cases has a little difference when they are compared to the total exergy, but the difference is significative on the period of one year.
With the validated model, focuses on another system of separation, the distillation of the isomers n-butane and i-butane. The column was fed with a flow of saturated liquid 100kmol / h. The pressure was 12bar and target was 95% purity of the products. Two columns were simulated, which varied was the number of stages, 60/30 and 50/25. The results follow below:

Trays/Feed	Exergy Losses	Minimum Work	Exergetic efficiency
60/30	111500 KJ/h	123350 KJ/h	52.52%
50/20	147520 KJ/h	123570 KJ/h	45.58%

Table 2: Results of simulation of system with n-butane and i-butane.

Observe that in these simulations, the differences were significant. Due to difficulty of separation of these compounds, there is the need greater height of columns to hold it. In a difference of 10 stages of the reduction efficiency is almost 7%.
As the third simulation to evaluate the software, used a system with 50% propane and 50% n-butane. The pressure and flow of the supply current was the same also with previous simulations of saturated liquid, and a target was used for purity of 98%. To evaluated, it was tested columns with 18, 16 and 10 trays, and the feed stage, 7, 6 and 4. The results follow in the table 3:

Trays/Feed	Exergy Losses	Minimum Work	Exergetic efficiency
18/7	122470 KJ/h	170710 KJ/h	58.23%
16/6	147660 KJ/h	170670 KJ/h	53.61%
10/4	512320 KJ/h	170730 KJ/h	24.99%

Table 3: Results of simulation of system with propane and n-butane.

Similarly to the previous example, reducing the number of stages results in higher exergy losses. The system in question has a minimum theoretical stages equal to 6.5. When approaching this value, simulating with 10 stages, occurs a high exergy destruction

The results obtained using the computational program were excellent. One can seen that it is necessary more exergy for components which are more difficult to separate, which can be compensated using more trays. When the number of stages is close to the minimum number of stages there is a big reduction of efficiency.

5. Conclusions

A software was developed and the results demonstrated that it can be used as a powerful tool for process engineer in the evaluation of energetic efficiency of distillation column. Neverthe less it is necessary to use a process simulator to obtain the column data. The software was capable to identify the better process operational conditions and to decrease the exergy lost and the energy consume.

Using this program together an economic evaluation software it is possible to select the best operational conditions, optimizing in all aspects (efficiency and economic).

References

A. B. Araujo, R. P. Brito, L. S. Vasconcelos, 2007, Exergetic analysis of distillation processes—A case study, Energy Journal, 29, 1185-1193.

G. de Koeijer, R. Rivero, 2003, Entropy production and exergy loss in experimental distillation columns, Chemical Engineering Science 58,1587–1597

A. Hepbasli, 2008, A key revew on exergetic analysis and assessment of renewable energy resources for a sustainable future, Renewable and Sustainable Energy Reviews, 12, 593-661

A. P. Hinderink,F. P. J. M. Kerkhof, A. B. K. Lie, J. De Swaan Arons,H. J. Van Der Kooi, 1996, Exergy analysis with a flowsheeting simulator—I. Theory: calculating exergies of material streams. Chem Eng Sci,51,4693–700.

M. L. O. Maia, 2001, Síntese e otimização de Sistemas de Destilação utilizando a Análise Exergética, UNICAMP, doctoral thesis.

L. S. Moussa, 2001, Análise Termodinâmica de Colunas de Destilação Visando à Otimização Exergética, UNICAMP, masther thesis.

F. L. P. Pessoa, 2005, O uso eficiente de energia em colunas de destilação, UFRJ

R. Rivero, 2001, Exergy Simulations and optimization of adiabatic and diabatic binary destillation, Energy Journal, 26, 561-593.

R. Rivero, M. Garcia, J. Urquiza, 2004, Simulation, exergy analysis and application of diabatic distillation to a tertiary amyl methyl ether production unit of a crude oil refinery, Energy Journal, 29, 467-489.

A. Røsjorde, S. Kjelstrup, 2005, The second law optimal state of a diabatic binary tray distillation column, Chemical Engineering Science 60, 1199 – 1210

J. M. Smith,H. C. Van Ness, M. M. Abbott,2000, Introdução á Termodinâmica da Engenharia Química. Rio de Janeiro, LTV Editora.

J. Szargut, D. R. Morris, F. R. Steward, 1988, Exergy analysis of Thermal, Chemical and metallurgic Process.

H. Yeomans, I. E. Grossmann, 1999, Nonlinear disjunctive programming models for the synthesis of heat integrated distillation sequences, Computers and Chemical Engineering 23, 1135–1151

R. J. Zemp, 1994, Thermodynamic Analysis of Separation Systems. Inglaterra, UMIST, doctoral thesis

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Hybrid Desalination Systems: Alternative Designs of Thermal and Membrane Processes.

Marian Marcovecchio[a] , Sergio Mussati[a] , Nicolas Scenna[a] and Pio Aguirre [a]

[a]*INGAR/CONICET Instituto de Desarrollo y Diseño, Avellaneda 3657 C.P. 3000 Santa Fe, Argentina E-mail: mariangm@santafe-conicet.gov.ar*

Abstract

In this paper, desalination systems integrating thermal and membrane processes are investigated. Specifically, a hybrid desalination plant integrating reverse osmosis (RO) and multi stage flash (MSF) systems is mathematically modeled. The non linear programming problem is developed in order to optimize the configuration and operating conditions in order to satisfy fresh water demands at minimum costs.
Both implementations, with one and two reverse osmosis stages are considered at the particular framework RO-MSF adopted. The system is modeled in such a way that the RO unit receives brine feed from the MSF. In addition, partial extractions of flashing brine stream from flash chambers of MSF can be performed in order to feed the RO system. In fact, the flow-patterns and flow-rates of the brine streams of MSF are optimization variables. Heat transfer areas of pre-heaters and the geometric design of stages are design variables to be optimized. The optimization procedure will also decide between one and two RO stages and the number of modules operating in parallel at each stage. Moreover, all the operative conditions will be optimized.
Different case studies considering various seawater conditions were successfully solved without convergence difficulties. It can be concluded that the optimal arrangement of hybrid system depends strongly on the seawater conditions (salinity and temperature) and the fresh water demand as well.

Keywords: optimization, reverse osmosis, multi stage flash desalination, hybrid RO/MSF desalination systems

1. Introduction

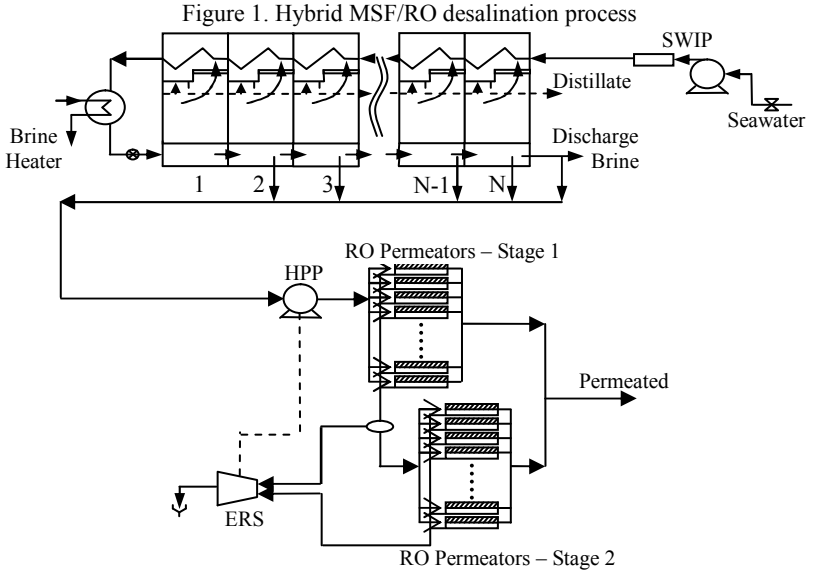
Desalination of seawater has been considered as one of the most promising techniques for supplying fresh water in the regions suffering water scarcity.
Desalination plants require significant amounts of energy in the form of heat and/or electricity. The processes used for desalination include distillation, crystallisation and ion exchange. Reverse osmosis and electrodialysis processes have emerged as alternative for thermal methods. Furthermore, hybrid systems combining thermal and membrane processes are being studied as promising options for improving the economics aspect. Particularly, the hybridization of RO and MSF plants is being investigated by many researchers (Hamed, 2005; Helal et al., 2003; Agashichev, 2004). The hybrid RO-MSF plants have potential advantages of a low power demand, improved water quality and lower running cost as compared to stand-alone RO or MSF plants. Several models have already been described in the literature to find a functional relationship between the MSF/RO hybrid process by assuming a fixed configuration (Helal et al., 2003; Marcovecchio et al., 2005).
In this paper, the particular hybrid configuration modeled by Marcovecchio et al., 2005

has been appropriately reformulated to include more integration between the processes. The possibility to select one or two RO stages and potential brine extractions from the flashing chambers are taken into account for the optimization. Precisely, the partial brine extractions may increase the integration level between the two processes, since not only the mixing temperature could be optimized by the extractions but also the salinity mixing. Moreover, the MSF model is developed by stage and component, allowing the prediction of temperature, composition and flow-rate profiles along the evaporator. Meanwhile the RO model incorporates better approximations for the resolution of the differential equations involved in mass and salt concentrations balances by implemented the finite differences method.
This paper is outlined as follows. The process is briefly described in Section 2. Section 3 introduces the problem formulation. The mathematical model is summarized in Section 4. Section 5 presents applications of the developed model and results analysis. Finally, the conclusions and future works are presented in Section 6.

2. Process description

Figure 1 schematizes the process diagram of the RO/MSF hybrid plant. As it is shown, both processes are coupled by the discharge brine of the MSF desaltor, since part of this stream is fed into the RO process while the remaining part is discharged to the sea.

Figure 1. Hybrid MSF/RO desalination process



The chemically treated seawater is fed into the MSF desaltor, then it is heated stepwise up to the brine heater inlet temperature as it flows through the condensers by recovering the heat of condensation of the vapors released in the flash stages. In the brine heater it is heated up to the brine top temperature by condensing a controlled mass of saturated heating steam coming from an external boiler at 115 °C. The heated seawater flows to the first flashing chamber via an integrated spray pipe. As the brine is superheated, spontaneous boiling takes place and vapor is released until the brine has reached its saturation temperature, the formed vapor flows through a wire mesh demister to remove the entrained brine droplets and then condenses over the condenser tubes. These processes continue right down to all next stages at decreasing temperature and pressure. The brine is flowing from stage to stage by inter-stage brine flow devices being

designed for optimum flexibility in operation and minimized stage to stage vapor leakages. The residual concentrated brine is discharged to the blowdown pump. The distillates forming in each stage are collected in an external distillate pipe and are discharged with the distillate pump.
On the other hand, the RO system involves four major parts: a High Pressure Pumps (HPP), two RO stages and the Energy Recovery System (ERS). At a stand-alone RO plant, the seawater stream is treated chemically before its pressure is raised by the high pressure pumps. Part of the power for the pumps is supply by the energy recovery system, and the rest is provided by an external source. The high pressure stream enters to the permeators of the first stage of reverse osmosis. The feed stream of the second stage, in case that it does exist, is formed with part or the total rejected stream leaving the first stage. Then, the permeated streams of both stages are blended with the product water from the MSF. The remainder rejected flow rate of the first stage, if it does exists, and the rejected flow rate of the second stage will pass through an energy recovery system, before being discharged back to the sea.
In next sections, the optimization problem is described and the hypotheses assumed are presented.

3. Definition of the problem

As was mentioned, the particular hybrid configuration studied in Marcovecchio et al., 2005 is properly extended allowing partial extractions of flashing brine stream from flash chambers of MSF that feed the RO system.
The optimization problem is defined in order to determine the best configuration and optimal operating conditions simultaneously. Flow-patterns and flow-rates of the brine streams of MSF are optimization variables. Water demands, seawater conditions (temperature and salinity), and unitary cost for capital investment and operating costs are given data. The goal is to determine the optimal configuration and operating conditions minimizing the cost per m^3 of product water.

4. Mathematical model and assumptions

The MSF mathematical model involves real-physical constraints for the evaporation process and is derived on the rigorous energy, mass and momentum balances. Physical and chemical properties of all streams are computed by specific correlations. Heat transfer areas of pre-heaters and the geometric design of stages are design variables to be optimized. Due to space limitations the mathematical model is not here presented. The complete mathematical model of the MSF process used in this paper can be found in Marcovecchio et al., 2005.
As regards the reverse osmosis process, dimensions and properties from the hollow fiber DuPont B10 modules are assumed. For these particular modules, the membranes are arranged as bundles of hollow fibers. Some hypotheses are established to formulate the model for the permeation process. It is supposed that the flow of the bulk stream inside the shell is mainly radial. Thus, null axial velocity is assumed. Axial permeated flow inside the fibers is assumed. Module parameters are supposed constant within the permeator. It is considered that the permeator have been operating for enough time to considerate stationary state. Constant fluid properties as density, viscosity and diffusion coefficient are assumed within the permeator.
Under these hypotheses, the permeation phenomenon is modelled. The concentration polarization is modeled according to the film theory. Hagen-Poiseuille equation is used to compute the pressure drop in bore side of hollow fiber. The pressure drop on the shell

side of the fiber bundle is estimated by using the Ergun equation.
An evaluation of the salt concentration of the permeated flow rate accumulated along the fiber is implemented. The flow velocities and concentrations are considered with radial and axial distribution, as well as the mass transfer coefficient.
An appropriated finite difference mesh for a hollow fiber module is implemented (Chatterjee et al., 2004), dividing the module in its axial and radial directions. The finite differences method is applied to solve the involved differential equations. The number of grid points guarantees that the approximation error for each final variable be within an acceptable range.
The complete mathematical model for the RO system implemented in this paper can be found in Marcovecchio, 2007.
Beside, the model includes the mass and energy balances on the mixers and splitters for the interconnection equipments at the hybrid configuration shown in Figure 1. Precisely, the following constraints are included in the model.

$$W_{out,j}^{MSF} = W_{inlet,j+1}^{MSF} + W_{j}^{MSF\_RO} \qquad j=1,...,N \qquad \text{(Eqn. 1)}$$

$$\sum_{j=1}^{N} W_{j}^{MSF\_RO} = W^{RO} \qquad \text{(Eqn. 2)}$$

$$\sum_{j=1}^{N} W_{j}^{MSF\_RO}\ H_{j}^{MSF\_RO} = W^{RO}\ H^{RO} \qquad \text{(Eqn. 3)}$$

where $W_{out,j}^{MSF}$ and $W_{inl,j}^{MSF}$ denote the outlet and inlet flow-rates at stage "j" of the flashing brine while $W_{j}^{MSF\_RO}$ refers to the brine flow-rate coming from stage "j" and feeding the RO system. N is the total number of flashing stages. Finally, $H_{j}^{MSF\_RO}$ and H^{RO} are the enthalpies of the corresponding streams.

It is important to note that the model was formulated in such way that the optimization procedure is able to eliminate internal flows, equipments even a whole system, and it is accurately reflected in the cost function.
The resulting mathematical model to be solved is a NLP problem involving 2533 constrains and 2350 optimization variables. The mathematical model and the solution procedure are implemented in General Algebraic Modelling System GAMS, Brooke et al., 1997 and CONOPT is used as NLP solver.
In the next section, the optimal solution of a representative case study is discussed.

5. Study Case

Many case studies have been solved by varying the seawater conditions but they are not included here due to space limitations. Indeed, a study case is presented in order to illustrate the application of the model. Table 1 lists the constants and parameters used to solve the optimization problem.
It is important to mention that the levels of fresh-water production of both processes are optimization variables.
The optimal process configuration is composed by the MSF desaltor and only one RO stage. Table 2 reports the optimal values of the obtained solution for the MSF system. Similarly, Table 3 shows the obtained solution for the RO system.
The second RO stage was not selected from the superstructure, as it can be expected. In fact, for high feed salinity the RO plants operate with only one stage at optimal

conditions.

Table 1. Parameters for the RO-MSF desalination model

Problem parameters	
Total fresh water demand, m^3/h	2000
Maximum allowed salt concentration, ppm	570
MSF-OT Evaporator	
Max. operating temperature, K	390
Sea-water temperature , K	298
Sea-water salinity, ppm	45.000
Number of stages	29
RO plant	
DuPont's B10 hollow fiber characteristics	
Pure water permeability constant, $kg/m^2.s.atm$	$4.3508955\ 10^{-5}$
Salt permeability constant, m/s	$4.0\ 10^{-9}$
Inner radius of the fiber bundle, m	$1.27\ 10^{-2}$
Outer radius of the fiber bundle, m	$5.334\ 10^{-2}$
Length of fiber bundle, m	0.75
Membrane area, m^2	152
Inner fiber radius, m	$2.1\ 10^{-5}$
Outer fiber radius, m	$5.0\ 10^{-5}$
Void fraction	0.4
Specific surface diameter, m	$1.2\ 10^{-4}$
Operating pressure range, bar	55.2-68.8
Maximum brine flow rate m^3/h	0.917

Part of the brine leaving the MSF desaltor is fed into the RO process. On the other hand, no partial brine extractions from the flashing stages resulted at optimality.
The optimum objective value is 0.8877 \$ / m^3 of fresh water. This cost can be straightly compared to the obtained for the stand-alone RO plant. In fact, under identical conditions, e.g. 45000 ppm for the seawater salinity and a total production of 2000 m^3/hr, the RO single plant produces fresh water at 0.9896 \$ / m^3 at optimality. (Marcovecchio, 2007).

Table 2. MSF optimal values

MSF desaltor	
Production [m^3/hr]	770.9
Feed [m^3/hr]	5595
Hot Utility [Gcal/h]	71.6
Total Heat Transfer Area [m^2]	30624.5
Stage Length [m]	0.6
Stage Width [m]	19.0
ΔTml per stage [K]	7.9
Brine velocity in pre-heater [m/s]	15.0
Number of tubes per stage	595

It is well known that the RO technology is recommended when a desalination plant is built, since it is the best option from an economical point of view. However, the solution obtained here shows that the integration with the MSF system results in 10.3% of saving compared to a RO single plant. Note that the salinity of the RO permeated stream exceeds the maximum allowed. This constraint is satisfied because of the

distillated stream produced by the MSF system.

Table 3. RO optimal values

RO plant	
Production [m^3/hr]	1229.1
Feed flow rate [m^3/hr]	4488.8
Feed salt concentration [ppm]	51607
Operating Pressure [atm]	67.9
Number of permeators, stage1	7107
Number of permeators, stage 2	0
Feed flow rate per module, stage 1 [m^3/hr]	0.632
Brine flow rate, stage 1 [m^3/hr]	3259.7
Permeated concentration, stage 1 [ppm]	854
Brine concentration, stage 1 [ppm]	70745

6. CONCLUSIONS.

A detailed NLP optimization model for a hybrid MSF-RO system has been developed. The hybrid configuration studied in Marcovecchio et al., 2005 has been appropriately reformulated to include more integration between both systems. At the same time, more rigorous model describing RO and MSF processes are been included. For the same problem parameters and cost items, the configuration obtained in this paper leads to a lower specific production cost of fresh water. On the other hand, authors concluded that the optimal arrangement of hybrid systems depend strongly on the seawater conditions (salinity and temperature) and the fresh water demand as well. Therefore, the influence of the main problem parameters (e.g. seawater conditions and fresh water demands) on the process efficiency will be studied in detail in future work.

ACKNOWLEDGEMENTS.

The authors acknowledge financial supports from 'Consejo Nacional de Investigaciones Científicas y Técnicas' (CONICET) and 'Agencia Nacional de Promoción Científica y Tecnológica' (ANPCyT).

References

A. Hamed, 2005, Overview of hybrid desalination systems - current status and future prospects, Desalination, 186, 207-214.

A.M. Helal, A.M. El-Nashar, E. Al-Katheeri, E. Al-Malek, 2003, Optimal design of hybrid RO/MSF desalination plants. Part I: Modeling and algorithms, Desalination, 154, 43-66.

S. P. Agashichev, 2004, Analysis of integrated co-generative schemes including MSF, RO and power generating systems (present value of expenses and "levelised" cost of water), Desalination, 164, 281-302.

M.G. Marcovecchio, S.F. Mussati, P.A. Aguirre, N.J. Scenna, 2005, Optimization of hybrid desalination processes including multi stage flash and reverse osmosis systems, Desalination, 182, 111-122.

A. Chatterjee, A. Ahluwalia, S. Senthilmurugan and S.K. Gupta, 2004, Modeling of a radial flow hollow fiber module and estimation of model parameters using numerical techniques, Journal of Membrane Science, 236, 1-16.

M.G. Marcovecchio, 2007, Process modeling and global optimization methods applied to the synthesis of desalination processes, Ph.D. In Engineering. Mention: Computational Mechanics. FICH, Universidad Nacional del Litoral, Santa Fe. Dissertation.

A. Brooke, D. Kendrick, A. Meeraus, R. Raman, R., 1997, GAMS Language Guide, Release 2.25, Version 92. GAMS Development Corporation.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Feasibility of separation of ternary mixtures by pressure swing batch distillation

Gabor Modla, Arpad Kopasz, Peter Lang

BUTE Department of Building Services and Process Engineering, Muegyetem rkp. 3-5 H-1521 Budapest, Hungary, E-mail:lang@mail.bme.hu

Abstract

Feasibility of the pressure swing batch distillation separation of ternary homoazeotropic mixtures in different single and double column configurations is investigated by assuming maximal separation. Feasibility regions where the ternary mixture can be separated into its pure components (by applying in at least one step pressure swing) and the separation steps for different configurations are determined. The method is presented for the most frequent types of ternary mixtures with minimum azeotrope(s).

Keywords: Pressure Swing, Batch Distillation, Azeotropes

1. Introduction

Azeotropic mixtures cannot be separated into their components by traditional distillation. Several mixtures form a homoazeotrope, whose position can be shifted substantially by changing system pressure (pressure sensitive azeotropes). These mixtures can be separated by pressure swing distillation (PSD). Knapp et al. (1992) studied the use of entrainers for continuous PSD. Hence such azeotropes can be separated by PSD which are not pressure sensitive if the entrainer forms a pressure sensitive boundary and the number of mixtures separable by PSD can be increased. Phimister and Seider (2000) studied the separation of a minimum azeotrope by semi-continuous PSD. Repke et al. (2007) investigated experimentally the application of PSD in batch. They studied the separation of a minimum azeotrope (acetonitrile-water) by PSD in a batch rectifier (BR) and in a batch stripper (BS). Modla and Lang (2008) compared different batch configurations for the separation binary (maximum and minimum) homoazeotropes by PSD. The best results were obtained with the two new double column configurations (double column batch rectifier (DCBR) for maximum and double column batch stripper (DCBS) for minimum azeotropes) equipped with a common bottom or top vessel, respectively.
Modla et al. (2008) stated that by the methods of classification of Serafimov (1970) and Matsuyama and Nishimura (M&N, 1977)) the pressure sensitivity of an azeotrope is not indicated though this information is indispensable for the synthesis of PSD separations. They extended the method of M&N by indicating the pressure sensitivity of an azeotrope by writing ‘P’ after the M&N number of an azeotrope if it is pressure sensitive. (They also mentioned that if the type of RCM varies (an azeotrope disappears or the volatility order of components varies with the variation of pressure), it must be given for both pressures.)

The aim of our work is to investigate the feasibility of the batch PSD separation of ternary mixtures in different (single and double) configurations by assuming maximal separation, to determine feasibility regions where the ternary mixture can be separated

into its pure components (by applying in at least one step pressure swing) and to determine separation steps for ternary azeotropic mixtures of different type.

2. Feasibility method

When making feasibility studies we suppose that maximal (perfect) separation (Bernot et. al. 1990) can be produced. This involves the following assumptions: 1) large number of stages (*N*), 2) large reflux/reboil ratio, 3) negligible liquid plate hold-up, 4) negligible vapour hold-up.

The method is based on the determination of the feasible compositions of products (continuously withdrawn) and those of residues (remaining in the vessel) by analysing the *vessel path*s in the residue curve maps at the two different pressures (P^1, P^2).

For the different column configurations separation steps and *feasibility region(s)* of the separation, which means all feed compositions, from where all components can be purely recovered by maximal separation at the given pressure or by applying pressure swing, are determined.

In order to determine the sequence of the cuts (products) for a given charge composition by maximal separation we have to study the batch distillation (BD) regions. From any point of a BD region we get the same cuts (fractions) by rectification. The method to divide the composition diagram into BD regions can be summarised as follows:

1. First the composition triangle is divided into regions containing one unstable node.
2. In each of these regions the other batch distillation boundaries are found by joining the unstable node to all the saddles and nodes of its region.

Since we make feasibility studies not only for BRs but also for batch strippers we have to define also the batch stripping (BS) region: from any point of a BS region we get the same cuts by stripping. In the case of strippers the maximal separation means large reboil ratio (and *N*). In order to distinguish better the BD regions (used for rectification) and BS ones in this paper the BD regions will be called batch rectification (BR) regions.

In a ternary mixture containing pressure sensitive azeotrope(s) the *location of the boundaries* involving this (these) azeotrope(s) varies. (We suppose that the boundaries are straight lines.)By changing the pressure we can cross these boundaries. Between the boundaries concerning to the two different pressures there is a region from where at the different pressures different products (distillate in a BR, bottoms in a BS) can be obtained.)

3. Feasibility results

Results will be shown for the two most frequent Serafimov's types of mixture, only. Sketch of the residue curve map (RCM), simple distillation, BR/BS regions, separation steps and vessel-paths for the recommended configuration(s) are presented.

3.1. Separation of mixtures of class 1.0-1a

The classification of original residue curve map is 1.0-1a (21.6%) by Serifamov (S) and it is 1-0-0 by N&M.

The azeotrope *L-I* is the unstable, vertex *H* is the stable node, respectively (*Fig. 1a*). Vertices *L* and *I* are saddles. The RCM (M&N: 1P-0-0, e.g. CCl4(*L*)-ethanol(*I*)-acrylic acid(*H*)) contains only one simple distillation (the whole triangle) but two BR and BS (maximal separation) regions. For this type of mixture the BR and BS regions coincide but the sequence of the products (cuts) is just the opposite as it is indicated in *Table 1* where sequence of the vertices of the different regions (triangles) corresponds to that of the cuts.

The location of the boundary (line between Az_{LI} and vertex *H*) considerably varies and so the areas of the BR and BS regions change with the pressure.
Configurations recommended: batch stripper or double column batch stripper.
The separation steps are as follows:
1. Production of *H*:
In the case of DCBS only one column is operated; at the end of this step the vessel path (*Fig. 1b*) arrives at the *LI* edge. This step is feasible at either of the two pressures but if the charge composition is located
- in the area I and P^2 is applied or
- in the area II and P^1 is applied
there is no need for pressure change at the end of this step.
2. Production of *L* and *I* by pressure swing distillation (in the case of DCBS both columns are operated).
Remarks:
-The separation of the ternary mixture can be reduced to that of binary systems.
-This type of ternary mixture cannot be separated in one operation step by double column system.
The feasible region is the whole area of the triangle.

Region	I.	II.
BR	Az_{LI}-*I*-*H*	Az_{LI}-*L*-*H*
BS	*H*-*I*-Az_{LI}	*H*-*L*-Az_{LI}

Table 1. BR and BS regions for a mixture 1-0-0

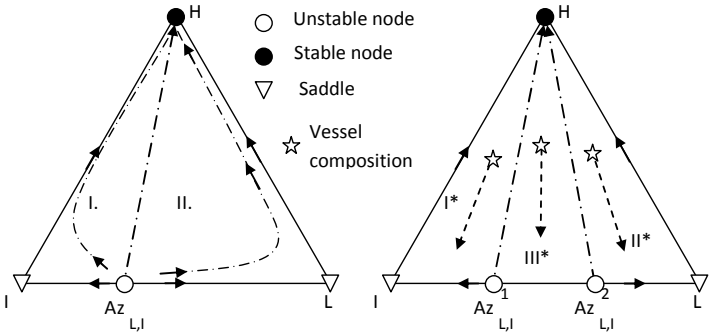


Fig. 1. Sketch of the RCM (a) and vessel paths of a BS for a mixture 1P-0-0

3.2. Separation of mixtures 2.0-2b
The class of original residue curve map is 2.0-2b (21%) by Serifamov and it can be 1-2-0 or 1-0-2 or 0-2-1 by N&M.
The azeotrope *I*-*H* and vertex *L* are saddles, vertices *I* and *H* are stable nodes and the azeotrope *L*-*I* is the unstable node, respectively (*Fig. 2*). The RCM contains two simple distillation and three BR/BS regions (*Table 2*).

If only one azeotrope is pressure sensitive it can be
- the unstable node, (by extended classification of RCM: 1P-2-0 (Case1 e.g. acetone(*L*)-methanol(*I*)-benzene(*H*)) or 2-1P-0 or 0-2-1P) or
- a saddle (by extended classification of RCM: 1-2P-0 (e.g. methanol(*L*)-toluene(*I*)-n-butanol(*H*)) 1-0-2P, 0-2P-1; not shown).
The other possibility is that both azeotropes are pressure sensitive (Case 2).

Region	I.	II.
BR	Az_{LI}-*I*-*H*	Az_{LI}-*L*-*H*
BS	*H*-*I*-Az_{LI}	*H*-*L*-Az_{LI}

Table 2. BR and BS regions for a mixture 1-0-0

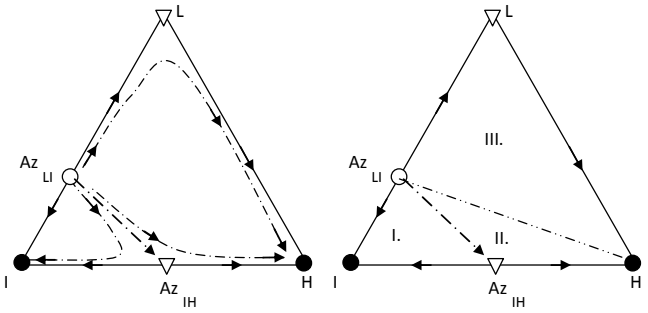
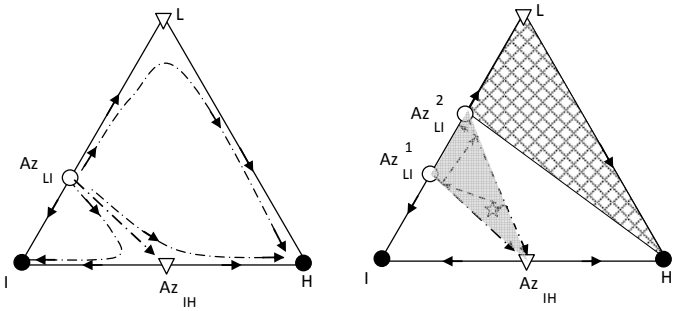


Fig. 2. Sketch of RCM (a) mixture and BR/BS regions (b) for a 1-2-0 mixture (2.0-2b)

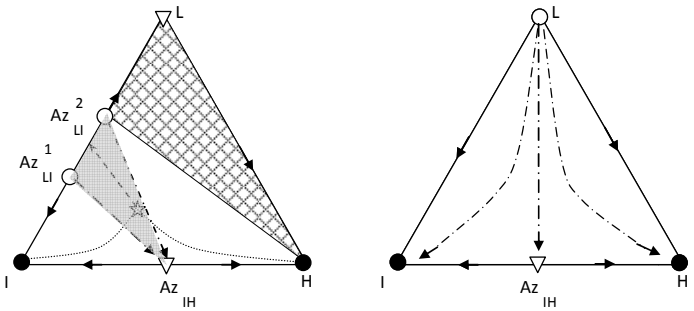
3.2.1. Case 1 – Only one azeotrope is pressure sensitive

The pressure sensitive azeotrope is the unstable node (e.g. Az_{LI}, *Figs. 3*)

The location of the boundary (line between Az_{IH} and Az_{LI}) considerably varies with the pressure (due to the movement of Az_{LI}) .



a. Sketch of RCM at P1 *b. Vessel path for the PS-BS*



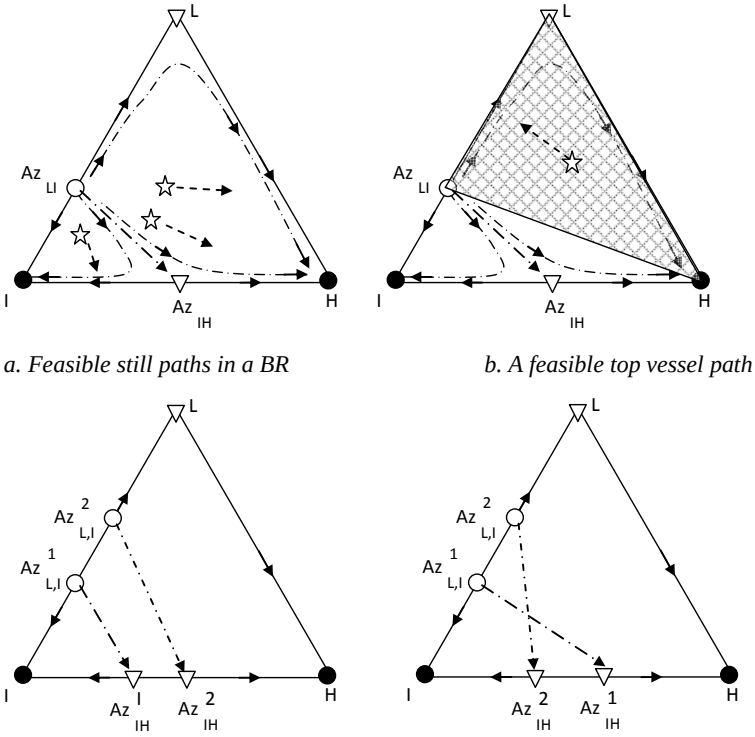
c. Vessel path for the DCBS *d. Sketch of RCM at P3*

Fig.3. Separation of 1P-2-0 mixtures

The configurations recommended: batch stripper (*Fig. 3b*) or double column batch stripper (*Fig. 3c*).
The separation method contains the following steps:
0. (optional). Production of some *I* or *H*: in the case of DCBS by operating only one column, if the charge composition is out of the darkened triangle.
1. Production of *H* or *I* alternately in a BS, (vessel path (dashed) in *Fig. 3b*) or simultaneously in a DCBS (*Fig. 4c* the liquid concentration profiles are dotted)). (In the case of BS this step consists of several parts and requires several changes of *P*.)
2. Production of *L* and *I* by PSD.
If the charge composition is located in the BS Region III (hatched area for P^2 in *Fig. 3b*) in the first production step the whole amount of *H* can be withdrawn as product.
It must be still noted that with the variation of the pressure (e.g. at P^3) the azeotrope *L-I* can even disappear (*Fig. 3d*). In this case the topology of the residue curve map varies (to Class S: 1.0-2, M&N: 0-2-0). The recommended configurations are the same as above. In this case the separation *L/I* (at P^3) does not require pressure swing.
The feasible region is the whole area of the triangle.

3.2.2. Case 2 – Both azeotropes are pressure sensitive
In this case there are two boundaries whose location varies with the pressure. We have to focus on the variation of the concentration of that component (e.g. component *I* in *Fig. 4*) which is present in both azeotropes. Two different cases are possible. The concentration of that component in the two azeotropes can vary with the pressure
- in the same direction (increases or decreases (*Fig. 4c*) in both azeotropes, e.g. acetone(*L*)-methanol(*I*)-ethyl-acetate(*H*))
- in the opposite direction when the boundaries intersect.
Recommended separation methods:
1. Removal of Az_{LI} by BR which is pressure sensitive so it can be separated by pressure swing BS or DCBS (*Fig. 4a*). The residue can be binary mixture *I-H* or *L-H*. The mixture *L-H* is zeotropic so it can be separated by both BR and BS while the mixture I-H can be separated by pressure swing in BS or DCBS.
2. If the charge composition is located in the BS region III (the hatched area in *Fig. 4b*) the first step is production of *H* in BS. In this case the remaining binary mixture *L-I* which can be separated by pressure swing in BS or DCBS.
The feasible region is the whole area of the triangle.

The most frequent types of ternary mixtures were investigated. We studied 5 types of mixture by the classification of Serafimov corresponding to 11 different types of mixture by the classification of Nishimura and Matsuyama (N&M), and 41 different cases by the extended N&M classification taking into account the pressure sensitivity of the azeotropes, as well. From these 41 cases studied
-in 14 cases the whole area of the composition triangle proves feasible,
-in 15 cases in the triangle there is at least one region where the separation is feasible,
-in 12 cases there is no feasible area within the interior of the triangle.
We stated that, the composition triangle contains a feasible part only if
- both the azeotrope which is unstable node and
- the separatrice(s) between stable node(s) and saddle(s)
are pressure sensitive.
Moreover we concluded that if the cuts of a BR/BS region are only pure component(s) and/or pressure sensitive binary azeotrope(s) this region is feasible.



a. Feasible still paths in a BR *b. A feasible top vessel path in a BS*

c. Movement of the boundaries with P in the same direction *d. in the opposite direction*

Fig. 4. Separation of 1P-2P-0 mixtures:

Acknowledgement

This paper was supported by the Janos Bolyai Research Scholarship of the Hungarian Academy of Sciences and the Hungarian Research Funds (OTKA) (No: T-049184).

References

Bernot Ch., M. Doherty, M. F. Malone, (1990). Patterns of Composition Change in Multicomponent Batch Distillation, Chem. Eng. Sci., 45, 1207

Knapp J.P. and M.F. Doherty, (1992). A new pressure swing-distillation process for separating homogeneous azeotropic mixtures, Ind. Eng. Chem. Res., 31, 346-357.

Matsuyama, H., H. Nishimura, (1977). "Topological and Thermodynamic Classification of Ternary Vapor-Liquid Equilibria", *J. Chem. Eng. Japan,* **10,** 181.

Modla G. and Lang P. (2008). Feasibility of new pressure swing batch distillation methods, Chem. Eng. Sci., 63 (11) 2856-2874.

Modla G., P. Lang and A. Kopasz, (2008) Entrainer selection for pressure swing batch distillation, *ESCAPE-18, in CD*, ISBN 978-0-444-53228-2.

Phimister, J.R.; Seider, W.D. (2000). Semicontinuous, Pressure Swing Distillation, Ind. Eng.Chem. Res., 39, 122-130.

Repke .J. U., Klein A., Bogle D., Wozny G., (2007). Pressure Swing Batch Distillation for Homogenous Azeotropic Separation, ChERD 85 (4) 152, 492-501

Serafimov, L.A. (1970). The Azeotropic Rule and the Classification of Multicomponent Mixtures VII. Diagrams for Ternary Mixtures. Russ. J. Phys. Chem. 44(4), 567-571

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Published by Elsevier B.V.

Activity and Information Infrastructure for Risk-Based Process Design

Yasunori Kikuchi,[a*] Masahiko Hirao [a]

a The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
**kikuchi@pse.t.u-tokyo.ac.jp*

Abstract

This study presents a set of models of activity and information infrastructure required for practical risk-based process design. In this process design, local risk and global impact are evaluated and interpreted comprehensively by integrating risk assessment (RA) and life cycle assessment (LCA). The type-zero method of Integrated DEFinition language, or IDEF0 and the Unified Modeling Language for information system, or UML, were applied for enabling systematic and effective definitions of requirements for process designers and supporting information infrastructure. In this study, a case study is performed on metal cleaning process, where various chemical substances have been utilized. With the description of developed IDEF0 and UML models, risk-based process design involving on-site engineers was proposed.

Keywords: life cycle assessment, risk assessment, activity and information modeling

1. Introduction

Global environmental impact and local risk due to the use of chemicals have become issues in decision making in process management, which covers planning, design and operation phases, and process improvement of production processes. A decision making in industrial process design is imposed to various constraints. Process managements should focus on the way how to apply and what to be utilized for process chemicals or raw materials under such constraints. For implementing global impact and local risk into such process management, risk assessment (RA) and life cycle assessment (LCA) have been key technologies to quantify such non-monetary evaluation indices. RA can consider the critical concerns for decision makers in industry, such as occupational health, and safety issues. Based on LCA results, decision makers can identify the life cycle impacts associated with their decisions. For risk-based decision making, integrated application and interpretation of RA and LCA has been discussed with actual case study (Kikuchi and Hirao 2008).

Business modeling approach is useful to activate a smooth implementation of new business activities (Naka 2006). Because environmentally-conscious design of processes needs systematically connected activities and information in process evaluation, simulation and optimization (Chen and Shonnard 2004), effective information technology support is important for process engineers (Shneider and Marquardt 2002).

In this paper, plant-specific risk and life cycle assessments are integrated with sustainable process design. This integration is visualized as business and information models. Activity model represents a method of integrating RA and LCA for a risk-based

decision making in industrial process management. Required information infrastructure is visualized by information model. A case study is performed on the design of metal cleaning process, which has been a big issue in Japan on environmental managements by small and medium-sized enterprises (SMEs).

2. Activity and information infrastructure

2.1. Process design and evaluation knowledge

Chemical process should be continuously improved on global impact and local risk due to the change of the needs from society on such non-monetary issues. Fig.1 shows process improvement procedure with required knowledge. All knowledge for process design can be divided into five knowledge units: mechanisms of risk specification (RSM), evaluation (EvM), alternative generation (AGM), process simulation (PSM), and risk interpretation (RIM). Knowledge has individual roles of performing the phases of process improvement. As well as the single role of knowledge, the combination among them was addressed, and activities were visualized with appropriate knowledge to be referred. The systematization of the five units of knowledge enables the continuous improvement on site according to the change of chemical risk concerns.

2.2. Activity and information modeling

The type-zero method of integrated definition language or IDEF0 have widely been used for breaking down all administrations and operation procedures into “activities” and systematic relationships among them are described by ICOM: Input, Control, Output, and Mechanism. Each activity can be hierarchically decomposed into subactivities. A developed model can make it visible what knowledge, technique and information are actually required for certain activity. In process design field, several authors applied the activity and information modeling approach to integrate new or existing engineering methods and tools for environmental protection and EHS risks (Gabber et al. 2004, Sugiyama et al. 2008). Unified Modeling Language (UML) for systems modeling has a large potential to visualize clearly system/data requirements for software tools from design stages. Several researches on risk-based decision making have been developed in a couple of decades. They could have solved actual problems based on tactful application of environmental management methodologies by researchers. To enhance the efficiency and practicability of risk-based decision, such research achievements should be generalized and shared systematically enough to implement them into other fields and sectors.

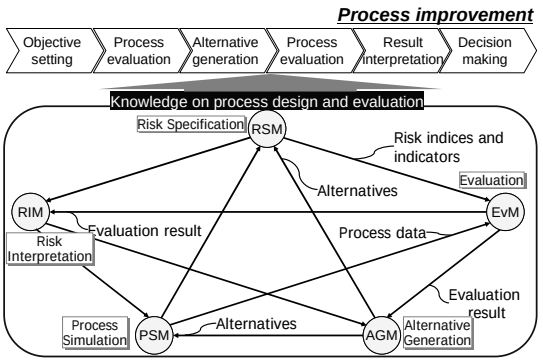


Figure 1 Knowledge structure for process design and evaluation

Fig. 2 shows the process evaluation activities utilizing different assessment methodologies. The evaluation indices must be specified on the industry-specific conditions on chemical usage and the social sense of value, which could be formalized into the constraints on alternative process and design. Based on the specified evaluation indices, the sufficient assessment methodologies are selected and executed in the centered activities in Fig.2. The information required assessments can be shared in the last activity "Collect information". The first and the last activities enable the effective settings of evaluation objectives and execution for avoiding redundant procedures during evaluation. Fig. 3 shows one of the developed diagrams, use case diagram of evaluation tool. This diagram describes a snapshot of processing in system with its user. By developing the software system of performing RA and LCA on individual conditions, the applicability and practicability of assessment methodologies for non-monetary issues can be enhanced significantly.

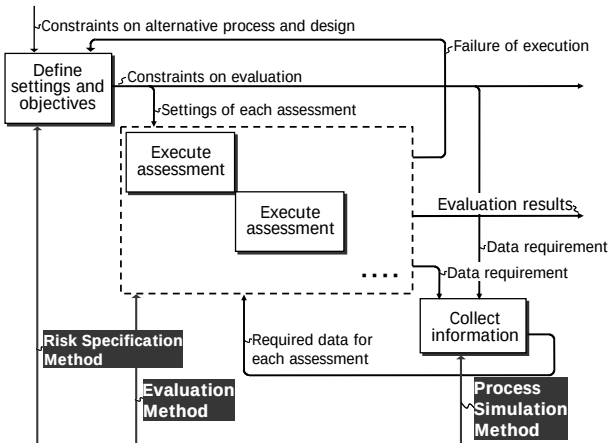


Figure 2 Process evaluation activities with information infrastructure

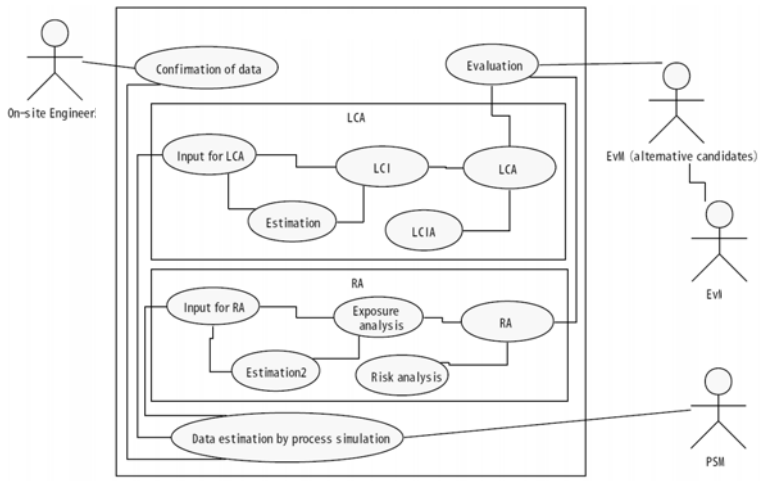


Figure 3 UML use case diagram for information infrastructure of implementing LCA and RA into process design by on-site engineers

3. Case study of process design in metal cleaning

Actual case study of enhancing the practicability of RA and LCA was addressed in the design of metal cleaning process. Metal surface treatments are inevitable processes for any kinds of products including metal parts, which are electroplated and surface converted. A metal cleaning process is important for the quality of such metal parts treatment. For supporting risk-based process design in metal cleaning process, the developed activities model was applied and the requirements of actual information infrastructure were defined. The evaluation procedures are based on the existing researches (Kikuchi and Hirao 2007, 2008), where RA and LCA are performed simultaneously to design an economically viable process on the basis of local risk and global impact.

Fig. 4 shows the developed UML sequence diagram for RA in metal cleaning process design. This diagram illustrates the temporal sequence among existing domains inside of software program. The domains include graphical user interface (GUI) and databases assembled in the software tool. The exposure and effect analyses in RA can be supported by the software system illustrated in Fig. 4. According to the sequence, the databases storing the background exposure and the hazardous properties of targeted chemical substances should be available. The connection with existing risk assessment tools is also required for effective RA. Fig. 5 shows the UML sequence diagram for LCA in metal cleaning process. The main steps in LCA are the life cycle inventory analysis and life cycle impact assessment. In these steps, a wide range of information is required such as process inventories in the production plants of inputted materials. Databases storing such information should be available. In addition to the databases required for both RA and LCA, the calculation modules for the steps of assessments should be able to accessible in the system. The temporal sequences of such modules were developed by UML sequence diagrams as well.

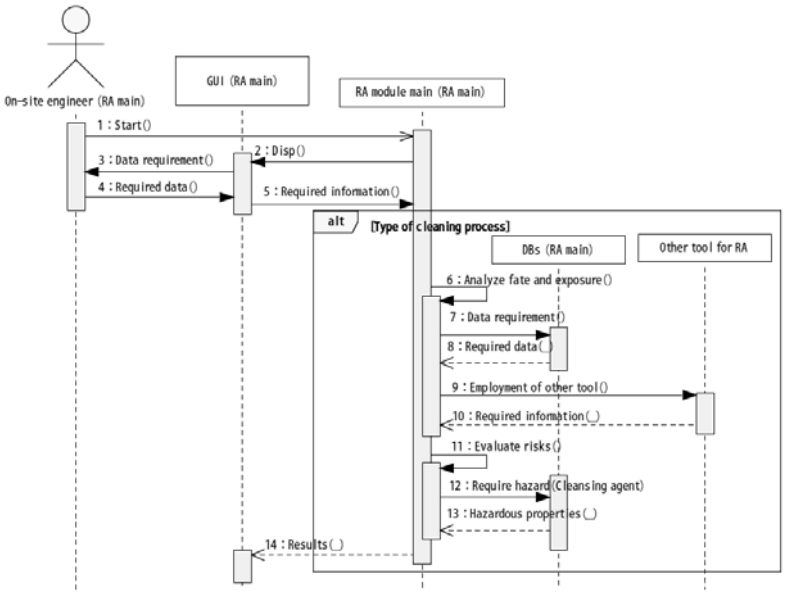


Figure 4 UML sequence diagram for risk assessment in metal cleaning process design

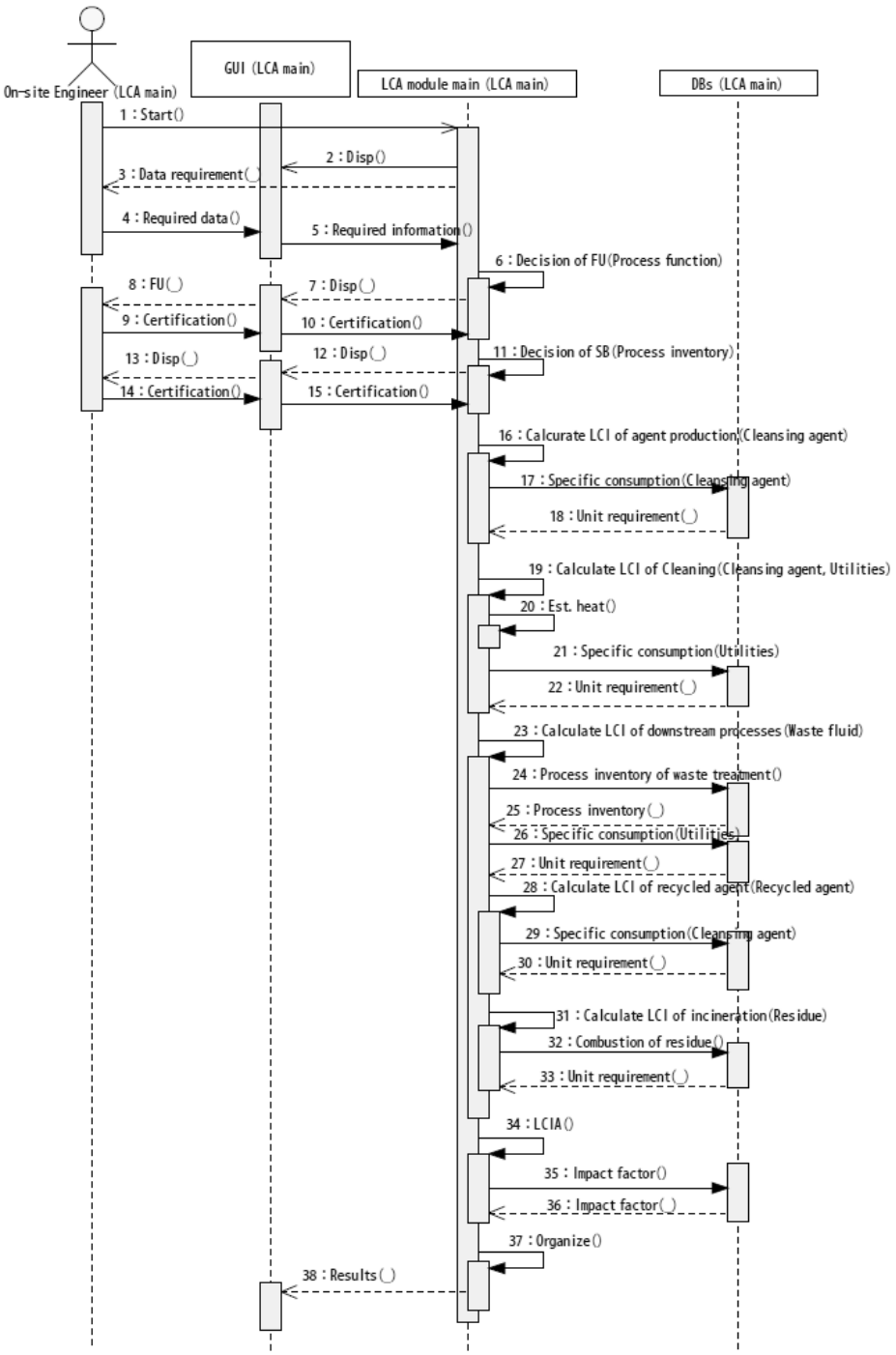


Figure 5 UML sequence diagram for life cycle assessment in metal cleaning process

The effectiveness of software tool to implement RA and LCA into practice was verified through the development of actual software tool of designing metal cleaning process on Microsoft ® Excel. Based on the software system, the on-site engineers became able to

perform RA and LCA only by their own process information, because the system implements all other required information with calculation modules.

4. Conclusion

This study aims to enhance the applicability and practicability of developed engineering knowledge on risk-based process design. The knowledge on process design and evaluation is classified and categorized into five different units, the function of which was defined to establish logical connection. Especially on assessment methodologies for non-monetary issues, the actual procedures and information infrastructure of them was visualized in clearly and logically-practical styles in business process by IDEF0 activity modeling language. In addition to the business activity modeling, the required information system for implementing the assessments into practice was designed by UML diagrams, which can visualize the system requirements. Actual case study on the design of metal cleaning process was performed to show the effectiveness of software system to enhance the applicability of RA and LCA. The databases and calculation modules to be implemented in software system could be defined by UML based on the business model defined by IDEF0. Actual software system verified the enhancement of the applicability of RA and LCA by system development.

Business activity and supporting software system models can stimulate a fast and appropriate implementation of structured knowledge on process design and evaluation. Implemented system enables the continuous improvement of processes, which had been designed correctly on the previous constraints. The visualization of such activities and information enables the effective discussion on scenario-based selection of engineering and scientific knowledge for improving decision making toward sustainable future.

Acknowledgement

The part of this study was supported by the Japan Society for the Promotion of Science through a Grant-in-Aid for Research Fellowships (No. 1811314) and Scientific Research (B) (No. 18310052).

References

H. Chen, D. R. Shonnard, 2004, Systematic Framework for Environmental Conscious Chemical Process Design: Early and Detailed Design Stages, Industrial & Engineering Chemistry Research, 43, pp.525-552

H. A. Gabber, A. Aoyama, Y. Naka, 2004, Model-Based Computer-Aided Design Environment for Operational Design, Computers & Industrial Engineering, 46, pp.413–430

Y. Kikuchi, M. Hirao, 2007, Integrated Design of Process and Operation Considering Local Risks and Global Impacts: A Case Study on Metal-Degreasing Process Design, Proceeding of 17th ESCAPE, pp.1223-1228

Y. Kikuchi, M. Hirao, 2008, Practical Method of Assessing Local and Global Impacts for Risk-Based Decision Making: A Case Study of Metal Degreasing Process, Environmental Science & Technology, 42, (12), pp.4527-4533

Y. Naka, 2006, Introduction to Integration Engineering (original title in Japanese), Kogyo Chosakai, Tokyo

R. Schneider, W. Marquardt, 2002, Information Technology Support in the Chemical Process Design Life Cycle, Chemical Engineering Science, 57, pp.1763-1792

H. Sugiyama, M. Hirao, U. Fischer, and K. Hungerbühler, 2008, Activity Modeling for Integrating Environmental, Health and Safety (EHS) Consideration as a New Element in Industrial Chemical Process Design, Journal of Chemical Engineering of Japan, 41, (9), pp. 884-897

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling and simulation the fixed bed column extraction of inulin from garlic (*Allium sativum L.* var. Chonan)

Raquel M. Galante[a], Marintho B. Quadri[a*], Ricardo F. Machado[a], Mara G. N. Quadri[a], Leandro Osmar Werle[a]

[a]*Santa Catarina Federal University, Chemical Engineering Department. University Campus, Mail Box: 476, Florianópolis-SC, Brazil, Zip Code: 88010-970 Fax: 55-48-3721-9554 *Email: m- quadri@enq.ufsc.br*

Abstract

Inulin is a fructan found in many plants as a reserve carbohydrate. It is an ingredient that integrates the group of "functional foods" improving physiological function, preventing diseases, replacing fat and sugar, and acting as a stabilizer. This work carried out the modeling and experimental data simulation obtained in the inulin extraction process from garlic in continuous flow. The extraction system comprehends a column of 60cm^{3}filled with garlic and glass pearls. The study was conducted in flow rate of 0.59, 2.4, and 4.7 cm^3/min at 60°C with single passage of solvent and recirculation. For all configurations studied, the model described satisfactorily the inulin extraction tendency. Were obtained low values of Peclet number, indicating the importance of the bed dispersive nature.

Keywords: extraction, simulation, inulin

1. Introduction

Inulin is an important carbohydrate resource in the plants. It is a fructan synthesized by a wide range of plants, approximately 36,000 species, representing 10 families (Roberfroid, 1993).

By presenting a very high level of fructose (around 94%), the inulin has excellent prospects for obtaining pure fructose syrup. Their functional properties, among them the low calorific value, contribution to improving the balance and mineral factor "bifid", which is the regulation of the intestinal flora. Inulin was recently identified as an ingredient to replace fat or sugar, but commercial production in Brazil does not yet exist (Narinder et al., 2002; Gibson et al., 1993).

In this work was carried out the modeling and experimental data simulation obtained by the extraction process in fixed bed column. This work has great importance, because it has the ability to predict the system behavior, besides facilitating the improvements in process implementation, allowing more efficient design of extraction systems.

2. Methodology

2.1. Experimental Procedure

The extraction system in continuous flow was composed of a cylindrical glass column with internal capacity of 60 cm^3, filled with sliced garlic and glass pearls in the proportion of 1:2 (inert:garlic) in mass (Figure 1).

The inulin extraction study in continuous process was carried out with a single step and solvent recirculation, at flow rate of 0.59, 2.4 and 4.7 cm^3/min for both cases. In all conditions study were accomplished the inulin extraction kinetics and sugars exhaustion from garlic to the equilibrium.

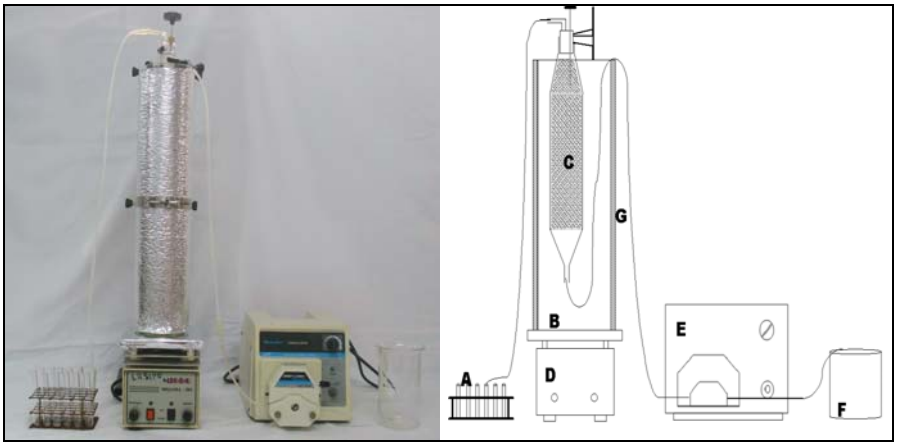


Figure 1 - Extraction system scheme: (A) collector test tubes, (B) glass cylinder, (C) column, (D) heating plate, (E) peristaltic pump, (F) solvent tank and (G) insolating.

2.2. Modeling for extraction in column with only passage of solvent

The used model divides the liquid phase in two fractions: a mobile, which flows along the bed and other immobile, consisting of liquid stagnated that doesn't participate in the flow. It took into account the following considerations: was included in model the partition phenomenon of solute initially present in the solid phase, through the Freundlich isotherms; was considered despicable the diffusive fraction associated with immobile fraction in the balance differential of solute mass; was included the relative portions to extraction transient; a first order kinetics was assumed, therefore, was obtained:

$$\left(\theta_m + \rho_{ap}\ f\ k\ n\ C_{im}^{n-1}\right)\frac{\partial C_m}{\partial t} + \frac{3\,k_{conv}\,\theta_{im}}{R}\left(C_m - C_{im}\right) = \theta_m D_m \frac{\partial^2 C_m}{\partial z^2} - \theta_m\ v\frac{\partial C_m}{\partial z} \qquad (1)$$

where: θ_m and θ_{im} = volumetric fractions of mobile and immobile water respectively; ρ_{ap} = dry bed apparent density [M L^{-3}]; f = retention sites fraction of solute on solid phase in contact with the mobile liquid fraction; k and n = Freundlich isotherm constant; C_m and C_{im} = solute concentration [M L^{-3}]; t = time [t]; D_m = hydrodynamic dispersion coefficient [L^2 t^{-1}]; z = coordinated space [L]; v = average speed interstitial pore [L t^{-1}]; R = particles radius, spherical considered [L] and k_{conv} = convective coefficient of mass transfer [$L.t^{-1}$].

In the model construction that describes the physical situation of solvent injection in a finite column of length L, for permanent flow, apply to Equation 1, the following conditions:

- Initial condition:

Concentration of inulin in this column, uniform, constant and equal to the C_{ini}:

$$C_m = C_{im} = C_{ini} \quad \text{for } t = 0 \text{ and } \forall z \qquad (2)$$

- Contour condition:

Input: the solvent solution enters with concentration C_0 and speed v, for times greater than zero.

$$v\,C_m - D_m \left.\frac{\partial C_m}{\partial z}\right|_{z\to 0^-} = v\,C_0 \quad \text{for} \quad t > 0 \quad \text{and} \quad z = 0 \tag{3}$$

Output: disregarding any dispersion effect in column output and given to the continuity condition of the variable C_m have been:

$$\left.\frac{\partial C_m}{\partial z}\right|_{z\to L^+} = 0 \quad \text{for} \quad t > 0 \quad \text{and} \quad z = L \tag{4}$$

The mathematical model described is solved using the finite volume numerical method (Patankar, 1980). All differential equations are discretized, obtaining a linear algebraic equations system, which are resolved through the matrices solution traditional method. The formulation implemented is fully implied, with linear interpolation functions.

2.3. Modeling for extraction in column with solvent recirculation

To simulate the inulin extraction in column with solvent recirculation, the model suffers alteration in the outline condition in the entrance of the column (Equation 3). In the new situation, the concentration at the column entrance will be equal to the solvent tank ($C_{reserv.inst}$) (Equation 5).

Input contour condition:

$$v\,C_m - D_m \left.\frac{\partial C_m}{\partial z}\right|_{z\to 0^+} = v\,C_{reserv.inst} \quad \text{for} \quad t > 0 \quad \text{and} \quad z = 0 \tag{5}$$

It is included in the model a relative equation to the mass balance in solvent recirculation tank (Equation 6).

$$A\,v\,C_m\big|_{z=L} - A\,v\,C_{reserv.} = V_{reserv.}\frac{dC_{reserv}}{dt} \tag{6}$$

Where: A = area of column traverse section [L^2]; v = interstitial velocity [$L\,t^{-1}$]; C_{reserv} = concentration in the recirculation tank [$M\,L^{-3}$]; $C_{reserv.\ inst}$ = concentration in the recirculation tank for some time t_{inst} [$M\,L^{-3}$] and V_{reserv} = volume of liquid accumulated in the recirculation tank on a permanent regime [L^3].

For column extraction, 3 tests were analyzed in total for each configuration, one test for each flow rate studied (0.59, 2.4 and 4.7 cm^3/min). For those cases the Table 1 shows the experimental conditions used.

The garlic particle was admitted spherical, the radius is considered as the ratio between the volume and area of the garlic slice, multiplied by 3 to maintain the same relationship with the volume and area of the sphere.

The operating conditions (Table 1) were corrected according to Equations 7 to 10 in order to take into account the glass pearls presence in bed. The introduced corrections preserved the initial inulin mass contained in garlic and the volume occupied by the particles bed.

$$v = \frac{Q}{A\,\varepsilon_{bed}} \tag{7}$$

$$\rho_{aps} = \frac{\left[m_g\left(1 - X_g\right) + m_s\right]}{A\,L_{bed}} \tag{8}$$

$$X_\infty = \varepsilon_{bed} + X_{gV}\left(1 - \varepsilon_{bed} - \frac{V_s}{V_{bed}}\right) \tag{9}$$

$$\theta_m = \frac{\varepsilon_{bed}}{X_\infty} \quad (10)$$

where: ε_{bed} = bed porosity; Q = flow rate [$L^3\ t^{-1}$]; ρ_{aps} = dry apparent density [$M\ L^{-3}$]; m_g = garlic mass [M]; m_s = spheres mass [M]; X_{gV} = garlic moisture (in volume terms); X_∞ = middle moisture; V_s = volume occupied by the spheres [L^3]; V_{bed} = bed volume [L^3] e θ_m = mobile water fraction.

Table 1 – Experimental conditions for column extraction and the parameters adjusted in the simulations.

Experimental data
Extraction Temperature - 60°C
n – 1 (isotherm equilibrium parameter)
k – 0.0268 (isotherm equilibrium parameter)
Bed particle radius (R_{bed}) – 0.1336 cm
Proportion used in bed– 34% glass pearls / 66% sliced garlic
Bed porosity (ε_{bed}) – 0.5800
Solution volume within the column (V_0) – 35 cm^3
Molecular diffusivity Coefficient – $2.048.10^{-6}$ (cm^2/seg) (Bunim et al., 1937)
Middle moisture (X_∞) – 0.8220
Dry apparent density (ρ_{aps}) – 0.3628 g/cm^3
Interstitial velocity (v) - 0.0096 cm/seg (0.59 cm^3/min)
0.0390 cm/seg (2.40 cm^3/min)
0.0764 cm/seg (4.70 cm^3/min)
Hydrodynamic dispersion coefficient(D_{ap}) –
0.0082 cm^2/seg (0.59 cm^3/min)
0.0184 cm^2/seg (2.40 cm^3/min)
0.0922 cm^2/seg (4.70 cm^3/min)
Initial solute concentration in immobile phase –
0.1650 g/ml (0.59 cm^3/min)
0.2315 g/ml (2.40 cm^3/min)
0.0689 g/ml (4.70 cm^3/min)

In simulations of column process, the adjusted parameters were the convective mass transfer coefficient (k_{conv}) and mobile water fraction (θ_m). The mobile water fraction, although it was estimated by Equation 10, has been changed according to the need for adjustments to the experimental data. This is due to the fact that stagnant water fraction depends on the bed arrangement and the flow regime.

3. Results

3.1. Modeling for column extraction with only passage of solvent

The Table 2 presents the adjusted parameters in the simulation in only passage of solvent in the flow rate of 0.59, 2.4, 4.7 cm^3/min.

Table 2 – Adjusted parameters in the simulation without solvent recirculation.

	Flow rate (cm^3/min)		
	0.59	2.40	4.70
Mobile water fraction (θ_m)	0.55	0.55	0.55
Convective mass transfer coefficient (k_{conv}) (cm/seg)	$1.0.10^{-5}$	$1.4.10^{-5}$	$2.2.10^{-5}$

*Modeling and simulation the fixed bed column extraction of inulin from garlic (*Allium sativum L. *var. Chonan)*

During the simulations, the model capacity was verified of adjusting to the experimental data. The Figure 2 shows the simulations results accomplished together with the experimental data for the flow rate of 0.59, 2.4 e 4.7 cm^3/min. The abscissa axis shows the ratio of the solvent volume injected in the column and the pore volume (empty spaces); already the ordinates axis represents the ratio between the solute concentration in the column output and the solute concentration in the immobile phase, in other words, the inulin concentration in the liquid phase interns to the garlic.

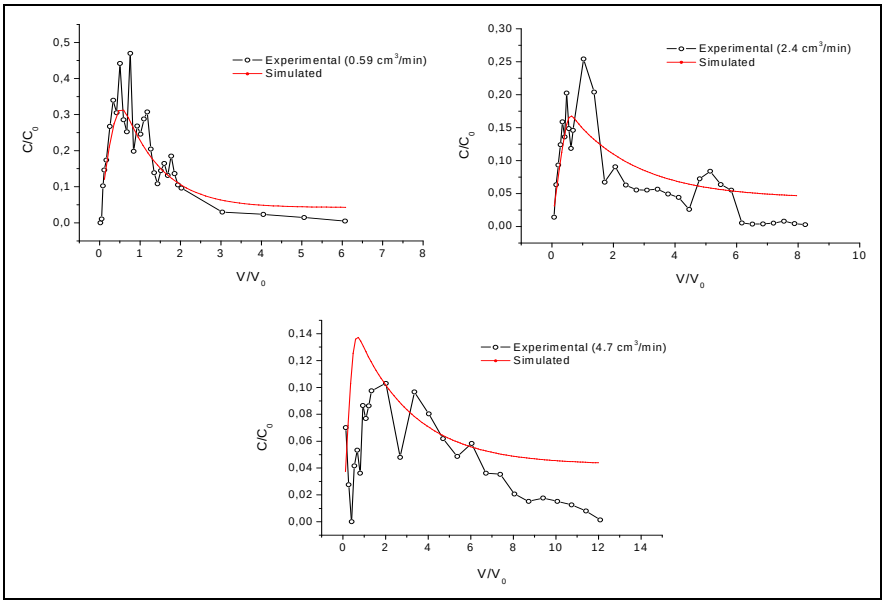


Figure 2 – Column experiments simulations with single passage of solvent.

Observed in Figure 2 deviations among the adjusted curve for the model and of the experimental data, being more evident in the flow rate of 4.7 cm^3/min. It is noted that, in adjustment the parameters, for a coherence subject, the same values were used for both column extraction configurations (single pass and recirculation).
Therefore, those deviations could be improved if each case had been analyzed individually. But, in general, adequately describes the behavior of inulin extraction from garlic.

3.2. Modeling for extraction in column with recirculation of solvent

The input data were the same used in the configuration of single passage, presented in the Table 1.
The Figure 3 shows the simulations results accomplished together with the experimental data relating to the column extraction with solvent recirculation in the flow rate of 0.59; 2.4 and 4.7 cm^3/min.
In the adjustment of parameters, tried to preserve the values for both configurations, with and without recirculation (Table 2). Therefore, the experimental data simulation represents an effort in the sense of assisting the results group through a single model and with the broad parameters of all tests. In general, the model presented good agreement with the experimental data, reproducing the inulin extraction tendency appropriately along the process.

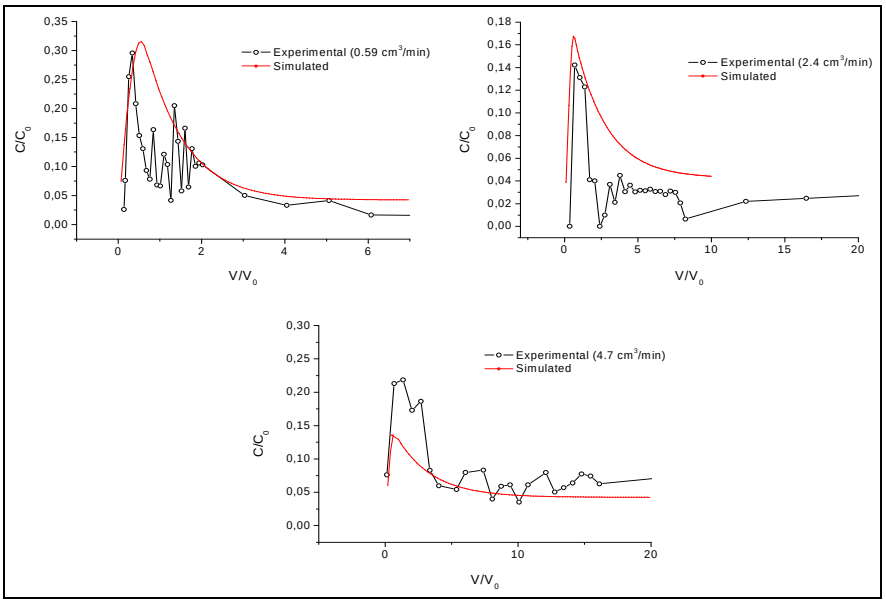


Figure 3 – Column experiments simulations with solvent recirculation.

The values found for column Peclet number were close to 40, 72 and 28 for the flow rate of 0.59 cm^3/min, 2.4 cm^3/min and 4.7 cm^3/min respectively. The column Peclet number lists the hydrodynamic dispersion mechanisms and convection in the solute displacement, even taking into account the bed length. These values are considered low and indicate the importance of the bed dispersive nature (porous medium) for carrying out such experiments.

4. Conclusion

In the column process simulation, were obtained low values for Peclet number, indicating the bed dispersive nature.

Overall, the studied models acted in a satisfactory way the inulin extraction tendency, for the experimental data adjustment.

References

K. Narinder; A. K. Gupta, 2002, Applications of inulin and oligofructose in health and nutrition. Journal of Biosciences, v.27, n.7, p.703-714.

G. R. Gibson; M. B. Roberfroid, 1993, Dietary modulation of the human colonic microbiota - introducing the concept of prebiotics, Journal of Nutrition, v. 125, p.1401-1412

J. J. Bunim, W. W. Smith, H. W. Smith, 1937, The diffusion coefficient of inulin and other substances of interest in renal physiology, The Journal of Biological Chemistry. p.667-677

M. Roberfroid, 1993, Dietary fiber, inulin and oligofructose: a review comparing their physiological effects. Critical Reviews in Food Science and Nutrition, v.33, p.103-148

S. V. Patankar, 1980, Numerical Heat Transfer and Fluid Flow. Hemisphere Publishing Corporation, New York.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Particle Swarm Optimisation Applied in Retrofit of Heat Exchanger Networks

Aline P. Silva[a,b], Mauro A. S. S. Ravagnani[a], Evaristo C. Biscaia Jr.[b]

[a]*DEQ, Universidade Estadual de Maringá, Av. Colombo, 5790 – Bl D90, CEP 87020-900, Maringá – PR, Brasil, E-mail address: aline@deq.uem.br, ravag@deq.uem.br*
[b]*PEQ/COPPE, Universidade Federal do Rio de Janeiro, Av. Horácio Macedo, 2030 - Centro de Tecnologia - Bloco G - Sala G-115, CEP 21941-914, Cidade Universitária / Ilha do Fundão, Rio de Janeiro - RJ, Brasil, E-mail address: evaristo@peq.coppe.ufrj.br*

Abstract

In this work, a methodology for the retrofit of heat exchanger network (HEN) considering simultaneously the possibilities of reuse of exchangers for different junctions and the inclusion of new units is proposed. The developed methodology is based in a superstructure simultaneous optimisation model for the HEN synthesis considering stream splitting, with an additional constraint of including all the existent equipments among the possibilities of heat exchange. This superstructure takes into account the operational and capital costs, by maximizing the recovery of energy and minimizing the installation costs. The case is treated as a constrained optimisation problem whose objective is to minimize the global cost of the HEN. The object function considers the total cost composed by the cost of the utilities used in the HEN and the cost of the new equipments. The constraints are the thermodynamics limits and the obligation of reusing all the existing heat exchangers. The problem was solved using Particle Swarm Optimization (PSO), giving consistent and promising results. The main contribution of this paper is, besides the retrofit of heat exchanger networks in industrial processes, the optimization technique used. PSO is able to avoid local optima and suits very well to non-linear problems.

Keywords: Retrofit of Heat Exchanger Network, Optimization Problem, Particle Swarm Optimization.

1. Introduction

The synthesis of heat exchanger networks (HEN) is an important field in process systems engineering, and has been the subject of considerable research efforts over the last 40 years. Many studies and methodologies were proposed to make possible the energy recovery between process streams, minimizing the utilities consumption and the number of heat transfer equipment. Most of these formulations have been developed to the synthesis of new plants. Industrial processes in operation, however, can often be made more energy-efficient by a retrofit.

As pointed in Briones and Kokosis (1999), retrofit problems represent a large amount of industrial process design. However, retrofit technology when compared to grassroots designs has a less intense progress, considering both, mathematical programming and methods based on thermodynamic analysis.

According to Westerberg (1989) and Yee and Grossmann (1991), the systematic work on retrofit is considerably harder than grassroots designs. It is due to the necessity of considering reuse of existing heat exchangers in despite of new unities design.
There are two groups of papers related to the retrofit of heat exchanger networks. The first one is based on thermodynamic concepts and Pinch Analysis (Asante and Zhu, 1996 and Silva and Zemp, 1998 and Panjeh Shahi *et al*. 2008). The second one is based on mathematical programming (Singh and Castillo 2002, Mizutani, 2003 and Bjork and Nordman, 2005).
The first papers in the retrofit of heat exchanger networks were published during the eighties, involving thermodynamic concepts. Tjoe and Linnhoff (1986) proposed for the retrofit the calculus of the optimum heat transfer area based in Pinch Technology. The authors related this optimum area with the current network area, but they didn't take into account the equipment location.
The first paper based on mathematical programming was the work of Ciric and Floudas (1989). The authors presented a two stage methodology to the retrofit of heat exchanger networks. In the first stage a MILP formulation selects the matches between the streams, minimizing the investment costs. In a second stage, a NLP formulation is used to optimize the network with the matches defined in the first step.
Yee and Grosmman (1990) proposed a MILP formulation model to establish the structural modifications on the heat exchanger network.
In the works of Ciric and Floudas (1990) and Yee and Grosmman (1991) the authors used superstructures whose models has the formulation of MINLP problems.
Using both, Pinch Analysis and mathematical programming, Zhu and Asante (1999) developed a methodology considering the costs of modifications in the process utilities as well as in the network topology. Firstly the network pinches are identified as well as the heat exchangers that cause a bottleneck in the process due to the violation of the minimum temperature difference. This is done by a LP formulation in which the maximum energy recovery is searched with modifications on the network topology. After this, by using a MILP the modifications on the network topology are calculated (new heat exchanger incorporation, stream splitting and position changes in the equipment) aiming to maximize the energy recovery without considering the costs of installation of these modifications. The propositions of modifying the original network can be adopted or discharged by the user, in an ulterior evaluation stage. Finally, the proposed chosen changes can be optimized considering the global, operational and fixed costs, with the topologies maintained unaltered, i. e., the variables of optimization in this stage are the flowrates of the stream split and stream temperatures maintaining fixed the network structure. The authors did not take into account pumping costs in their proposition.

2. Problem Definition

The developed methodology is based on the superstructure model presented by Yee and Grossmann (1990). The difference is that the authors considered the number of stages equal to the maximum number of hot or cold streams. In the present paper, the number of stages of the superstructure is a variable to be optimized and not a parameter as fixed by Yee and Grossmann (1990).
The difference between the grassroots design and the retrofit is in the inclusion of the condition of including obligatorily all the existent heat exchangers in the heat exchange possibilities. The equipments are used in the same matches or in other positions in the future heat exchanger network structure.

The problem consists in determining a heat exchanger network considering a set of hot and cold streams with known inlet and outlet temperatures and flowrates. Besides, it must be known hot and cold utilities and the corresponding temperatures.
To the current heat exchanger network must be considered the known matches in each existent heat exchanger in the network as well as its heat transfer areas. The costs of the minimum utility consumption and for the addition of new heat transfer equipment must be also considered.
In the present paper for the solution of the constrained optimization problem of the retrofit of heat exchanger networks an objective function considering the utility costs used in the generated heat exchanger network associated with the cost involving the increase of new heat exchangers. The constraints are the thermodynamic restrictions in the heat matches. Equations (1) to (4) present the constrained optimization problem.
The proposed problem was solved with Particle Swarm Optimisation (PSO). PSO is a stochastic optimisation technique developed by Kennedy and Elberhart (2001), inspired in the social behaviour of bird flocking or fish schooling. PSO as an optimisation tool provides a population-based search procedure in which individuals called particles change their position (state) with time. In a PSO system, particles fly around a multidimensional search space. During flight, each particle adjusts its position according to its own experience, and according to the experience of a neighbouring particle, making use of the best position encountered by itself and its neighbour.
According to the proposed methodology to solve the problem, 5 variables are initially randomly generated. These variables are modified at each iteration according to Equations (5) and (6). Each particle is formed by the following variables: Stream splitting fraction for the hot and cold streams, number of stages, position at each heat exchanger is allocated and the heat duty for each heat exchanger.
After the particle generation the area of each new heat exchanger is calculated, as well as the hot and cold utilities demand. This is done for all the particles, being the particle the solution of the problem or not. The objective function value (global network cost) is achieved and case the particle is not the problem solution (some constraint is violated), the objective function is penalized.

Minimize: $\mathrm{C}_{\text{total}} = \mathrm{C}_{\text{area}} + \mathrm{C}_{\text{utility}}$

Subject to:
$$\begin{cases} \left(T_h^{in}\right)_k > \left(T_c^{out}\right)_k \\ \left(T_h^{out}\right)_k > \left(T_c^{in}\right)_k \end{cases} \quad (1)$$
$\mathrm{k} = 1,\ldots,$ number of equipments

$$\mathrm{C}_{\text{area}} = \sum_{1}^{k} \mathrm{aa}_1 \cdot (A)^{\mathrm{aa}_2} \quad (2)$$
$\mathrm{k} = 1,\ldots,$ number of equipments

$$\mathrm{C}_{\text{utility}} = \mathrm{C}_{\text{HU}} \cdot \mathrm{HU} + \mathrm{C}_{\text{CU}} \cdot \mathrm{CU} \quad (3)$$

$$k^{MAX} = N \cdot ncq \cdot ncf \quad (4)$$

$$v_i^{k+1} = w \cdot v_i^k + c1 \cdot r_1 \cdot \left(p_i^k - x_i^k\right) + c2 \cdot r_2 \cdot \left(p_{GLOBAL}^k - x_i^k\right) \quad (5)$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad (6)$$

In Equations (5) and (6), $\mathbf{x}_i$ and $\mathbf{v}_i$ are, respectively, position vectors and velocity of the particle *i*, *w* is the inertial weight, *c1* and *c2* are constant, $\mathbf{r}_1$ and $\mathbf{r}_2$ are vectors containing random numbers with uniform distribution in [0, 1], $\mathbf{p}_i$ is the position with the best value of the objective function that the particle *i* has found and $\mathbf{p}_{global}$ is the position of the best value found for the set of particles.

3. Case Study

An example with various set of different parameter for the PSO was used and the influence of each set in the algorithm performance was studied (Table 1).

The case studied was presented in Panjeh Shahi *et al*. (2008). The heat exchanger network has 2 heat exchangers, 3 heaters and 2 coolers. The utilities demands are 23.498 kW for the cold utility and 1.708 kW for the hot utility. One can assume that the individual heat transfer coefficients are 0.2 kW/m2 ºC for all the streams. Figure 1 presents the existent heat exchanger network and Table 2 shows the heat transferred and the area values for the existing heat exchangers.

Table 1 – Parameters for the PSO

c1	c2	w	Npt
1.0	1.0	0.65	30

By applying the developed methodology, the new heat exchanger network presents 1 heater, 3 coolers and 4 heat exchangers, being 2 new and 2 old reused equipments. Table 3 shows the heat duty, the area to be increased in the network (2 new equipments) and the costs. Hot and cold utilities are, respectively, 23.00 $/kW/year and 1.52 $/kW/year. Additional heat exchangers costs were calculated using Equation (7), with *A* in m^2.

$$C_{area}(\$) = 37643 + 7529 \cdot A^{0.7} \tag{7}$$

The global cost for the new heat exchangers network is $ 68,068. It means saving 21,478 $/year in utilities use. Table 4 shows the consumption and utilities cost for the current and for the proposed heat exchanger network. The payback period is 3.2 years. Figure 2 shows the new heat exchanger network. (The new equipments are in bold).

The computational time at Pentium IV 170 GHz was approximately 500 minutes. Success rate of PSO was 32%.

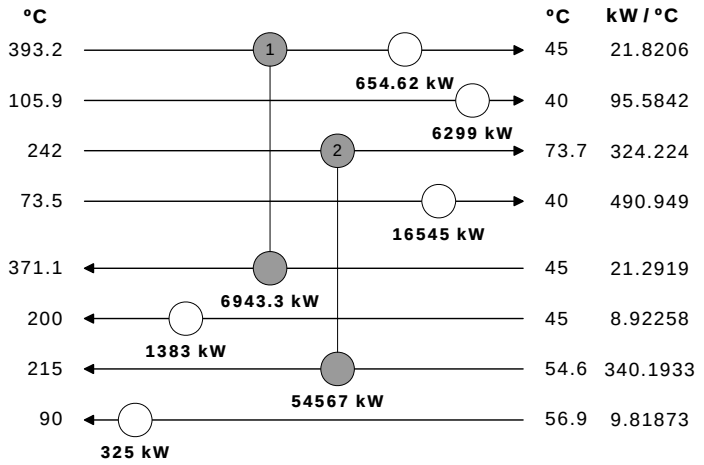


Figure 1 – Current heat exchanger network

Table 2 – Heat duty and heat transfer area for the existing equipments

Equipment	Q (kW)	Area (m^2)
1	6943	2686
2	54567	23909

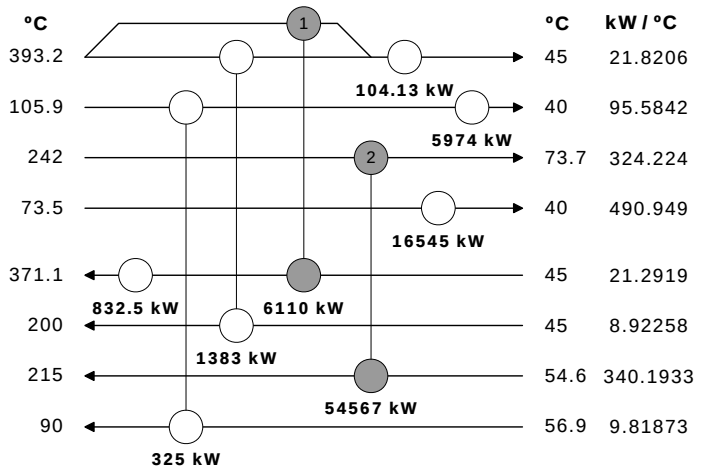


Figure 2 – Final heat exchanger network (after the *retrofit*)

Table 3 – Heat duty, area and new heat exchangers cost

Equipment	Situation	Q (kW)	Area (m^2)	Cost ($)
1	Current	6110	2686	*
2	Current	54567	23909	*
3	New	325	115	24656
4	New	1383	288	43412

* Equipment existent with no cost

Table 4 – Cost and utilities demand

	Current		New	
	Consumption (kW)	Cost ($/year)	Consumption (kW)	Cost ($/year)
UQ	1708	39284	832	19136
UF	23498	35717	22623	34387
Total		75001		53523

4. Conclusions

In this paper it was presented a new methodology for the retrofit of heat exchanger networks. The goal is to find the best configurations for a heat exchanger network with minimum cost, taking into account the area and utilities costs, considering streams physical properties and cost parameters. The existing heat exchangers must be considered in the analysis. The methodology is based on the superstructure proposition of Yee and Grossmann (1990) and the optimization problem is solved using Particle Swarm Optimization.

An example from the literature was used to test the developed methodology. Results showed the best configurations for the heat exchanger network, considering the existing equipments and saving energy.

Being a highly complex problem, the computational time is acceptable. Success rate is low because of considering the existing heat exchangers as a constraint to the problem, decreasing the convergence in the viable region solution.

Nomenclature

A	heat transfer area
aa_1, aa_2	constants
C_{area}	heat transfer area cost
C_{CU}	cold utility cost
C_{HU}	hot utility cost
C_{total}	global cost
$C_{utility}$	utilities cost
CU	cold utility demand
HU	hot utility demand
k^{MAX}	maximum number of heat transfer equipments
N	number of stages in the superstructure
ncf	number of cold streams
ncq	number of hot streams
T^{in}	inlet temperature
T^{out}	outlet temperature

References

Asante, N. D. K. and X. X. Zhu, , 1996, "An Automated Approach For Heat Exchanger Network Retrofit Featuring Minimal Topology Modifications". Comp. Chem. Eng. V. 20, Pp. S7–S12.

Bjork, K. and Nordman, R., 2005, "Solving Large-Scale Retrofit Heat Exchanger Network Synthesis Problems". Chem. Eng. Proc. V. 44, Pp. 869-876.

Briones, V. and Kokossis, A. C., 1999, "Hypertargets: A Conceptual Programming Approach For The Optimization Of Industrial Heat Exchanger Networks – Ii. Retrofit Design", Chem. Eng. Sci., V. 54, Pp. 541-561.

Castillo, E., Acevedo, L. and Reververi, A., 1998, "Cleaner Production Of Nitric Acid By Heat Transfer Optimization: A Case Study", Chem. Biochem. Eng., 12, 157 – 165.

Ciric, A. R., and Floudas, C. A., 1989, "A Retrofit Approach For Heat Exchanger Networks", Comp. Chem. Eng., V. 13, N. 6, Pp. 703-715.

Ciric, A. R. and Floudas, C. A., 1990, "A Mixed Integer Nonlinear Programming Model For Retrofitting Heat-Exchanger Networks", Ind. Eng. Chem. Res., V. 29, 239-251.

Kennedy, J. and Eberhart, R.; Swarm Intelligence. Academic Press, London, 2001.

Mizutani, F. T., 2003, Síntese De Redes De Trocadores De Calor Com O Projeto Detalhado Das Unidades Via Programação Matemática, Tese De D.Sc., Escola De Química/Ufrj, Rio De Janeiro, Rj, Brasil.

Panjeh Shahi, M. H., Ghasemian Langeroudi, E., Tahouni, N., 2008, "Retrofit Of Ammonia Plant For Improving Energy Efficiency", Energy, V. 33, Pp 46–64.

Silva, M. L. and Zemp, R. J., 1998, "Retrofit De Redes De Trocadores De Calor Utilizando Matriz De Área Com Restrições De Queda De Pressão", In: Anais Do 12º Congresso Brasileiro De Engenharia Química (Cobeq), Porto Alegre/Brasil (Setembro), Em Cd-Rom.

Singh, H. and Castillo, F., 2002, "Process Life Cycle Solutions For The Case Of Automated Heat Exchanger Network Retrofit". Applied Thermal Eng. V. 22, Pp. 949-958.

Tjoe, T.N. and Linnhoff, B., 1986, "Using Pinch Technology For Process Retrofit", Chemical Engineering (April 28), Pp. 47 - 60.

Westerberg, A. W., 1989, "Synthesis In Engineering Design", Comp. Chem. Eng., N. 13, Pp. 365-376.

Yee, T. F. and Grossmann, I. E., 1990, "Simultaneous Optimization Models For Heat Integration – Ii Heat Exchanger Network Synthesis", Comp. Chem. Engng., 14(10): 1165-1184.

Yee, T., F. and Grossmann, I. E., 1991, "A Screening And Optimization Approach For The Retrofit Of Heat-Exchanger Networks", Ind. Eng. Chem. Res., N. 30(1): Pp. 146-162.

Zhu, X. X. and Asante, N. D. K., 1999, "Diagnosis And Optimization Approach For Heat Exchanger Network Retrofit", Aiche Journal, N. 45(7): Pp. 1488-1503.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Improving energy performance of a hydrazine hydrate plant by applying hot feed

Xiao Feng, Junkun Yang, Taofeng Cao

Department of Chemical Engineering, Xi'an Jiaotong University, Xi'an 710049, China

Abstract

When analyzing a heat exchanger network, it can be often found, that a given stream is cooled first and then heated or in reverse. In such a situation, one of the available energy-saving methods is a hot feed strategy. In this paper, the heat exchanger network in a hydrazine hydrate plant is analyzed and retrofitted. From the analysis, it can be found that the extensive energy consumption takes place is some coolers above the pinch. Additionally, the product is cooled before leaving the unit, and heated afterwards when entering the downstream plant. By eliminating cooling and the subsequent heating, some streams will be the hot feeds for the downstream processes. In this way, 60.7% of the current heating utility can be saved after the optimization.
Keyword: hot feed, heat integration, hydrazine hydrate plant, pinch technology

1. Introduction

Energy saving is a pressing issue. In the chemical industry, where the processes are usually complex and the energy consumption is high, the optimization of heat exchanger networks has been always a concern. Since Hwa proposed the structure optimization of heat exchanger networks in 1965(Hwa, 1965), the heat exchanger network synthesis and retrofit became an important branch of process integration. It is worth to note that one of the most widely used methods (Mubarak and Kawari, 2000; Feng, 2004) for heat exchange network optimization is pinch technology.

When integrating a heat exchange network via pinch technology, the first step is to extract stream data, next is to determine the "bottleneck" of the system, and then to optimize it. (Mubarak and Kawari, 2000) The optimization of energy system is especially complex in large-scale process systems, where there are large numbers of streams linked with each other. When analyzing a heat exchanger network, it can be sometimes found that a stream may be cooled first and then heated, or heated first and then cooled. It means that a stream may be both a hot stream and a cold stream at different locations. More often, the feed of a plant is a cold stream, but it is a hot stream in the upstream plant. In this case, a more effective energy-saving method is to change the process itself. Although much research on the pinch technology and its application has been reported (Linnhoff and Townsend, 1982; Sorin and Paris, 1997; Singhvi et al, 2004; Herrera et al, 2003), there are no publications on changing process by using hot feed to achieve energy-saving effect.

Hydrazine hydrate is an important chemical raw material used in production of pesticides, pharmaceuticals, chemical intermediates, dyes, imaging agents, antioxidant materials. A hydrazine hydrate plant is consisted of three units: the reaction unit, the evaporation and fractionation unit, and the fractionation and concentrating unit. The heating utility in the process is steam.

In this paper, the heat exchanger network of the hydrazine hydrate plant in Yibin Tianyuan Group Company Limited, Sichuan, China, is analyzed. The unreasonable energy utilization is detected by pinch technology. The hot feed strategy is utilized to synthesize the heat exchange network.

2. The stream data of the hydrazine hydrate plant

There are 11 hot streams, as shown in Table 1, and 13 cold streams, as shown in Table 2, in the plant. There are some batch processes in the plant. Their heat loads are averaged, basing on 24 hour values, when calculating the pinch position. The batch processes are shown in Table 3. The current heat exchange network consumes 44195 kW heating utility and 33070 kW cooling utility.

Table 1. The hot stream data in the hydrazine hydrate plant

No.	Stream	Supply temperature/℃	Target temperature/℃	Heat load/kW
H1	Chloros	20	10	555.56
H2	Synthetic fluid	100	80	1388.89
H3	Synthetic fluid	80	65	471.06
		65	45	636.57
		45	-2	135.42
H4	Synthetic fluid	80	67	444.44
		67	50	597.22
		50	30	694.44
		30	-2	57.87
H5	Overhead product of fractionation columnT1101a	85	30	3055.56
H6	Overhead product of fractionation column T1101b	85	30	4888.89
H7	Overhead product of fractionation column T1101c	85	30	8250.00
H8	Overhead product of fractionation column T1101d	85	30	8250.00
H9	Raw product	100	70	2361.11
H10	Distillate from product column T1103a	113	70	1362.85
H11	Distillate from product column T1103b	113	70	1753.47

Table 2. The cold stream data in the hydrazine hydrate plant

No.	Stream	Supply temperature/℃	Target temperature/℃	Heat load/kW
C1	Mixture of chloros and carbamide	12	108	2527.78
C2	Mother Liquid	5	30	1388.89
		30	44	444.44
C3	Mother Liquid	44	110	6042.5
C4	Mixture of hydrazine hydrate and water	95	105	611.11
C5	Mother Liquid	44	110	5694.44
C6	Mixture of hydrazine hydrate and water	95	105	972.22
C7	Mother Liquid	45	95	9611.11
C8	Mixture of hydrazine hydrate and water	95	105	1638.89
C9	Mother Liquid	45	95	9611.11
C10	Mixture of hydrazine hydrate and water	95	105	1638.89
C11	Mixture of hydrazine hydrate and water	70	110	2583.33
C12	Raw product	60	123	1423.61
C13	Raw product	60	123	1840.28

Table 3. The batch processes in the hydrazine hydrate plant

Heat exchanger	Runtime h/day	Actual heat load/kW	Calculated heat load/kW
E1105a	11	1027.78	471.06
E1105b	11	1388.89	636.57
E1106	6.5	500	135.42
E1107a	12	888.89	444.44
E1107b	12	1194.44	579.22
E1107c	12	1388.89	694.44
E1108	5	277.78	57.87
E1123	15	2277.78	1423.61
E1124	15	2180.56	1362.85
E1125	15	2944.44	1840.28
E1126	15	2805.56	1753.47

3. System analysis

Considering the recovered energy, the heat transfer area, the retrofit cost, and the stability of operation, the minimal temperature difference is taken as 10 ºC. According to the stream data in Table 1 and Table 2, the average pinch temperature can be calculated as 50 ºC, that is, the temperature of the hot streams at the pinch is 55ºC, and that of the cold streams is 45ºC. The minimum heating utility is 21294 kW, and the minimum cooling utility is 10168 kW. Therefore, the energy-saving potential of the network is 22902 kW, which accounts for 51.8% of the current heating utility. The T-H composite curve is shown in Figure 1.

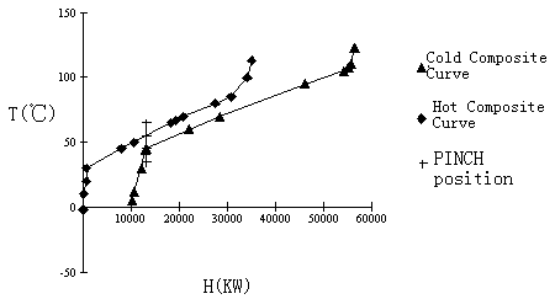


Figure 1. The composite curve of the hydrazine hydrate plant

The current heat exchange network of the hydrazine hydrate plant is shown in Figure 2. According to the three basic principles of pinch technology: (1) no heat transfer across the pinch; (2) no heating utility below the pinch; (3) no cooling utility above the pinch, we can find the unreasonable energy utilization in the hydrazine hydrate plant, as shown in Table 4. From Table 4, it can be found that in the current heat exchange network, ten coolers, two exchangers and three heaters are set unreasonably.

Table 4. The unreasonable energy utilization in the hydrazine hydrate plant

Heat exchanger	*Stream*	*Unreasonable heat load/kW*	*Type of unreasonable heat match*
E1105a	Synthetic fluid	471.06	Cooling utility above the pinch
E1105b	Synthetic fluid	318.29	Cooling utility above the pinch
E1107b	Synthetic fluid	421.57	Cooling utility above the pinch
E1111abc	Overhead product of fractionation column T1101a	1666.67	Cooling utility above the pinch
E1114abc	Overhead product of fractionation column T1101b	2666.67	Cooling utility above the pinch
E1117abcde	Overhead product of fractionation column T1101c	4500.00	Cooling utility above the pinch
E1120abcde	Overhead product of fractionation column T1101d	4500.00	Cooling utility above the pinch
E1122	Raw product	2361.11	Cooling utility above the pinch
E1124	Distillate from product column T1103a	1362.85	Cooling utility above the pinch
E1126	Distillate from product column T1103b	1753.47	Cooling utility above the pinch
E1104	Mother liquid	1388.89	Heat transfer across the pinch
E1107a	Mother liquid	444.44	Heat transfer across the pinch
E1102abc	Mixture of chloros and carbamide	868.92	Heating utility below the pinch
E1109	Mother Liquid	91.56	Heating utility below the pinch
E1112	Mother Liquid	86.28	Heating utility below the pinch
Total		22902	

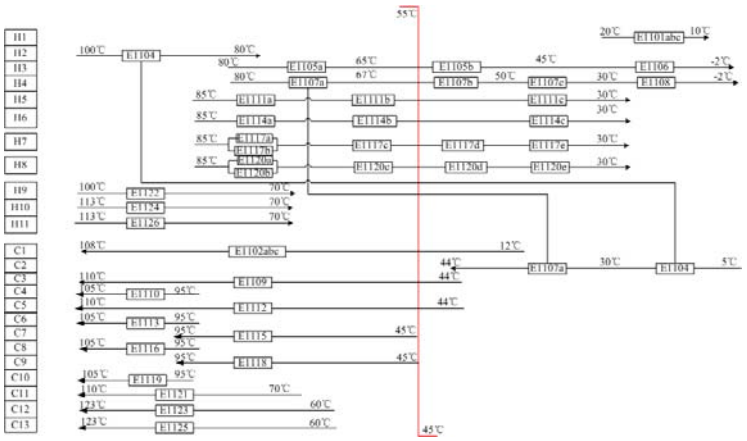


Figure 2. The current heat exchange network of the hydrazine hydrate plant

4. The retrofit scheme

The current heat exchange network of the hydrazine hydrate plant consumes steam and is composed of many batch processes. To recover the 22902 kW unreasonably used utility, large investment in equipment is needed. So, in such a situation, the economic benefit of the retrofit would be less. It can be found when inspecting the current process,

that the following heat matches are unreasonably set, and so can be retrofitted without affecting the production process.

(1). In the evaporation and fractionation unit, the mother liquid is fed into fractionation columns (T1101a, T1101b, T1101c, and T1101d). The overhead products (H5, H6, H7 and H8) are cooled from 85ºC to 30ºC by coolers, and then flow out of the plant and are fed into the ammonia stripping process in the downstream plant. These coolers use cooling utility above the pinch, which accounts for 58% of the unreasonably used energy. However, these streams are heated when entering ammonia stripping column. Figure 3 gives the flow sheet of the current process.

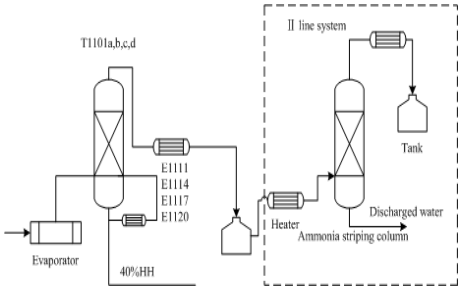


Figure 3.The current process between the plant and its downstream plant

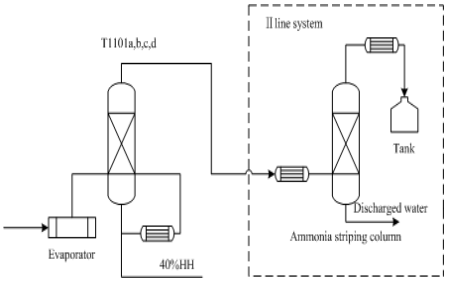


Figure 4. The retrofit scheme between the plant and its downstream plant

Therefore, the overhead products of the fractionation columns go through the processes first cooled then heated before they are fed into the ammonia stripping column. The overhead products of the fractionation columns can be fed into the ammonia stripping column directly without cooling. After such retrofit, 24445 kW steam and the equivalent of cooling water can be saved. The process after the retrofit is shown in Figure 4. It is seen from the pinch analysis, presented in Section 2, that H5, H6, H7 and H8 consume 13333 kW cooling utility above the pinch. That means that at most 13333 kW can be saved by retrofit of the heat exchanger network according to the traditional pinch technology. However, by using hot feed, 24445 kW heating utility can be saved without adding a new heat exchanger. (2). In the fractionation and concentrating unit, the raw product flows into the product column (T1102). The overhead product (H9) is cooled from 100ºC to 70ºC in E1122 and flows into the raw product tank. Then it is split into two streams (C12, C13), heated from 60ºC to 123ºC, and fed into product columns (T1103a, T1103b) separately. The current process is shown in Figure 5.

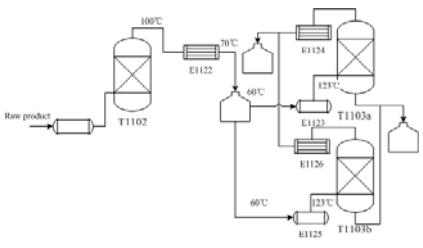


Figure 5. The current process around the product columns

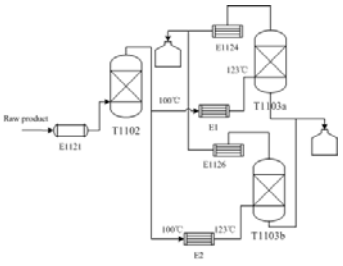


Figure 6. The retrofit scheme around the product columns

Therefore, the overhead product of column T1102 go through the processes first cooled and then heated before it is fed into the product columns (T1103a, T1103b). In consequence, the cooling process can be skipped. Then the overhead product of column T1102 can be heated from 100ºC to 123 ºC and fed into the product columns (T1103a,

T1103b). In this way, 2361 kW steam and the equivalent cooling water can be saved. The process after retrofit is shown in Figure 6.
In the current process, the heat exchanger H9 consumes cooling utility above the pinch. The unreasonable energy utilization is 2361 kW. By using hot feed, 2361 kW heating utility can be saved without adding a new heat exchanger. The results obtained from pinch analysis show that the energy saving potential is 22902 kW. However, after the retrofit by using hot feed, 26806 kW heating and cooling utilities can be saved, which accounts for 60.7% of the current heating utility. Therefore, the retrofit gets higher energy saving effect and reduces two heat exchangers and two tanks. The economic benefits are obvious.

5. System analysis after considering hot feed option

After introducing the above hot feeds, the stream data will change, as shown in Table 5 for hot streams and in Table 6 for cold streams. The heat exchange network consumes 41834 kW heating utility and 6265 kW cooling utility.

Table 5. The hot stream data after considering hot feeds

No.	Stream	Supply temperature/°C	Target temperature/°C	Heat load/kW
H1	Chloros	20	10	555.56
H2	Synthetic fluid	100	80	1388.89
H3	Synthetic fluid	80	65	471.06
		65	45	636.57
		45	-2	135.42
H4	Synthetic fluid	80	67	444.44
		67	50	597.22
		50	30	694.44
		30	-2	57.87
H10	Distillate from product column T1103a	113	70	1362.85
H11	Distillate from product column T1103b	113	70	1753.47

Table 6. The cold stream data after considering hot feeds

No.	Stream	Supply temperature/°C	Target temperature/°C	Heat load/kW
C1	Mixture of Chloros and carbamide	12	108	2527.78
C2	Mother liquid	5	30	1388.89
		30	44	444.44
C3	Mother liquid	44	110	6042.5
C4	Mixture of hydrazine hydrate and water	95	105	611.11
C5	Mother liquid	44	110	5694.44
C6	Mixture of hydrazine hydrate and water	95	105	972.22
C7	Mother liquid	45	95	9611.11
C8	Mixture of hydrazine hydrate and water	95	105	1638.89
C9	Mother liquid	45	95	9611.11
C10	Mixture of hydrazine hydrate and water	95	105	1638.89
C11	Mixture of hydrazine hydrate and water	70	110	2583.33
C12	Raw product	100	123	393.77
C13	Raw product	100	123	509.01

The minimal temperature difference is still taken as 10 ºC. According to the stream data in Tables 5 and 6, the average pinch temperature is calculated as 15ºC. The minimum heating utility is 35951 kW, and the minimum cooling utility is 391 kW. The energy-saving potential of the network is 5884 kW, which accounts for 14ºC of the current heating utility. The T-H composite curve is shown in Figure 7. This time the unreasonable energy utilization is shown in Table 7.

Because E1105a, E1105b, E1106, E1107c, E1107b, E1108, E1124 and E1126 are batch operations, and the small heat load in E1104 will result in high retrofit cost, the heat exchanger network is not retrofitted further.

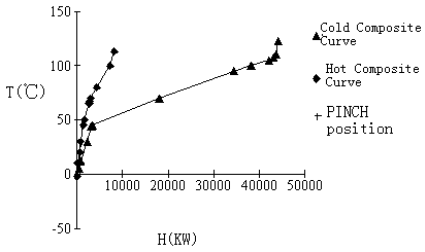


Figure 7 The composite curve after considering hot feed

Table 7. The unreasonable energy utilization after considering heat feed

Heat exchanger	*Stream*	*Unreasonable heat load/kW*	*Type of unreasonable heat match*
E1104	Mother liquid	277.8	Heat transfer across the pinch
E1105a	Synthetic fluid	471.06	Cooling utility above the pinch
E1105b	Synthetic fluid	636.57	Cooling utility above the pinch
E1106	Synthetic fluid	72.03	Cooling utility above the pinch
E1107b	Synthetic fluid	597.44	Cooling utility above the pinch
E1107c	Synthetic fluid	694.44	Cooling utility above the pinch
E1108	Synthetic fluid	18.08	Cooling utility above the pinch
E1124	Distillate from product column T1103a	1362.85	Cooling utility above the pinch
E1126	Distillate from product column T1103b	1753.47	Cooling utility above the pinch
Total		5884	

6. Conclusion

In this paper, the heat exchange network in a hydrazine hydrate is analyzed. It is found that the unreasonable energy utilization is coolers above the pinch. The main hot streams which are cooled above the pinch will be heated afterwards. Based on the analysis, a retrofit scheme with hot feeds is proposed, by eliminating all the streams which are first cooled and then heated. In this way, 26806 kW heating utility and cooling utility is saved, which account for 60.7% of the current heating utility and 81.1% of the current cooling utility.

Acknowledgements

Financial support provided by the National Natural Science Foundation of China under Grant No. 50876079 is gratefully acknowledged.

References

X. Feng, 2004, Energy Saving Principle and Technology in Chemical Engineering, 2nd Edition, Beijing: Chemical Industry Press.

A.Herrera, J.Islas,, A.Arriola, 2003, Pinch technology application in a hospital, Applied Thermal Engineering, 23 (2), 127 - 139.

C S. Hwa, 1965, Mathematical Formulation and Optimization of Heat Exchange Network Using Separable programming. AIChE - Intern Chem Eng Symp Series, 4, 101.

B. Linnhoff, D. W. Townsend, D. Boland, et al., 1982, A user guide on process integration for the efficient use of energy, The Institution of Chemical Engineers.

E. Mubarak, A I. Kawari, 2000, Pinch technology: an efficient tool for chemical-plant energy and capital-cost saving. Applied Energy, 65 (1-4), 45-49

A. Singhvi, K. P. Madhavan, U.V. Shenoy, 2004, Pinch analysis for aggregate production planning in supply chains. Computers and Chemical Engineering, 28 (6-7), 993 - 999.

M. Sorin,, J. Paris, 1997, Combined Exergy and Pinch Approach to Process Analysis. Computers & Chemical Engineering, 21, 23-28.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Water and Wastewater Management in a Petrochemical Raw Material Industry

Karla P. Oliveira-Esquerre[a], Asher Kiperstok[a], Ricardo Kalid[a], Emerson Sales[a], Lara Teixeira[a], Victor Matta Pires[b]

[a] *Federal University of Bahia (UFBA), Clean Technology Network of Bahia (TECLIM), Industrial Engineering Program (PEI), Rua Aristides Novis, 02, Federação, 40210-630, Salvador, Bahia, Brazil*
[b] *Braskem´s Raw Materials Unit (UNIB), Rua Eteno, 1561 Pólo Industrial de Camaçari, Camacari, Bahia, Brazil*

Abstract

Based on cleaner production concepts, a methodology for water use minimization has been developed by the Clean Technology Network of Bahia at one of the largest industrial complex in Latin America, in Camaçari, Bahia, Brazil. Here we present the methodological tools and results of the water and wastewater minimization program at a petrochemical plant operated by Braskem's Raw Material Unit (UNIB) at this complex. The set of tools used includes reconciled aqueous stream balances, database of aqueous streams, definition and analysis of water and wastewater indicators, as well as regional analysis in a search for more environmentally friendly water sources. Water mass balances had to be built with assumed values as poor measurement or rather lack of water measurement is a constant inside the industrial complex. Because of this, the proposed indicators had to be analyzed with caution. The studies have identified the possibility of recovering about 1,500 m^3/h of the overall 5,700 m^3/h water consumed by the industrial complex at the moment. The effluent flow rate has been reduced by 250 m^3/h since 2002. Process variability was found to be an important cause of both effluent generation in total flow rate and its continuity.

Keywords: Water Mass Balance; Environmental Indicators; Petrochemical Industry.

1. Introduction

Water and wastewater minimization techniques have been widely studied and case studies applied to petrochemical/refinery industries (Bagajewicz, 2000; Zbontar and Glavic, 2000) and water user surveys on industrial sites have been reported in the literature (Al-Muzaini, 1998; Féres et al., 2008; Kiperstok et al., 2006). A methodology for water use minimization based on cleaner production concepts has been developed by the Clean Technology Network of Bahia. This method has evolved over a ten-year period in research cooperation programs with chemical, petrochemical and copper metallurgy industries located in Camacari's Industrial Complex in Brazil.

1.1. Clean Technology Network of Bahia (TECLIM)

The Clean Technology Network of Bahia (TECLIM) was set up in 1997 by the Departments of Environmental Engineering and Chemical Engineering at the Federal University of Bahia, Brazil, to promote cleaner practices in industry. Both water and effluent minimization are parts of these objectives. Institutions taking part in this program include universities and technical schools, public offices, industrial

associations and centers for technology training, environmental agencies as well as industrial and consulting firms, and industrial water and wastewater treatment firms (Kiperstok, 2000). As a clear result of the interaction between these institutions, cleaner production attitudes have been widely recognized at the Camaçari Industrial Complex. This, together with the financial support from the Research and Development Sectorial Funds from the Brazilian Ministry of Science and Technology, has stimulated the setting up of cooperative research projects to optimize water and energy use on industrial sites. The cooperative project between TECLIM and Braskem's Raw Material Unit (UNIB) is the most long-lasting of these projects, having started in 2002.

1.2. The Industrial Complex of Camaçari and Braskem's Raw Material Unit (UNIB)

The Industrial Complex of Camaçari is one of the most industrialized area in Latin American and consumes approximately 5,700m^3/h of water to produce more than 11.5 million tons per year of primary, intermediate and final chemical and petrochemical products. The majority of the companies at the Complex are connected through a pipe network to UNIB. This company processes petroleum-by-products, mainly naphtha, into primary olefins (mainly ethylene) and aromatics. Its utility facilities also include water treatment units which produce clarified, drinking and demineralized water for internal use and for almost all the other companies at the Industrial Complex.

About 2,500 m^3/h of inorganic and organic effluent are produced by the Complex, and through two separate wastewater systems the organic effluent is directed to an activated sludge treatment system and then to a submarine outfall. This outfall also receives inorganic effluents. Major contributions come from UNIB and from a metallurgic industry with *c.a.* 350 m^3/h and 140 m^3/h, respectively. Both systems are operated by Cetrel, an environmental protection company.

2. Methodology

TECLIM's methodology includes, among others, three basic tools: water mass balances, building environmental indicators and a regional analysis in the search for more environmentally friendly water sources.

2.1. Water Mass Balances

Water mass balances were built from the larger to the more specific areas and processes. Water intake from surface and underground water reservoirs, final effluent destination and rainfall were considered in the larger control volume. Smaller internal control volumes were built for each operational unit. With the difficulties of getting quality data where identified, reconciliation techniques came as the natural step. Lack of information related to flow and quality of aqueous streams hampered the task of building the desired water balance. This difficulty was initially overcome by getting available data from process flow sheets, experience and feeling of operators and expedite measurements. Then, a quality of information record (QI) was attached to each of the values considered in the data bases (Kiperstok et al., 2006). QI values were defined from a range from 0.4 to 10.0 and enabled operators to understand how good their information about water flow rate was. Afterwards, QI values were adapted to behave as the inverse of uncertainty value which were then used to define the minimization problem for data reconciliation. Eq. (1) describes the objective function of this problem.

$$\min \sum_{i=1}^{N} \left[\left(V_{Ri} - V_{Mi}\right)^2 \cdot \frac{QI_i^{\,2}}{V_{Mi}^{\,2}} \right] \tag{1}$$

V_R and V_M are the reconciled and estimated flow rates for each stream i; N is the total number of input and output streams for each control volume.

2.2. Building Environmental Indicators

Environmental practices lead to a more comprehensive and quantitative characterization of environmental performance (Perotto et al., 2008). Thirty eight environmental indicators were proposed in the first survey exercise where eighteen aimed to cover water consumption and the other twenty to evaluate effluent generation by all UNIB plants. Here, effluent generation refers to aqueous streams contaminated by organic compounds from the processes and inorganic compounds from utilities. Water losses evaporation in cooling towers is not considered as effluent. However, it was considered in another project dealing with heat exchange optimization. A systematic and transparent approach was needed to derive these indicators which should both reflect eco-efficiency development and could be implemented with available data. Some of the proposed indicators are shown in Table 1.

Table 1: Proposed environmental indicators.

Environmental Indicator	Status
• OE * / petrochemical production • IE * / petrochemical production	Statistical analyses performed
• OE from each plant / petrochemical production • IE from each plant / petrochemical production	Effluent data not available at the moment
• Clarified water / petrochemical production • Demineralized water / petrochemical production	Statistical analyses are not reliable because water flow rate are poorly measured

* OE and IE: organic and inorganic effluent flow rates, respectively.

The data records were analyzed on a monthly basis going back over a seven-year period, 2001 to 2007. Time series plots included the indicator and its respective raw data. Plant maintenance periods were highlighted in these graphs and were correlated with water and wastewater flow rate variations. Figure 1 shows a close view of UNIB's total organic effluent per petrochemical production indicator. The thicker the horizontal line, the greater the expected impact of the maintenance procedures on effluent generation.

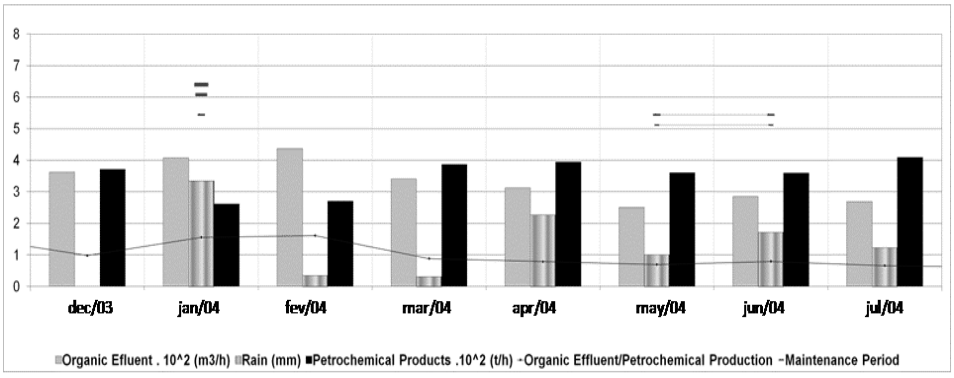


Figure 1: UNIB's total organic effluent per petrochemical production indicator.

Since UNIB's drainage effluent system is mainly composed of open channels and the site is located in a tropical region runoff contribution was estimated and discounted from both organic and inorganic effluents flow rates.

2.3. Considering Regional Analysis Seeking in the search for more Environmentally Friendly Water Sources

Two alternative centralized water reuse options were investigated while water conservation practices were partially proposed. Harvesting stormwater from three wastewater reservoirs built to operate as accumulation basins for pumping inorganic effluents and rain water into a submarine outfall is one of the options (Oliveira-Esquerre et al., 2008). The Complexo Básico, Bandeira and Cobre basins show a reserve capacity of *c.a.* 1,600, 800 and 400 thousand m^3, respectively. A mass balance was performed with historical data to assess water availability from all three basins based on the daily volume variation and inorganic effluent inflow from 2001 to 2007. Even though stormwater is usually of better quality than industrial discharges (Mitchell et al., 2002), few cases of industrial stormwater use have been reported in the literature (Thomas et al., 2002).

Another option considers reusing groundwater removed by a pump-and-treat system operated in the Industrial Complex and designed to protect both the São Sebastião Aquifer and surface water sources. This system is composed of fourteen wells strategic located and screened at different levels to groundwater extraction. Identification, assessment and remediation of large-scale groundwater characteristics require a profound knowledge of the conceptual model to predict the water quality in the subsurface (Wysic et al., 2003). To identify reuse alternatives, time series and dot plots were used to evaluate pumping flow rates, water quality levels and their variability.

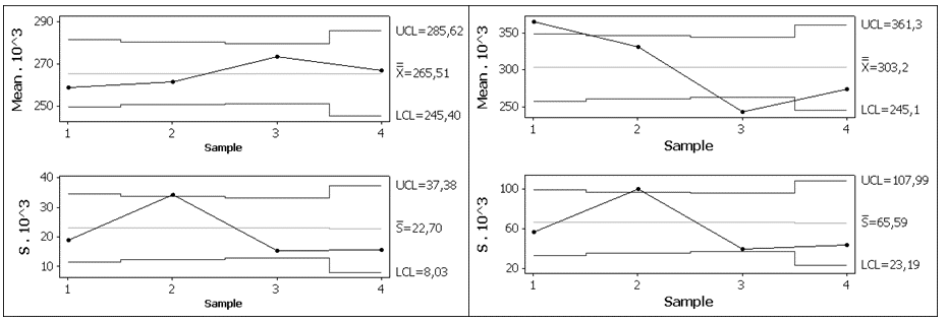
3. Discussion and Results

3.1. Water Mass Balances

In total 1,400 flow rates are considered in the water mass balance, where about 20% were not registered by UNIB information system. Although water use and wastewater generation minimization tools are well known, lack of quality data in plants, limited the contribution that this tool could offer. Braskem's effluent flow rate has been falling in a reasonably constant pattern over recent years, both in absolute and relative numbers. The activities related to the quantification of the aqueous streams in reconciled mass balances and the search for reduction opportunities among the plants' collaborators may account for the reduction in specific waste water generation by more than 40% in less than two years. According to Smith and Petela (1992), traditional approaches to water minimization such as changing washing operations, when complemented with water pinch analysis methodology have been shown to achieve 30% to 60% fresh water savings in industrial applications. Further work is needed to identify means to reduce the contributions to the inorganic system both at the source and outside production limits. After reconciliation, the balances will be transferred to the operational areas and updated with online data from existing flow rate metering and manual reports. Reconciled water balances and the efforts to maintain them up to date partially fulfill the need to enable operators to know how water is used inside the industrial plants. This knowledge is considered to be a basic requirement to enhance water conservation practices at source. It also provides basic support for the identification of reuse and recycling opportunities closer to the processes.

3.2. Environmental Indicators

No reliable analysis of water indicators could be carried out because of either the lack of water metering or the low QI of the measurements. The absence of adequate water metering may be due to a persistent regional view of water as an infinite and low cost resource. In the case of the UNIB's inorganic and organic effluents, the flow rates are measured by Parshall flumes. The estimated runoff contribution areas for the organic and inorganic effluents were 0.7 and 4.0 km2, respectively. Time series plots showed a reduction in the petrochemical production variability and organic effluent flow rate. The cause effect relationship was confirmed by comparing the biannual mean (X-bar) and standard deviation (S) for both measurements (Figure 2).



(a) (b)

Figure 2: Control charts (X-bar, S) of the (a) petrochemical production (t/month) and (b) organic effluent (m^3/h).

3.3. Considering Regional Analysis Seeking for more Environmental Friendly Water Sources

The use of both water accumulated in the attenuation basins and the extracted water from the pump-and-treat system means the decentralization of industrial water withdrawal at the Industrial Complex. Time series and frequency data indicate an averaged annually water recovery from Complexo Básico, Cobre and Bandeira basins of 1,140, 870 and 440 m^3/h, respectively. For the pump-and-treat system, 50 to 120 m^3/h of extracted water may be used. These waters have been considered for use as cooling water after simple dilution. Another option is to incorporate these new sources into the UNIB's water treatment input. This option requires the segregation of potable water production which is easily feasible considering the system's characteristics. Any 1000 m^3/h of water from new sources represent a reduction of 17% in the actual industrial water withdrawn from regional resources.

4. Conclusion

This paper describes some results and methodological tools used in a cooperative project proposed by the Clean Technology Network for water and wastewater minimization at a petrochemical plant. An understanding of the water flows throughout water mass balance development, together with other initiatives on the part of the company, have contributed to a reduction of about 40% of total effluent generation in the first three years of the studies. A systematic and transparent approach was needed to derive indicators which on the one hand reflect the concerns of eco-efficient development and on the other are defined in terms of data requirements. This approach

has enabled, among others, the identification of the process variability as a relevant cause of both organic effluent flow rate and variability. Lack of quality data in current plants limited both the development and the interface of mass balance and indicator studies. New water sources were proposed here taking advantage of wastewater streams which indicate a reduction of at least 17% in extraction from regional water sources. This has stimulated the development of engineering projects and the proposal of new operational procedures that guarantee the minimizing of water use and wastewater generation as well as discharge.

Acknowledgment

The authors gratefully acknowledge the financial support given by FINEP (State Agency for Innovation) and CNPq the Brazilian Agency for Scientific and Technological Development; Braskem/UNIB for the financial and technical support and, Cetrel S.A. for the technical support. Special thanks go to the researchers of the Ecobraskem Project, Mario Cezar Mattos, Leonardo Souza, Elicelma Carvalho and Márcio Martins, and the under graduate student Carlos Mendes.

References

A. Kiperstok, A., 2000, Implementation of Cleaner Production Practices with the Support of a Diploma Course, J Cleaner Prod, 8: 375-379.

A. Kiperstok, R. Kalid, E. Sales, 2006, Development of Water and Wastewater Minimization Tools for the Process Industry: The Experience of the Clean Technology Network of Bahia, Brazil. Proc. Global Conf. on Sustainable Product Development and Life Cycle Engineering, 4, 3-6, São Carlos, SP, Brazil.

E. Perotto, R. Canziane, M. Renzo, P. Butelli, 2008, Environmental Performance, Indicators and Measurement Uncertainty in EMS Context: A Case Study, J Cleaner Production, 16: 517-530.

J.-S. Thomas, J. Le Pol, S. Gillet, L. Phan, 2002. SIRRUS: Evaluation Software of the Technical and Economic Profitability of Rainwater on an Industrial Site. Water Sci Technol, 46(6-7):71-76.

K.P. Oliveira-Esquerre, A. Kiperstok, E. Cohim, R. Kalid, E. Sales, V.M. Pires, Taking Advantage of Storm and Waste Water Retention Basins as Part of Water Use Minimization in Industrial Sites, Resources, Conservation and Recycling J, submitted on August 2008.

M. Bagajewicz, 2000, A Review of Recent Design Procedures for Water Networks in Refineries and Process Plants, Comp Chem Eng, 24(9):2093-113.

P. Wycisk, H. Weiss, A. Kaschl, S. Heidrich, K. Sommerwerk, 2003, Groundwater Pollution and Remediation Options for Multisource Contaminated Aquifers (Bitterfeld/Wolfen, Germany), Toxicology Letters, 140 -141:343-351.S.-H.

S. A1-Muzaini, 1998, Industrial Wastewater Management in Kuwait, Desalination, 115: 57-62.

V. G. Mitchell, R.G. Mein, T.A. McMahon, 2002. Utilising Stormwater and Wastewater Resources in Urban Areas, Australian J Water Resources, 6(1):31-43.

Z. Zbontar, P. Glavic, 2000, Total Site: Wastewater Minimization, Wastewater Reuse and Regeneration Reuse, Resources, Conservation and Recycling, 30: 261-275.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Synthesis of Energy efficient Complex Separation Networks

Seon B. Kim, Gerardo J. Ruiz, Jeonghwa Moon, Libin Zhang, and Andreas A. Linninger

Departments of Chemical Engineering and Bioengineering, University of Illinois at Chicago Laboratory for Product and Process Design. Chicago, IL 60607, USA

Abstract

Separation processes account for about fifty percent of capital and operating costs including the highest energy demand in the chemical industry. The rise in energy consumption, the high cost of the fuels, and combined with the environmental impact increases the demand for energy saving separation processes such as optimal complex column networks which are estimated to achieve energy savings of up to 70%. This paper introduces a novel algorithm to create optimal complex column arrangements which encode the cost and states of global solutions with minimum user input. In the complex column networks, several combinations and internal connections increase the difficulty to optimize these processes. The proposed configuration algorithm generates all possible network configurations expressed as a continuous sequence of column section profiles and discriminates suboptimal solutions taking the column network structure as input. Simultaneously, a robust feasibility test is applied to the generated systematic network based in thermodynamic transformations called temperature collocation where the operating conditions, structure, and length of the separation network for realizing the desired product purities are not predefined. This robust hybrid algorithm combines the advantages of deterministic and stochastic search techniques.
This computational approach guarantees rigorous column profiles validated with industrially accepted simulation software such as ASPEN. The capabilities will be illustrated using multicomponent realistic case studies especially quaternary systems.

Keywords: Complex Column, Temperature Collocation, flowsheet simulator

1. Introduction

Chemical distillation is one of the most predominantly used and versatile method of separation, in the petrochemical and commodity industry. It continues to occupy 40-70% of capital and operating costs of chemical manufacturing. Distillation processes account for more than 60% of the total process energy for the manufacture of commodity chemicals (DOE 2005).
Previously energy consumption in separations was a minor important factor, because of low energy prices and less stringent environmental standards. Thus, mainly simple distillation column configurations were built. However the rising energy cost and concerns over atmospheric carbon emissions redefine the design objectives for industrial separations with a new focus on energy conservation and the emission reduction associated with it. In the past ten years, several research groups have advanced systematic methods for identifying optimal distillation network.
Recently, computer-aided separation synthesis methods for energy savings in distillation have been revisited critically. Specifically it was demonstrated that

substantial energy savings can be achieved by designing the network simultaneously. These complex column configurations have the potential of achieving up to 70% energy savings over simple column networks (Hilde K. Engelien 2005). The well considered Petlyuk complex column configuration is an excellent example to realize energy efficient separations.

Complex column configuration is a kind of mathematical model which has problems for two reasons. First, problem size increases as the number of stages approximates infinity close to singular points of the composition space (e.g. saddle, pinch points). Secondly, in column design problems it is not known aprior whether a set of desired specifications has a feasible solution. In case of an infeasible specification, all direct solution methods don't converge. Hence, we have developed a set of collocation based methods specially for its use in column design including simple and complex column configuration. The method proposes variable transformations which enable it to successfully solve columns design problems even near pinch and saddle points with acceptable accuracy. We also present a new algorithm to test the feasibility of a distillation network. The key novelty of our procedure exploits a view of physically meaningful design variable to distinguish between feasible region and unattainable composition range. The method is shown to be effective in constant relative volatility, idea and even non-ideal equilibrium models.

2. Methodology

Our approach towards a general solution of the separation synthesis problem has three major components: (i) systematic network generation, (ii) thermodynamic problem transformation to tame numerical challenges and reduce size, and (iii) advanced algorithm engineering to seek simultaneously in the structural and parametric design space.

2.1. Generating Complex Column Configurations

Complex distillation configurations have been studied previously by Agrawal and Fidkowski (Agrawal and Fidkowski 1998; Agrawal 2003). They systematically evaluated the energy consumption of basic feasible configurations of complex networks and advocate the energy savings that can potentially be achieved through feasible thermal coupling. Another important advancement is Agrawal's procedure for the exhaustive enumeration of complex flowsheets, which always contains the network structure of the most energy efficient solution. This complete synthesis of candidate flowsheet structures is an important element in a fully automatic separation synthesis methodology. To complete the synthesis, the optimal operating parameters like reflux, intermediate product purities, column height and diameter for all the candidate networks in the superstructure would have to be determined.

We propose to combine network generation methods like Agrawal's with robust global search to systematically identify optimum distillation schemes for given product purity targets. The solution seeks to determine both structural as well as parametric design decisions. Unfortunately, the resulting synthesis problem is NP-hard and suffers from challenges associated with the thermodynamic topology of the feed mixture which cannot be anticipated a priori for multi-component streams with different vapor-liquid interactions. At present, our evolving design method comprises two stages:

Step 1. Generic structure synthesis: Generate all possible basic network configurations using only structural information thus yielding a superstructure containing all potentially optimal networks for the separation of a given feed. The superstructure could be enumerated exhaustively or incorporate equipment constraints when addressing retrofit problems with existing column inventory.

Step 2. Network Task Optimization: Rigorous optimization of all candidate solutions in their respective parametric design spaces to find the globally optimal structure. Network optimization task should take into account the specific feed composition and product targets as well as the specific thermodynamic properties of the chemical species to be separated. In a brute force approach to the synthesis problem, one could advantageously solve independent optimization sub-problems for each network in parallel on separate processors. More elegant would be deployment of thermodynamically motivated implicit pruning criteria, which would allow us to discard inferior designs without explicitly performing network optimization. Such an *admissible heuristics*, however, have not been developed for complex column networks.

2.2. Continuous Complex Column Model

Complex columns have multiple feed and product steams requiring some modifications to profile computations. Fortunately, Hildebrandt and Glasser worked out a continuous model called difference point equations model, valid for complex column sections in any possible network configuration (Tapp, Holland et al. 2004). A complex column section is defined as a column segment between feed or product trays. Eq. (1) expresses the liquid composition profiles of a complex section in terms of the generalized reflux R_Δ and the difference point composition $X_{\Delta i}$. It also applies to simple sections as a special case, thus rendering a single mathematical expression for all sections in any separation network composed of complex or simple columns. Columns – simple or complex – are said to be feasible if and only if every pair of composition profiles belonging to adjacent column sections intersect. A network is feasible, if all its columns are feasible.

$$\frac{dx_i}{dn} = \left(1+\frac{1}{R_\Delta}\right)(x_i - y_i) + \frac{1}{R_\Delta}(X_{\Delta i} - x_i) \qquad (1)$$
$$R_\Delta = L/\Delta \quad X_{\Delta i} = (Vy_i - Lx_i)/\Delta$$

Where x_i is the liquid composition, y_i is the vapor composition, n is the number of stages, and $\Delta = V - L$ is the flow rate difference point. Unfortunately, in higher dimensional space of multicomponent mixtures, it is not a simple task to ascertain the intersection of column section profiles with modest computational effort. We discovered that a thermodynamics transformation of the column stage number, n, into the equilibrium bubble point temperature, T. This thermodynamically motivated transformation offers massive size reductions in the column profile computations and is a key feature in a rigorous feasibility criterion based on minimum bubble point distance functions which eases the search.

2.3. Bubble-Point Distance(BPD)

We eliminate the tray number in favor of the bubble point temperature as independent variable. Theoretically, this transformation requires the implicit function theorem for a one-to-one relationship between column height and temperature; this assumption may not hold in highly non-ideal mixtures, but has been tacitly assumed for well-behaved separation tasks we have so far investigated. Applying the thermodynamic transformation called *temperature collocation* changes the continuous column profile equations in eq. (1) to a new composition profiles with temperature instead of height as new independent integration variable like shown in eq. (2)

$$\frac{\partial x_i}{\partial T} = -\left(\left(1+\frac{1}{R_\Delta}\right)(x_i - y_i) + \frac{1}{R_\Delta}(X_{\Delta i} - x_i)\right) \frac{\sum_{i=1}^{c}\left(\frac{\partial K_i}{\partial T} x_i\right)}{\sum_{i=1}^{c}\left[\left(\left(1+\frac{1}{R_\Delta}\right)(x_i - y_i) + \frac{1}{R_\Delta}(X_{\Delta i} - x_i)\right) K_i\right]} \qquad (2)$$

Design for distillation process relies on the specifications of distillate composition and reflux ratio which make it feasible. For the binary and ternary mixtures the reflux ratio can be found within some ranges obviously with the existing graphical methods. The *bubble-point distance* (BPD) is defined as the Euclidean distance between two equilibrium compositions belonging to the stripping and rectifying profiles with the same bubble-point temperature. In the case that there exists a pair of compositions with BPD = 0, it guarantees feasible design and other cases vice versa.

Once compositions profiles have been calculated as functions of temperature, the column height can easily be recovered with the help of eq. 3. The calculation of the column height from bubble point temperatures has been used in the case studies to estimate the capital cost.

$$\frac{dn}{dT} = \frac{\partial n}{\partial x_i}\frac{\partial x_i}{\partial T} = -\frac{\sum_{i=1}^{c}\left(\frac{\partial K_i}{\partial T}x_i\right)}{\sum_{i=1}^{c}\left[\left(\left(1+\frac{1}{R_\Delta}\right)(x_i - y_i) + \frac{1}{R_\Delta}\left(X_{\Delta i} - x_i\right)\right)K_i\right]} \tag{3}$$

To apply this method onto complex column, this analysis should be applied for the adjacent sections *i* and *j* in complex column *k* simultaneously. According to this generalization of the minimum bubble point distance approach, a complex column is feasible if the sum of all profile distances of all adjacent sections *i* and *j* at column *k* in the network is within a small tolerance (eq. 4). It is worth pointing out that the aggregated bubble point distances for a complex column is still a scalar delineating the vicinity to a realizable design. The ability to judge the closeness of a design to realizable specifications is of key importance in the genetic search deployed successfully for the synthesis of simple configurations (Zhang and Linninger 2006) . The entire network is feasible if all its columns are feasible as in eq. 5. Where *N* is the total number of column sections in each column *k*.

$$Z(k) = \sum_{i=1}^{N-1} BPD(i, i+1) < \varepsilon_1 \tag{4}$$

$$\Psi(k) = \sum_{k=1}^{K} Z(k) < \varepsilon_2 \tag{5}$$

3. Case Study

In mixtures with four or more components, the design problem involves the search for one or more compositional degree of freedom of any product stream. This multidimensional global optimization problem is solved using a hybrid method that combines stochastic genetic algorithm (i.e. product specification) with rigorous finite element collocation of column profiles (Zhang and Linninger 2006).

3.1. Separation of Quaternary Mixture

First example illustrates the potential energy saving of complex column configuration over simple column sequences for the separation of a quaternary mixture of methanol, ethanol, 1-propanol, and acetic acid. This complex network uses two simple columns and one complex column as shown in *Figure 1*(a). The composition profiles diagram for column I and column III are showed in *Figure 1*(b) and (c). The dimension (height) and minimum vapor rate (V_{min}) for all columns are collected on Table 1. We also estimated the total cost of the complex network. The complex configuration analyzed uses half total vapor rate and saves 20% of total costs compared to the simple column

configuration. This case study illustrates a feasible solution to the synthesis problem using a complex network, and potential energy savings realizable with complex columns.

Table 1. Height and V_{min} for each column in the network

	Col. I	Col. II	Col. III
Height (m)	10.9	26.6	30.5
V_{min} (kmol/h)	13.2	58.9	71.3

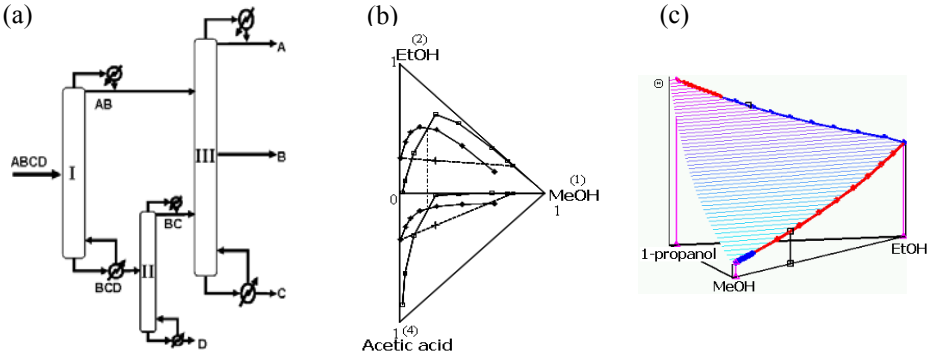


Figure 1. (a) Complex column configuration, (b) The quaternary simple column profile, and (c) Temperature Collocation profile for ternary complex column

3.2. Initialization of complex distillation networks with Aspen HYSYS

Figure 2(a) shows the complex network used to separate the same quaternary mixture of the first case study. The composition profiles found with *Aspen* HYSYS are indicated in *Figure 2*(b). *Figure 2*(c) shows the composition profiles using temperature collocation transformation. These profiles as well as the equivalent number of equilibrium stages were computed using eq. (3).

We initialize the *Aspen* HYSYS flowsheet with our column profiles as initial guesses. The *Aspen* simulation converged with only one or two iterations. This result confirms the accuracy of our methodology is acceptable for a rigorous simulator. In a similar approach we anticipate all separation networks synthesized with our methodology will be easily verified using automatic interface with *Aspen*, this will be a very useful tool for implementing this design idea in the industrial practice.

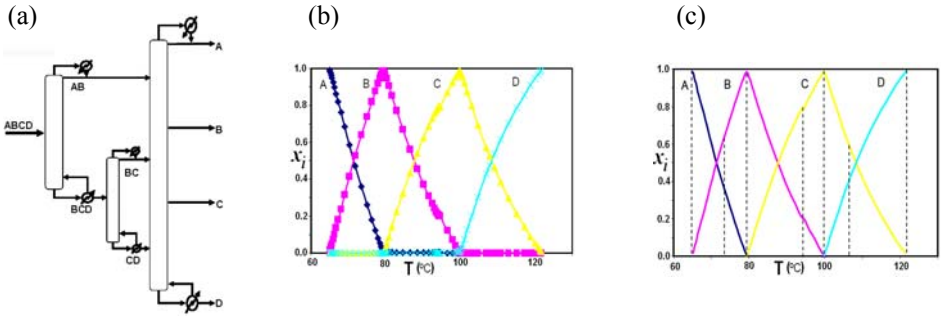


Figure 2. Complex column configuration (a), Aspen HYSYS simulation profiles for tray temperature (b), and Temperature Collocation profiles (c)

4. Conclusions

In this article, we advocated the recent progress for synthesizing energy efficient complex distillation networks. An approach towards a generalized solution of the separation synthesis problem is presented. Difference point equation with temperature as an independent variable instead of tray number was employed to find the composition profiles of complex column sections. Temperature collocation and minimum bubble point distance (MIDI) algorithm were effective to find a feasible separation by intercepting profiles. The first case study demonstrates the potential to 50% of energy savings using a complex column network compared to the simple column configuration. The second case study demonstrates current state of the art of separation synthesis in conjunction with computer simulations to fully integrate complex separation networks. Finally, the seamless integration of rigorous flowsheet simulators to validate the predictive results of our scientific method was demonstrated.

Acknowledgements

Financial support by DOE Grant: DE-FG36-06GO16104 is gratefully acknowledged. We also like to thank Dr. Agrawal's helpful discussion on complex column. We acknowledge Dr. Chau-Chyun Chen for providing an Aspen software research license.

References

Agrawal, R, 2003, Synthesis of multicomponent distillation column configurations, AIChE J, 49, 2, 379-401.

Agrawal, R. and Z. T. Fidkowski, 1998, Are thermally coupled distillation columns always thermodynamically more efficient for ternary distillations?, I&EC Research, 37, 8, 3444-3454.

DOE, 2005, Technical Topic Description, www.doe.gov.

Hilde K. Engelien, S. S, 2005, Minimum energy diagrams for multi-effect distillation arrangements, AIChE J, 51, 6, 1714-1725.

Tapp, M., S. T. Holland, et al., 2004, Column Profile Maps. 1. Derivation and Interpretation, I&EC Research, 43, 2, 364-374.

Zhang, L. and A. A. Linninger, 2004, Temperature collocation algorithm for fast and robust distillation design, I&EC Research, 43, 12, 3163-3182.

Zhang, L. and A. A. Linninger, 2006, Towards computer-aided separation synthesis, AIChE J, 52, 4, 1392-1409.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Comparative Study of Gibbs Free Energy Minimization in a Real System using Heuristic Methods

C. E. Figueira[1], R. R. Soares[2], F. S. Lobato[3], V. Steffen Jr[4]

[1]milaemif@yahoo.com.br, [2]rrsoares@ufu.br,
[3]fslobato@mecanica.ufu.br, [4]vsteffen@mecanica.ufu.br
[12]*School of Chemical Engineering,* [34]*School of Mechanical Engineering*
Universidade Federal de Uberlândia, Av. João Naves de Ávila 2121, Campus Santa Mônica,P.O. Box 593, 38400-902, Uberlândia – MG, Brazil.

Abstract
The knowledge of phase and chemical equilibrium is clearly important in the analysis and design of a wide variety of chemical processes. These problems are described mathematically by complex systems due to the nonlinear nature of the thermodynamic models. This equilibrium state is characterized by the minimization of the Gibbs free energy. Usually, this problem is calculated by using the method of Lagrange multipliers together with the mass balance conditions as necessary subsidiary conditions. However, the convergence is highly dependent of initial estimates of the Lagrange multipliers. The development of robust and efficient methods for the computation of thermodynamic multi-phase systems has long been a challenge in both chemical engineering and materials science. In this context, both heuristic methods, Genetic Algorithms and Differential Evolution, are able to escape from local minima and saddle points. As main disadvantage the high number of objective function evaluations as compared with classic methods can be mentioned. To overcome this disadvantage, this work proposes the dynamic updating of the population size to reduce the number of objective function evaluations. The methodology is applied to determine phase compositions at the chemical equilibrium of real systems. The results demonstrated that the methodology used represents a promising alternative for the problem studied.

Keywords: Gibbs free energy minimization, real systems, heuristic methods, dynamic population.

1. Introduction

The phase and chemical equilibrium problem is extremely important for predicting fluid phase behavior for a very large number of separation process applications, such as reactors, reactive distillation columns, azeotropic and three-phase distillation. Moreover, the chemical equilibrium problem can also predict the maximum conversion as well as the selectivities of pre-fixed reactants and products molecules sets. Process simulators need to be able to reliably and efficiently predict the correct number of phases that will exist at equilibrium, and the distribution of components within those phases (McDonald and Floudas, 1995). In this sense, the thermodynamic function of foremost interest is the Gibbs free energy as it can be used to describe equilibrium at conditions of constant temperature and pressure. A global minimum of the Gibbs free energy corresponds to the true equilibrium solution (Castillo and Grossman, 1981).
The main difficulties in solving Gibbs free energy are the selection of the number of phases and the initial conditions used as starting point in the optimization solver. Traditionally, this problem has been solved by using the method of Lagrange multipliers. However, the convergence is highly dependent on initial estimates of the Lagrange multipliers (Castillo and Grossman, 1981; Xiao *et al*, 1989; McDonald and Floudas, 1995). Other methodologies can be found in the literature, such as the one by Ohanomah

and Thompson (1984), where the authors used the Newton method for the computation of multiphase equilibria. Lantagne *et al.* (1988), on the other hand, presented a mixed penalty function method for the computation of complex equilibria.
To overcome the difficulty associated with the initial estimates, non-deterministic solvers have been employed for this purpose (Lin and Miller, 2004; Srinivas and Rangaiah, 2007). However, theses methodologies use fixed population sizes during the evolutionary process that can affect the robustness and the computational cost of the algorithms associated with these methods. Small population size may result in local convergence; large population size will increase computational efforts and may lead to slow convergence. So, an appropriate population size can assure the effectiveness of the algorithm (Vellev, 2008).
This work is organized as follows. Section 2 presents the general aspects regarding the formulation of the Gibbs free energy minimization. A review about Genetic Algorithms and the Differential Evolution Approach and its extension to deal with dynamic updating of the population size are presented in Sections 3 and 4, respectively. In Section 5, the proposed methodology is applied to a real system. Finally, the conclusions are outlined in Section 6.

2. Gibbs Free Energy Minimization

For a multi-component and multi-phase system at temperature T and pressure P, the Gibbs free energy function *Gibbs* can be expressed in terms of fugacity and Gibbs free energy of formation:

$$Gibbs = \sum_{i=1}^{N}\sum_{k=1}^{\pi} n_{ik}\left(\Delta G_{ik}^{f} + RT\ln\frac{f_{ik}}{f_{ik}^{o}}\right) \tag{1}$$

where ΔG_{ik}^{f} is the Gibbs free energy of formation of component i in phase k at T, R is the gas constant, f_{ik} is the partial fugacity of component i in phase k and f_{ik}^{o} is the fugacity of pure component i at the standard state.
At equilibrium, *Gibbs* should be the global minimum with n_{ik} satisfying mass balances. For phase equilibrium, the conservation of moles of the individual component must hold:

$$\sum_{k=1}^{\pi} n_{ik} = n_{iT}, \quad i = 1,\ 2,\ ...,\ N \tag{2}$$

where n_{iT} is the total number of moles of component i in the system (feed).
For simultaneous phase and chemical equilibrium, the conservation of chemical elements must hold:

$$\sum_{i=1}^{N}\sum_{k=1}^{\pi} a_{ji} n_{ik} = b_{j}, \quad j = 1,\ 2,\ ...,\ M \tag{3}$$

where a_{ji} represents the number of gram-atoms of element e in component i, b_j is the total number of gram-atoms of element j in the system and M is the number of elements.
Non negativity conditions can be satisfied. For phase equilibrium:

$$0 \le n_{ik} \le n_{iT}, \quad i = 1,\ 2,\ ...,\ N;\ \ k = 1,\ 2,\ ...,\ \pi; \tag{4}$$

For simultaneous phase and chemical equilibrium:

$$0 \le a_{ji} n_{ik} \le b_{j}, \quad i = 1,\ 2,\ ...,\ N;\ \ k = 1,\ 2,\ ...,\ \pi;\ \ j = 1,\ 2,\ ...,\ M; \tag{5}$$

Thus, to obtain phase compositions at equilibrium it is necessary to find the minimum of the Gibbs free energy as given by Eq. (1) subject to the constraints imposed by Eq. (2) and Eq. (4) (or Eq. (3) and Eq. (5) for chemical equilibrium).

3. Heuristic Methods

3.1. Genetic Algorithm

Genetic Algorithms (GA) are probabilistic search algorithms combining the mechanics of natural selection and survival of the fittest. These algorithms are capable of efficiently finding an optimal solution for complex problems without necessitating reformulation for the evaluation of individual solution candidates. In this approach, a set of feasible designs constitutes a generation, which has a fixed number of individuals. A set of better designs is derived from the previous generation where the individuals are allowed to reproduce and cross among themselves with bias allocated to the fittest members. Combinations of the most favorable characteristics of the mating members of the population result in a new generation that is more fit than the previous one. The GA is implemented with three basic operations: reproduction, crossover and mutation. Besides, the main control parameters of the algorithm implemented are the number of individuals in the population, probability of crossover, probability of mutation and number of generations.

3.2. Differential Evolution

Differential Evolution (DE) is a recent optimization technique in the family of evolutionary computation. It is proposed as a variant of genetic algorithms to achieve the goals of robustness in optimization and faster convergence to a given problem. DE differs from other evolutionary algorithms in the mutation and recombination phase. Unlike some meta-heuristic techniques such as genetic algorithms and evolutionary strategies, where perturbation occurs in accordance with a random quantity, DE uses weighted differences between solution vectors to perturb the population (Storn and Price, 1995).

The crucial idea behind DE is a new scheme of mutation. DE executes its mutation by adding a weighted difference vector between two individuals to a third individual. Then, the mutated individuals will do discrete crossover and greedy selection with corresponding individuals of last generation to produce offspring. There are three strategy parameters in DE, i.e., population size, crossover rate and perturbation rate.

4. Update of Parameters in Heuristic Methods

All classical selection algorithms keep the population size fixed during the evolutionary process. This aspect simplifies the algorithms but it is an artificial restriction and does not follow any analogy to the biological evolution, where the number of individuals in a population varies continuously with time, increasing when there are highly-fitted individuals and abundant resources, and decreasing otherwise. Intuition hints that it may be beneficial for the population to expand in the early generations when there is high phenotype diversity and there is opportunity to "experiment" with different characteristics of the individuals, and to shrink with the increase of population convergence, when the unification of the individuals in terms of structure and fitness no longer justifies the maintenance of a large population and the higher computational costs associated with it (Vellev, 2008).

In this context, Sun *et al* (2007) studied the influence of two strategies for the dynamic updating of the individuals number of the population during the evolutionary process:

$$NP = \left(NP_{\max} - NP_{\min}\right)\left(\frac{Iter_{\max} - iter}{Iter_{\max}}\right) + NP_{\min} \tag{6}$$

$$NP = \max\left(\frac{NP_{\max}}{2}\sin\left(\frac{iter}{A}\right) + \frac{NP_{\max}}{2}, NP_{\min}\right) \tag{7}$$

where $Iter_{max}$ is the generation maximum number, *iter* is the current generation, NP_{min}, and NP_{max}, are the minimal number and maximum number of individuals in the population, respectively and A is the amplitude in *sin* function.

In this work, an adaptive population size method with partial increasing or decreasing number of individuals according to diversities in the end of each generation is adopted. Besides, initially the convergence rate is defined as:

$$r = \frac{f_{average}}{f_{worst}} \tag{8}$$

where $f_{average}$ is the average value of the objective function and f_{worst} is the worst value of the objective function.

Consequently, the defined convergence rate evaluates the homogeneity of the population in the evolutionary process: if r is close to zero, i.e., the value "worst" of the objective function is different of the average value of the objective function; if r is close to one, the population is homogeneous. Thus, a simple equation for dynamic updating of the population size is proposed:

$$NP = \text{round}(NP_{\min} r + NP_{\max}(1-r)) \tag{9}$$

where NP_{min} and NP_{max} represent the minimum and maximum values for the population size, respectively, and the operator round(.) indicates the rounding to the nearest integer.

It should be emphasized that the equation (9) updates the population size as based on the convergence rate, differently from equations (6) and (7) in which any information regarding the evolution of the process is considered.

The next section presents the application to a real system using Genetic Algorithm (GA), Adaptive Genetic Algorithm (AGA), Differential Evolution (DE) and Adaptive Differential Evolution (ADE).

5. Numerical Results

A classical example of complex chemical equilibrium calculations involving two phases is the esterification reaction in an equimolar mixture of ethanol (EtOH) and acetic acid (HAc) reacting reversibly to form ethyl acetate (EtAc) and water (H_2O) (Castillo and Grossman, 1981; Xiao *et al* 1989; McDonald and Floudas, 1995; Peng Lee *et al*, 1999):

$$\text{EtOH (1)} + \text{HAc (2)} = \text{EtAc (3)} + \text{H}_2\text{O (4)} \tag{10}$$

Gibbs free energy of formation of a component in liquid state is related to the one in vapor state (ΔG_{iV}^{f}) by (McDonald and Floudas, 1995):

$$\Delta G_{iL}^{f} = \Delta G_{iV}^{f} + RT \ln\left(P_i^s\right) \tag{11}$$

where P_i^s is the saturated vapor pressure for pure component i at T. R is the gas constant and the parameters in Antoine equation for calculating P_i^s are given in Table 1.

Table 1: Data and parameters used in case study (McDonald and Floudas, 1995).

	ΔG_{iv}^{f} (cal/mol)	τ_{ij} in the NRTL model				Antoine constants		
	355 K	EtOH	HAc	EtAc	H_2O	a_i (-)	b_i (K)	c_i (K)
EtOH	-37091.8	0.0000	1.3941	0.6731	-0.2019	9.95614	1440.52	-60.44
HAc	-87303.1	-1.0182	0.0000	0.0070	-0.4735	9.68450	1644.05	-39.63
EtAc	-79840.6	0.1652	0.5817	0.0000	1.7002	9.22298	1238.71	-56.15
H_2O	-54093.4	2.1715	1.6363	1.9257	0.0000	10.0917	1668.21	-45.14

$\alpha_{ij} = 0.3$ (with α_{ii}=0) and Antoine constants in $\log_{10}(P_i^s) = a_i - b_i/(T + c_i)$.

Equality constraints arising from the conservation of chemical elements expressed by Eq. (3) for the esterification of 0.5 mol of ethanol with 0.5 mol of acetic acid are:

$$2(n_{11} + n_{12}) + 2(n_{21} + n_{22}) + 4(n_{31} + n_{32}) + 0(n_{41} + n_{42}) = 2 \tag{12}$$

$$6(n_{11} + n_{12}) + 4(n_{21} + n_{22}) + 8(n_{31} + n_{32}) + 2(n_{41} + n_{42}) = 5 \tag{13}$$

$$1(n_{11} + n_{12}) + 2(n_{21} + n_{22}) + 2(n_{31} + n_{32}) + 1(n_{41} + n_{42}) = 1.5 \tag{14}$$

for the conservation of carbon, hydrogen and oxygen, respectively.

Non-Random Two Liquid (NRTL) equation (Prausnitz *et al*, 1986) for the activity coefficient of component i in liquid phase k (γ_{ik}) is given by:

$$\ln(\gamma_{ij}) = \frac{\sum_{j=1}^{N} \tau_{ji} G_{ji} x_{jk}}{\sum_{j=1}^{N} G_{ji} x_{jk}} + \sum_{j=1}^{N} \frac{G_{ji} x_{jk}}{\sum_{l=1}^{N} G_{lj} x_{lk}} \left(\tau_{ji} - \frac{\sum_{l=1}^{N} \tau_{lj} G_{lj} x_{lk}}{\sum_{l=1}^{N} G_{lj} x_{lk}} \right) \quad (15)$$

where x_{jk} is the mole fraction of the component j in the phase k, τ_{ij} are the non-symmetric binary interaction parameters (with τ_{ii}=0) and G_{ij} is a parameter defined as:

$$G_{ij} = \exp(-\alpha_{ij} \tau_{ij}) \quad (16)$$

Parameters used in algorithms: NP_{min}=10 (AGA, ADE), NP_{max}=50 (GA, AGA, DE, ADE), perturbation rate and crossover probability (0,8), mutation probability (0.01), $Iter_{max}$=1000 generations. Finally, the stopping criterion is associated to the difference between the results calculated by the algorithms used and the one obtained by Peng Lee *et al* (1999): this difference should be smaller than 10^{-8}.

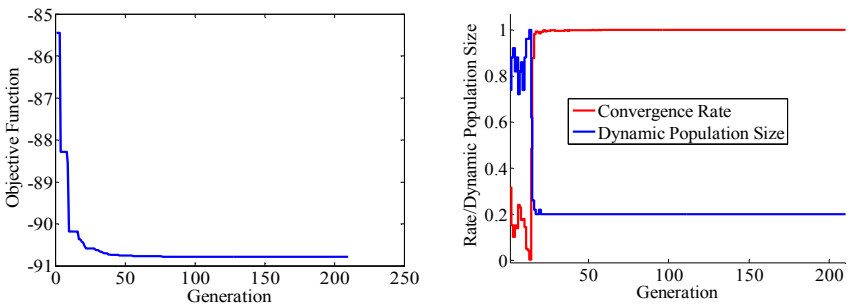
Table 2 shows the results obtained for liquid (x_i) and gas (y_i) phase using the Luus-Jaakola Algorithm (Peng Lee *et al*, 1999), GA, AGA, DE and ADE algorithms. All the algorithms were executed 100 times for obtaining of average values.

Table 2: Results obtained by LJ, GA, AGA, DE and ADE algorithms.

	LJ	GA	AGA	DE	ADE
x_1	0.039748	0.039747	0.039747	0.039748	0.039747
x_2	0.201850	0.201862	0.201847	0.201850	0.201847
x_3	0.081425	0.081425	0.081409	0.081421	0.081409
x_4	0.676977	0.676982	0.676981	0.676970	0.676981
y_1	0.078272	0.078273	0.078296	0.078277	0.078296
y_2	0.069894	0.069887	0.069882	0.069884	0.069882
y_3	0.441308	0.441311	0.441305	0.441312	0.441305
y_4	0.410526	0.410535	0.410521	0.410534	0.410521
Gibbs	-90.770474	-90.779501	-90.77952	-90.779491	-90.779499
OFE	-	**23950**	**17600**	**18100**	**10750**

In table 2 it is possible to observe a reduction of 24.43% in the number of objective function evaluations (*OFE*) when comparing the algorithms DE and GA without considering the population size updating. On the other hand, when the strategy for the population size updating is used, this reduction is 26.52% for the AGA in comparison with the GA, and 40.61 % for the ADE in comparison with the DE.

Figure 1 shows the objective function, convergence rate and population size as obtained by the ADE approach.



(a) Objective function (b) Dynamic updating of the population size

Figure 1: Results obtained using ADE approach.

In this figure it is important to observe that when the convergence rate (r) is close to one, the population size is at its minimum value, i.e., 10 individuals are found in the

population (normalized values). On the other hand, if the convergence rate is close to zero, the population size is at its maximum value, i.e., 50 individuals are found in the population. This aspect is very important, since in this case ADE or AGA preserve the characteristics of DE and GA algorithms, respectively.

6. Conclusions

In this paper a new strategy for the dynamic updating of the population size was presented for dealing with optimization problems. This methodology permits the reduction of the population size during the evolutionary process and, as a consequence, a reduction of the number of objective function evaluations. It was shown that the proposed methodology can be easily incorporated in classical evolutionary strategies, such as the DE and GA algorithms. The methodology was tested in a typical example of complex chemical equilibrium calculations involving two phases, namely the esterification reaction in an equimolar mixture of ethanol and acetic acid reacting reversibly to form ethyl acetate and water. The results showed that the proposed algorithm represents an interesting alternative for the treatment of optimization problems, once the same solution quality achieved by other techniques can be obtained by using a smaller number of generations. Further research work will be focused on the solution of phase stability problems combined with the methodology proposed in this paper.

Acknowledgments

Dr. Lobato greatly acknowledges FAPEMIG for the financial support to the present research work. Ms. Figueira greatly acknowledges CAPES for the financial support to the present research work. Dr. Steffen is thankful to CNPq for his research grant.

References

J. Castillo, I. E. Grossman (1981). Computation of phase and chemical equilibria. *Computers & Chemical Engineering*, 5, 99–108.

G., Lantagne, B. Marcos and B. Cayrol (1988), Computation of complex equilibria by nonlinear optimization. *Computers & Chemical Engineering*. 12(6), 589-597.

B. Lin, and D. C. Miller, (2004), Tabu search algorithm for chemical process optimization. *Computers and Chemical Engineering*, 28, 2287–2306.

C. M. McDonald, C. A. Floudas (1995). Global optimization for the phase and chemical equilibrium problem: Application to the NRTL equation. *Computers & Chemical Engineering*, 19, 1111–1139.

M. O. Ohanomah and D. W. Thompson (1984). Computation of multicomponent phase equilibria - Part I: Vapour-liquid equilibria. *Computers & Chemical Engineering*, 8 (3/4), 147-158.

Y. Peng Lee, G. P. Rangaiah and R. Luus (1999). Phase and chemical equilibrium calculations by direct search optimization. *Computers & Chemical Engineering*, 23, 1183–1191.

M. Srinivas and G. P. Rangaiah (2007), A study of differential evolution and tabu search for benchmark, phase equilibrium and phase stability problems, *Computers and Chemical Engineering*, 31, 760–772.

R. Storn and K. Price (1995). Differential evolution - A simple and efficient heuristic strategy for global optimization over continuous spaces, *Journal of Global Optimization*, vol. 11, pp. 341–359.

S. Y. Sun, Yc. Quiang, Y. Liang, Y., Liu and Q. Pan, Q. (2007), Dynamic population size based particle swarm optimization, ISICA 2007, Spring-Verlag Berlin, pp. 382-392, 2007.

S. Vellev (2008), An adaptive Genetic Algorithm with dynamic population size for optimizing join queries, In International Conference: Intelligent Information and Engineering Systems, INFOS 2008, Varna, Bulgaria, June-July.

W. Xiao, K. Zhu, W. Yuan and H. H. Chien (1989). An algorithm for simultaneous chemical and phase equilibrium calculation. *American Institute of Chemical Engineers Journal*, 35, 1813–1820.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Group Contribution Method for Mineral Flotation Circuit Design

Edelmira D. Gálvez[a], Luis A. Cisternas[b], Gonzalo Herrera[b], and Rafiqul Gani[c]

[a]*CICITEM, Department of Metallurgical Engineering, Universidad Católica del Norte, Antofagasta, Chile*
[b]*CICITEM, Department of Chemical Engineering, Universidad de Antofagasta, Antofagasta, Chile*
[c]*CAPEC, Department of Chemical Engineering, Technical University of Denmark, DK-2800 Lyngby, Denmark*

Abstract

Modelling and simulation of a flotation circuit usually involve identifying the structure of the flowsheet, deriving model equations to represent each operation, and solving the resulting total model equations according to one of various available simulation strategies. The flotation circuit synthesis problem determines the type of operations and their sequence needed to achieve the extraction of valuable metals from the ore minerals. The flotation circuit design problem determines the optimal values for the conditions of operation and for other operation/equipment related variables for the synthesized circuit. It can be noted that the circuit modelling, synthesis and design problems are related since for generation and screening of circuit alternatives (synthesis/design), some form of flotation circuit models are needed. Also, circuit models are needed for verification of the synthesis/design problem solution. In flotation circuit synthesis, as in other process synthesis problems, three types of approaches exist: a) the methods that employ heuristics or are knowledge based; b) the methods that employ mathematical or optimization techniques, and c) the methods that employ physical insights. d'Anterroches and Gani (2005) have introduced the concept of process-group contributions for process flowsheet property estimation and process flowsheet synthesis and design. In this work a group contribution model have been development for flowsheet properties for flotation circuit.

Keywords: Flotation circuit, process synthesis, group contribution.

1. Introduction

Flotation is a practical method to separate valuable minerals based on differences in surface properties of the particles from milled mineral mixtures. Also new applications as the separation of plastics of different density in waste streams containing a mixture of plastics have appeared. In froth flotation, various types of equipment exist which promote particles-bubbles encounter which contribute to controlling the balance between high recovery of the desired metal, and a high grade value of the metal in the product outflow. Past experience has shown single step separation is inefficient, and the inclusion of several complementary and supportive steps are required. Taking into account the large volume of material to be treated and its associated costs, choices related to the configuration of the separation system are critical.
Modelling and simulation of a mineral process flowsheet usually involve identifying the structure of the flowsheet, deriving model equations to represent each operation, and

solving the resulting total model equations according to one of various available simulation strategies. The flowsheet synthesis problem determines the type of operations and their sequence needed to achieve the extraction of valuable components from the raw materials. The flowsheet design problem determines the optimal values for the conditions of operation and related variables for the synthesized flowsheet. It can be noted that the flowsheet modelling, synthesis and design problems are related since for generation and screening of flowsheet alternatives (synthesis/design), some form of flowsheet models are needed. Also, flowsheet models are needed for verification of the synthesis/design problem solution. In mineral process synthesis three types of approaches exist: a) the methods that employ heuristics or are knowledge based (Connolly and Prince, 200); b) the methods that employ mathematical or optimization techniques (Goria et al., 2005; Cisternas et al., 2006), and c) the methods that employ physical insights (Gálvez, 1998). A review of methods for conceptual flotation circuit design has been recently published (Mendez et al., 2009). d'Anterroches and Gani (2005) have introduced the concept of process-group contributions for process flowsheet property estimation and process flowsheet synthesis and design, and they presented a property model for distillation operations. The objective of this work is to develop a group contribution model for flowsheet properties applied to flotation circuit.

2. 2. Group Contribution Model Development

The generation of the group contribution method was carried out in two steps: 1) data generation of flotation circuit properties and 2) model development and model adjustment.

2.2. Data generation

The supposition that the flotation process corresponds to a first order reaction is broadly used. If a group of solid particles transported in a pulp collide with bubbles within certain defined volume, the valuable (hydrophobic) mineral will adhere to upward bubbles, becoming separated from the gangue. This phenomenon can be considered as a simple mechanism where flotation is a pseudoreaction between the solid particles (*A*) and the bubbles (*B*) where $A+B \rightarrow AB$. If the concentration of bubbles is constant, then the flotation kinetic can be represented as a first order pseudoreaction. Then, the mineral flotation can be modeled considering that a mineral is formed of different classes or pseudospecies that have the same floatability. It means floatability represents the propensity of the particles to be floated according to its composition and size given certain flotation environment conditions (pH, Eh, reagent type, concentration), in such a way the most valuable particle, i.e., the most completely liberated valuable particle, has the greater floatability value in an intermediate particle size. Then, in figure 1, the following equations can be written:

$$C_i = T_i \, F_i \tag{1}$$

$$W_i = (1 - T_i) \, F_i \tag{2}$$

Where T_i is the mass ratio of concentrate and feed of class specie *i*. The ratio T_i may be obtained from plant data, values from pilot plants, or theoretical or empirical models. For example, for a bank of *N* cells,

$$T_i = 1 - \frac{1}{(1 + k_i \tau)^N} \tag{3}$$

Where k_i is the first order kinetic constant of class specie i, and τ the retention time in one cell.

In order to adjust a group contribution model, recovery values for different flotation circuits were generated by the simulation of several flotation circuits. Figure 2 shows a superstructure that represents a total of twenty four flotation circuits: two circuits with two flotation stages, four circuits with three flotation stages and eighteen circuits with four flotation stages. Each flotation stage was simulated using the equation 2, for multiple hypothetical classes of minerals. Each hypothetical mineral was generated with random values for the kinetic constant, cell retention time, and *N*=1, 3, 5 and 7 in equation 3. Figure 3 shows T_i values for different first order kinetic constant, retention time, and *N*=1, 3, 5 and 7.

2.2. Model adjustment

To generate the group contribution model two types of contribution were considered. First, it was considered each flotation stage can be characterized by the ratio of flow of concentrate and feed, T_i. Then the interconnections among flotation stages were considered, represented by the product of the ratios of concentrate/feed or tail/feed and concentrate/feed or tail/feed as it is the case: $T_i\ T_i, T_i\ (1-T_i)$, or $(1-T_i)\ (1-T_i)$. Then the model for recuperation of a pseudospecie has the following form:

$$R = \alpha + \sum_i \beta_i\ N_i\ T_i + \sum_i \sum_j \gamma_{i,j}\ N_{i,j}\ \Lambda_{i,j} \tag{3}$$

Where α, β_i and $\gamma_{i,j}$ are adjusted parameters, and T_i and $\Lambda_{i,j}$ are process-group contributions. The adjusted parameters were fitted to simulated recovery values for mineral class with low recovery (0 to 10%), medium recovery (10 to 60%) and high recovery (60 to 100%). The α values were -3.5542, -5.9192, and -2.5507 for low, medium and high recoveries. The values for β_i are given in table 1 and some $\gamma_{i,j}$ values in table 2. Good results have been obtained in the prediction of recovery (see figure 4).

3. Example

To explain the use of the equation 3, let us consider the flotation circuit shown in figure 5, where a mineral conformed by two pseudospecies, one with high recovery and another with low recovery, is fed. The equation 3 for this case corresponds to:

$$\begin{aligned} R = {} & \alpha + \beta_{T_R} T_R + \beta_{T_C} T_C + \gamma_{T_R T_C} T_R T_C + 2\gamma_{T_R (1-T_C)} T_R (1-T_C) + \\ & \gamma_{(1-T_R)(1-T_C)} (1-T_R)(1-T_C) \end{aligned} \tag{4}$$

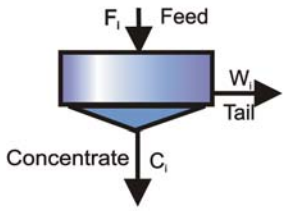


Figure 1. Flotation Stage

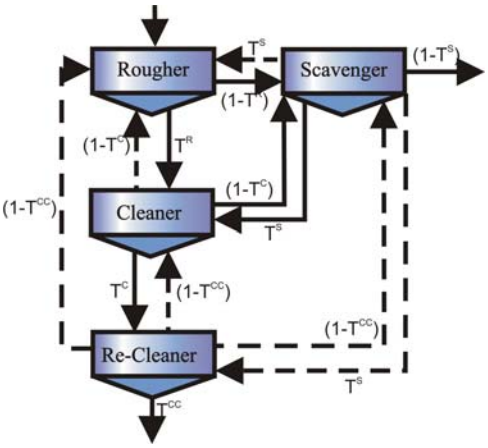
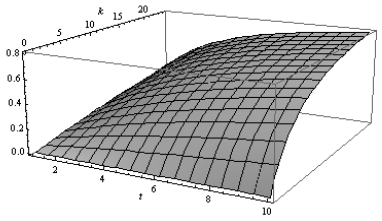
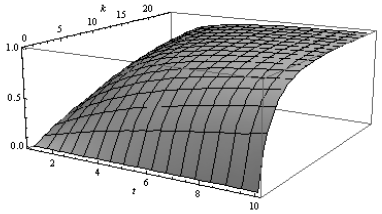


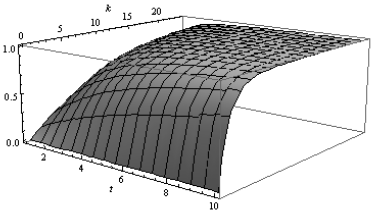
Figure 2. Superstructure of flotation circuit



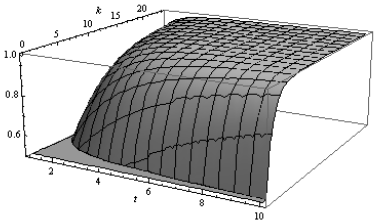
(N = 1)



(N = 3)



(N = 5)



(N = 7)

Figure 3. T_i values from equation 3 for different values of k_i, τ and *N*.

Table 1. Values for β_i groups.

Recovery	Flotation stage groups			
	T_R	T_C	T_{CC}	T_S
Low	1.8913	1.9870	-3.2947	0.0348
Medium	2.0104	2.3000	-5.1205	-1.0106
High	1.3011	1.2367	-1.0042	-1.3571

Table 2. Some values of $\gamma_{i,j}$

Flotation stages groups, $\Lambda_{i,j}$		$\gamma_{i,j}$ values		
I	j	low	medium	High
T_R	T_C	0.4251	2.6465	1.2645
$1-T_R$	T_S	0.7060	1.3288	0.6314
T_C	T_S	-0.3160	0.8209	2.6079
T_{CC}	T_S	-0.2917	1.1177	3.4985
T_R	$1-T_C$	1.8040	3.1885	1.1269
$1-T_R$	$1-T_C$	3.0154	6.0501	2.5505

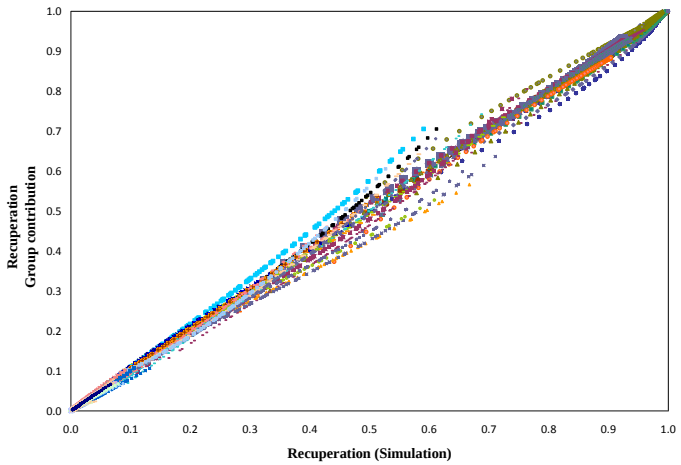


Figure 4. Recuperation values: Group contribution versus simulated values

For values of $T_R = 0.906$ and $T_C = 0.716$ for the high recovery speudospecie, the calculated valor of *R*, in equation 4, corresponds at 0.867, while the mass balance value of *R* corresponds at 0.873. On the other hand, for the low recovery speudospecie, with

values of $T_R = 0.07$ and $T_C = 0.05$, R correspond at 0.009 and 0.007 for equation 4 and mass balance respectively.

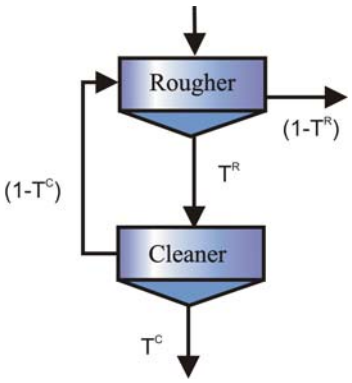


Figure 4. Flotation circuit for the example.

4. Conclusion and future work

A group contribution model was presented for the recovery estimates in flotation circuits. For the analyzed circuits the method gives acceptable results for process synthesis purpose. Works to include circuits with more flotation stages, development of an approach for the selection of T values, and an approach for the determination of the maximum number of flotation stages are under way.

5. Acknowledgements

We thank the CONICYT for financing and support of the work reported in this manuscript (FONDECYT 1060342).

References

A.F. Connolly, R.G.H. Prince, 2000, Performance improvement in minerals beneficiation circuits by retrofitting ,Separation and Purification Technology 19, 77–83

L.A. Cisternas, D.A. Mendez, E.D. Galvez, R. Jorquera, 2006, A MILP model for design of flotation circuits with bank/column and regrind/no regrind selection, International Journal of Mineral Processing, 79 (4), 253-263.

L. d'Anterroches, R. Gani, 2005, Group contribution based process flowsheet synthesis, design and modelling, Fluid Phase Equilibria 228–229 , 141–146

E.D. Gálvez, 1998, A shortcut procedure for the design of mineral separation circuits, Minerals Engineering, 11 (2), 113-123.

C. Guria, Verma M, Gupta SK, Mehrotra, 2005, Simultaneous Optimization of The Performance of Flotation Circuits And Their Simplification Using The Jumping Gene Adaptations of Genetic Algorithm, International Journal of Mineral Processing, 77 (3): 165-185.

D.A. Mendez, E.D. Gálvez, L.A. Cisternas, 2009, State of the art in the conceptual design of flotation circuits, Int. J. Miner. Process. 90(1-4): 1-15.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Fluid Dynamics Simulation for Design of a Biomass Gasifier

Cleiton B. da Porciúncula,[a] Nilson R. Marcilio,[b] Marcelo Godinho,[b] Argimiro R. Secchi[c]

[a,b]*GIMSCOP/LPR, Department of Chemical Engineering - Federal University of Rio Grande do Sul , CEP 90040-040, Porto Alegre, RS, Brazil.*
Rua. Eng. Luis Englert, s/n. E-MAIL address: {cleiton,nilson}@enq.ufrgs.br
[c] *COPPE/PEQ/LMSCP – Federal University of Rio de Janeiro*
Av. Horácio Macedo, 2030 - Centro de Tecnologia. CEP 21941-914, Rio de Janeiro - RJ, Brazil. E-MAIL address:arge@peq.coppe.ufrj.br

Abstract

The obtaining of alternative sources to generate energy has been a worldwide focus of research in the last years, mainly due to scarcity of fossil fuels in a close future. A very promising source to obtain combustion gases (especially methane and hydrogen) consists of gasification of biomass wastes generated by tanneries and footwear-leather industries. The Brazilian state of Rio Grande do Sul has a very privileged position concerning this industrial area, because it owns a large number of companies of these branches. The objective of this work is the design and simulation of a biomass gasification reactor, by using the technique of computational fluid dynamics. A commercial computational package of simulation, CFX 11.0, was the software employed for the development of this work. The gasification process may be carried out in two types of reactor configurations: fluidized and fixed bed. The present work has been developed upon fixed bed reactors in downdraft operation, considering four steps of gasification zones develop along the gasifier: drying, devolatilization, reduction, and combustion. The single and two-phase reactions were simulated coupled with the transport equations (momentum, mass, and energy) and the tracking of biomass particles. Several design configurations were tested in order to verify how they could influence the fluid flow and the temperature distributions inside the reacting chamber. Finally, one has been chosen a configuration that counterbalances project features, easiness, and operation costs. The remaining simulations have been run from this last configuration. The preliminary results show a good consistency with experimental data obtained previously from a pilot plant. Also, based on the results obtained with different scenarios, a new and modern pilot plant will be built. Forthcoming simulations will have the objective of optimizing and improving the operating conditions of the equipment.

Keywords: gasification, computational fluid dynamics, leather, simulation.

1. Introduction

The industrial activities that involve physical or chemical transformations of leather materials show a large environmental impact. Foot-wear industries are the most consuming raw materials based on leather, around 60% as observed by Bahillo *et al.* (2004). Solid wastes derived from tanneries have high levels of chromium, which is a

toxic and hazardous metal for health. Moreover, liquid and gaseous effluents are the consequence of several reaction steps that occur along the process of leather treatment. According to Vieira (2004), the state of Rio Grande do Sul generated over 189000 ton/yr of dangerous solid wastes from companies that handle with leather products, which demonstrates the importance to treat this issue. According to the same author, only 3% of these wastes are recycled and over 85% have been sent to deposits or urban embankments. A study carried out by Font *et al.* (1999) evaluated several gases as products of volatiles emitted from leather, as light hydrocarbons, some alcohols, and aromatic compounds.

2. Gasification: theoretical aspects

The definition of gasification considered by Belgiorno *et al.* (2003) is the thermal conversion of a liquid or solid material with large content of carbon, into a gaseous fuel with the use of a gasification agent (that is, another gas). The fuel generated in this process may be burnt to obtain energy or be purified to a further separation step, in order to manufacture several other substances. Gasification may be classified as direct (gasification agent partially oxidizes biomass), indirect (occurs in the absence of oxygen) and pyrolysis (indirect gasification that occurs in inert atmosphere), according to Godinho (2006). The reactors also may be classified as fluidized and fixed bed, as described by Belgiorno *et al.* (2003). Fixed-bed gasifiers have lower velocities of the gasification agent and may operate in counter-current (updraft) or co-current (downdraft) mode. In updraft configuration, biomass is inserted in opposite direction to the gasification agent. The major drawback is the dragging of tar together with fuel gas formed in the reactor. The downdraft operation minimizes this problem because the gasification agent is inserted perpendicularly to biomass feeding section. Fixed bed reactor presents four theoretical zones with different temperatures and chemical processes, as classified by Filho (1988):

- Drying zone: Placed on the top of the gasifier, it is where the coal or biomass depletes water and is pre-heated by hot gases that emerge from the bottom of the reactor.
- Devolatilization zone: The volatiles are removed in this zone. The most common gases obtained are CO_2, CO, light hydrocarbons, nitrogen, and sulphur-based compounds.
- Reduction or gasification zone: It is where the gasification reactions occur. Temperatures vary between 800ºC and 960ºC, approximately. The main reactions considered are:

$$C + H_2O \rightleftharpoons CO + H_2 \quad \text{(I)}$$
$$C + 2H_2 \rightleftharpoons CH_4 \quad \text{(II)}$$
$$C + CO_2 \rightleftharpoons 2CO \quad \text{(III)}$$

By inverting reaction (III) and adding it to equation (I), one obtains the water gas-shift reaction:

$$CO + H_2O \rightleftharpoons CO_2 + H_2 \quad \text{(IV)}$$

- Combustion zone: Occurs the complete oxidation of char. The heat of this reaction is transferred to the previous zones described. Temperatures may reach up to 1400ºC. The reaction that occurs is:

$$C + O_2 \rightleftharpoons CO_2 \quad \text{(V)}$$

3. Mathematical modeling and operating conditions

The governing equations applied to the problem are the following:

Moment Balance (Navier-Stokes equations)

$$\frac{\partial(\rho u_i)}{\partial t}+\frac{\partial(\rho u_j u_i)}{\partial x_j}=-\frac{\partial P}{\partial x_i}+\frac{\partial}{\partial x_j}\left(\mu\frac{\partial u_i}{\partial x_j}\right)+S^{u_i} \quad (1)$$

Total mass balance (Continuity equation)

$$\frac{\partial \rho}{\partial t}+\frac{\partial}{\partial x_j}(\rho u_j)=0 \quad (2)$$

Energy balance

$$\frac{\partial(\rho T)}{\partial t}+\frac{\partial}{\partial x_j}(\rho u_j T)=\frac{\partial}{\partial x_j}\left(\frac{k}{c_P}\frac{\partial T}{\partial x_j}\right)+S^T \quad (3)$$

Component mass balance

$$\frac{\partial C_i}{\partial t}+\frac{\partial(u_j C_i)}{\partial x_j}=\frac{\partial}{\partial x_j}\left(D\frac{\partial C_i}{\partial x_j}\right)+S^{C_i} \quad (4)$$

Turbulence model (k-ω) - Turbulent kinetic energy

$$\frac{\partial(\rho k)}{\partial t}+\nabla\cdot(\rho U k)=\nabla\cdot\left[\left(\mu+\frac{\mu_T}{\sigma_k}\right)\nabla k\right] \quad (5)$$

Turbulence model (k-ω) - Turbulent frequency

$$\frac{\partial(\rho\omega)}{\partial t}+\nabla\cdot(\rho U\omega)=\nabla\cdot\left[\left(\mu+\frac{\mu_T}{\sigma_\omega}\right)\nabla\omega\right]+\alpha\frac{\omega}{k}[(\mu_T\nabla U\cdot(\nabla U+\nabla U^T)-\beta\rho\omega^2] \quad (6)$$

In conservation equations above, u is the velocity vector in $m{\cdot}s^{-1}$, P is the pressure in Pa, x represents the dimensional variable in m, μ is viscosity in $Pa{\cdot}s$, T is temperature in K, ρ is the gas density in $kg{\cdot}m^{-3}$, C represents the concentration of each gaseous species in $kg{\cdot}m^{-3}$, and S accounts for the source terms in momentum (buoyancy force), energy (heat of reaction) and component (source of each gaseous species) balances. The system under consideration is three-dimensional and operates at steady-state conditions, so the transient terms (where t is time, in s) of the equations (1-6) are dropped out. About the turbulence model, k is the turbulent kinetic energy, in $m^2{\cdot}s^{-2}$, ω is the turbulent frequency, in s^{-1}, U is the velocity vector in $m{\cdot}s^{-1}$, μ_T is the turbulent viscosity in $Pa{\cdot}s$, and the parameters σ_ω, σ_k, α, and β are model parameters. The homogeneous gas phase reactions are:

$$H_2+1/2\,O_2\rightarrow H_2O \quad (VI)$$
$$CH_4+3/2\,O_2\rightarrow CO+2H_2O \quad (VII)$$
$$CH_4+2O_2\rightarrow CO_2+2H_2O \quad (VIII)$$
$$CO+1/2\,O_2\rightarrow CO_2 \quad (IX)$$
$$CO+H_2O\rightleftharpoons CO_2+H_2 \quad (X)$$

The rate expressions for reactions (II, and V-X) are the following (Filho (1988), ANSYS® CFX 11.0 User Manual):

Methane Formation

$$r_{II}=-k_{II}\left(p_{H2}{}^2-\frac{p_{CH4}}{K_{eq}}\right) \quad \text{Reaction (II)} \quad (7)$$

Water-gas shift reaction

$$r_{IV}=-k_{IV}(p_{CO}p_{H2O}-p_{CO2}p_{H2}) \quad \text{Reaction (X)} \quad (8)$$

Char Oxidation

$r_V=-\left(k_c{}^{-1}+k_d{}^{-1}\right)^{-1}p_{O2}A_P$, where: Reaction (V) (9)

$$k_C = A_C e^{-T_C/T_P} \quad - \quad (10)$$

$$k_d = \frac{D_{ref}}{R_P}\left(\frac{T_P + T_g}{2T_{ref}}\right)^{\alpha}\frac{P_A}{P} \quad - \quad (11)$$

Gas Phase Oxidation

$$r_{comb,n} = -A\frac{\epsilon}{k}\min\left(\frac{[I]}{\alpha_{I,n}}\right) \quad \text{Reactions(VI - X)} \quad (12)$$

The reaction rate r is expressed as $mol \cdot m^{-3} \cdot s^{-1}$ in equations (7) and (8), and as $kg \cdot m^{-3} \cdot s^{-1}$ in equation (9). The terms p_i represent the partial pressure of each gaseous compound i, expressed as Pa, K_{eq} is the equilibrium constant in Pa^{-1}, k_j is the specific reaction velocity of the reaction j in units according to each reaction order. The terms k_c and k_d are the Arrhenius reaction rate of char oxidation and resistance of oxygen diffusion into the particle, expressed in $kg \cdot m^{-2} \cdot s^{-1}$ and $m \cdot s^{-1}$, respectively. The parameters T_g, T_{ref}, T_P, and T_c represent the temperatures of gas, reference, particle, and activation (activation energy divided by perfect gas constant R), respectively, while R_P is the particle radius in m, A_p is the particle surface area in m^2 and α is an exponential parameter set equal to 0.75 in CFX 11. Other default parameters considered by the solver are T_{ref} (293K), the oxygen diffusivity of reference, D_{ref}, $1.8 \cdot 10^{-5}$ $kg \cdot m^{-1} \cdot s^{-1}$ and the parameter A_c in equation (10), set as 4. The terms $[I]$ and $\alpha_{I,n}$ are the molar concentration of the gas I and the stoichiometric coefficient of the respective gas in each combustion reaction n. The rate of gas phase combustion $r_{comb,n}$ in $mol \cdot m^{-3} s^{-1}$ is taken into account as an instantaneous process, given by the ratio of the local turbulent dissipation energy and turbulent kinetic energy. Transport of particles is considered by applying the second law of Newton:

$$m_P \frac{dU_P}{dt} = F_D + F_S \quad (13)$$

where:

$$F_D = \frac{C_D \rho_f A_P |U_F - U_P|(U_F - U_P)}{2} \quad (14)$$

$$F_B = \frac{\pi d_P{}^3 (\rho_P - \rho_F) g}{6} \quad (15)$$

In equations (13-15), m_p is the particle mass in kg, U_P is the particle velocity vector expressed as $m \cdot s^{-1}$, F_D and F_B are the drag and buoyancy forces in N, respectively. The drag force is considered to admit a constant non-dimensional drag coefficient $C_D = 0.44$, ρ_f is the fluid density in $kg \cdot m^{-3}$, d_P is the particle diameter in m and g is the acceleration of gravity field, 9.8 $m \cdot s^{-2}$. The equations (13-15) are solved numerical and simultaneously with the balance equations of moment, mass, and energy. The operating conditions are given in Table 1 and the corresponding scheme is depicted in Figure 1.

Table 1 - Temperature and flow in the gasifier

Location (see Fig. 1)	Flow ($kg \cdot h^{-1}$)	Temperature (K)
Top	200 (biomass)	673
Burner (2)	90	673
Air box (1) and air tubes (3)	270	303
Ashtray Air (6)	200	303

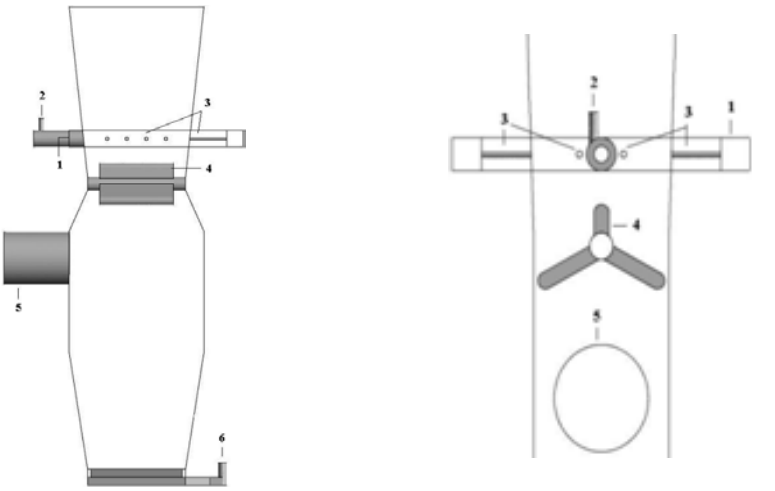


Figure 1 - Scheme of the gasifier.

The parts numbered as 4 and 5 are the metallic grate and the outlet of the reactor, respectively. The grate retains thick particulates (ash and tar), separating the sections of reduction (above) and starting of combustion (below) of the gases formed. The outlet is to be connected to a combustion chamber that is also being simulated apart. Wastes of leather are fed on the top of the gasifier, where devolatilizations occur (still not considered in this step). The remaining particles consist of char and ash, so that the fixed carbon is oxidized to generate heat that maintains the temperature that sustains the gasification reactions. Air is fed through the burner and tubes that redistribute it along the reactor. The ashtray air begins the combustion below the grate and oxidizes the remaining char particles.

4. Results and Conclusions

Estimated molar fractions of the main gases are presented in Table 2, compared with some previous results from the former reactor configuration. In Figure 2 are shown temperature contours and a sample of the three-dimensional vector field of biomass particles with varying char mass fractions.

Table 2 - Mean molar fractions at gasifier outlet

Component	Molar Fraction	Molar Fraction (Godinho, 2006)
CH_4	0,08	0,118
H_2	$2{,}09 \cdot 10^{-6}$	-
CO	0,0648	0,061
CO_2	0,081	0,132
H_2O	0,029	-

The obtained fuel gas has a very low concentration in hydrogen. Larger values may be reached by injecting steam into the gasifier to displace the water-gas shift equilibrium to the right. This approach is intended to further be applied in order to increase the calorific power of the gas.

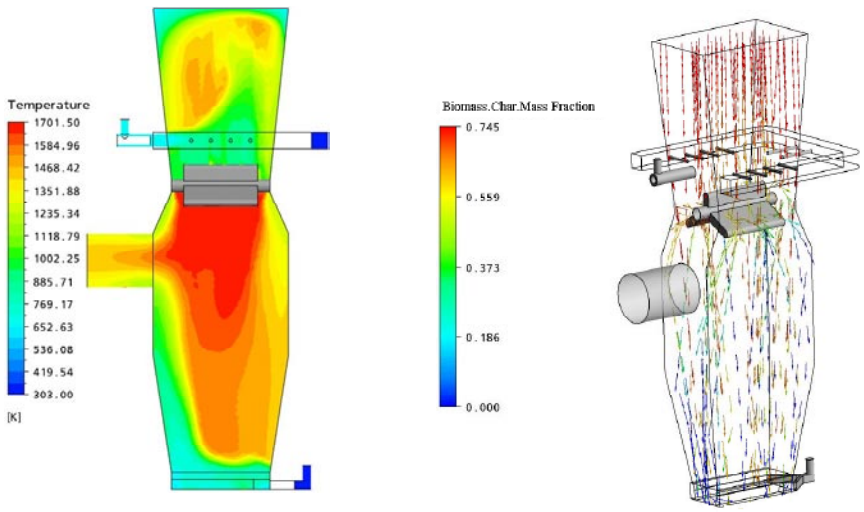


Figure 2 - Temperature profiles and vector field.

According to Figure 2, the largest temperatures occur immediately below the grate, where the homogeneous combustion reactions begin to occur. About the char particles, one verifies the bulk combustion occurs in the upper-half of the reactor. Although some non-burnt particles reach the ashtray, most of them are oxidized with the oxygen fed in that section.

These preliminary results show that gasification of leather wastes seems a very promising energy source due to capability of formation of useful gases for combustion and heat generation (mainly methane and likely hydrogen). The expermental results obtained with the plant operation shortly will be compared with the simulations carried out in this work to design the reactor.

Acknowledgments

The authors are very thankful to profs. Celso Martins, Leonardo Masotti and Marcelo Godinho for their valuable experience and help in the development of this work.

References

Ansys® CFX 11.0 User Manual, http://www.ansys.com/cfx, 2007.

A. Bahillo, L. Ernesto, A. Cabanillas, J. Otero, 2004, Thermal valorization of footwear leather wastes in bubblingfluidized bed combustion, Waste Management, 24, 935-944

V.Belgiorno, G. De Feo, C. Dellaa Rocca, R.M.A. Napoli, Energy from gasification of solid wastes, 2003, Waste Management, 23,1-15.

P.Filho, Gaseificadores de Leito Fixo - Análise e modelo para otimização de projetos e pesquisas, 1988.

R. Font, J.A. Caballero, M.M. Esperanza, A. Fullana,1999, Pyrolitic products from tannery wastes, Journal of Analytical and Applied Pyrolysis, 49, 243-256.

M.Godinho,2006, Gaseificação e combustão de resíduos sólidos da indústria calçadista, Post-Graduating Program of Mining, Metallurgic and Material Engineering, Porto Alegre, RS, Brazil.

M. Vieira, 2004, Recuperação do cromo contido nas cinzas provenientes da incineração de resíduos sólidos das indústrias coureira e calçadista visando a obtenção decromato de sódio(VI), Post-Graduating Program of Chemical Engineering, Porto Alegre, RS, Brazil.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Polystyrene produced by a Multifunctional Initiator

Paula F. de M. P. B. Machado[a], Liliane M. F. Lona[b]

[a]*LASSPQ/DPQ- School of Chemical Engineering - University of Campinas*
Cidade Universitária "Zeferino Vaz", s/n, CEP 13083-970, Campinas - SP, Brazil.
E-mail address: pmachado@feq.unicamp.br
[b]*LASSPQ/DPQ- School of Chemical Engineering - University of Campinas*
Cidade Universitária "Zeferino Vaz", s/n, CEP 13083-970, Campinas - SP, Brazil.
E-mail address: liliane@feq.unicamp.br

Abstract

Free-radical polymerization of styrene in the presence of trifunctional initiator was investigated to understand the behavior of initiators with functionality greater than two. The trifunctional initiator selected was the TRIGONOX 301 from Akzo Nobel. This study proposes a kinetic mechanism and a mathematical model to free radical polymerization of styrene using a trifunctional initiator. The complexity of the kinetic mechanism increases as the functionality of the initiator increases. The mathematical model was built and predicts results as conversion, molecular weight, radical concentration and polymer concentration. An experimental investigation was also explored to validate the simulations results. The experimental part followed the method of polymerization in ampoules. Experiments were also carried out to verify the effect of temperature and initiator concentration on the polymerization kinetics and molecular weights.

Keywords: multifunctional initiator, polystyrene, polymerization.

1. Introduction

The study of chemical initiators that have functionality superior than two has been explored in the scientific and industrial field. Free-radical polymerization in the presence of multifunctional initiator has been also investigated to understand the behavior of these initiators. The multifunctional initiator is able to increase the reaction rate in a free radical polymerization without decrease the molecular weight of the formed polymer and it can also changes the polymer structure.

Literature shows experimental and simulation studies related to free radical polymerization using bifunctional initiators (Villalobos et al., 1991; Dhib et al., 2000; Machado and Lona 2005 and others). Cerna et al. (2002) and Sheng et al. (2004) present detailed and important experimental studies with trifunctional initiators.

There are very few studies presenting models of free radical polymerization using multifunctional initiators (Scorah, 2005 and Scorah et al., 2007).

To understand the behavior of initiators with functionality greater than two, trifunctional initiators were chosen. The trifunctional initiator selected was the TRIGONOX 301 from Akzo Nobel. It is a cyclic triperoxide and very unstable at temperatures greater than 40ºC. It must be kept at room temperature. This study proposes a kinetic mechanism and a mathematical model to free radical polymerization of styrene using

this trifunctional initiator since there are very few studies presenting models of free radical polymerization using multifunctional initiators. Styrene is chosen because it is a very well known monomer and there are a lot of data about it. The mathematical model was built in Fortran 90, based on the Method of Moments Equations for all species, and predicts results as conversion, molecular weight (Number Average, Mn, and Weight Average, Mw), radical concentration and polymer concentration. An experimental investigation was also explored to validate the simulations results. The experimental part followed the method of polymerization in ampoules. Full-conversion-range experiments were carried out to validate the model and also to verify the effects of the temperature and initiator concentration on the polymerization kinetics and molecular weights. Conversion was calculated based on differences of masses (gravimetry). Gel Permeation Chromatography (GPC) was used to measure the molecular weights of polystyrene. The validation was made to conversion and molecular weights.

2. Experimental Procedure

Styrene was first washed with NaOH solution and deionized water. Then it is dried with $CaCl_2$. The washed styrene is distillate using a vertical rotative evaporator equipped with vacuum pump and a hot bath.

After that determined quantities of the solution initiator-monomer are transferred to ampoules. They are all sealed after the elimination of oxygen and put into a bath of oil at a desirable temperature.

At certain intervals of time, each ampoule is removed from the bath, cooled and weighted. Then the ampoule is broken and the solution polymer + monomer + initiator is dissolved in tetrahydrofuran (THF) and then the polymer is precipitated with ethanol. After ethanol evaporation, the sample of polymer goes to a dryer to guarantee the total evaporation of the solvent. After this, the dried samples are weighted, and conversion is calculated based on differences of masses (gravimetry). Finally, the polymer is dissolved in THF and characterized by GPC analysis, providing molecular weights data. These results bring elements to understand the behavior of trifunctional initiator in the production of polystyrene, since they provide data to validate the mathematical model and to estimate parameters.

3. Results

The simulation results and experimental data are shown in Figures 1, 2, 3 e 4. The experimental data are used to validate the model.

In Figures 1 and 2 it was considered temperature equal to 130°C and initiator concentration of 0.005M.

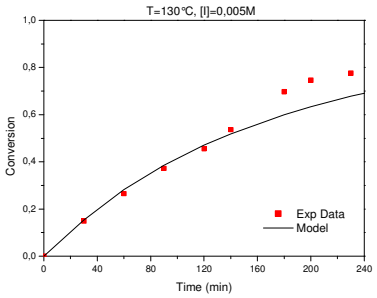


Figure 1. Conversion x Time.

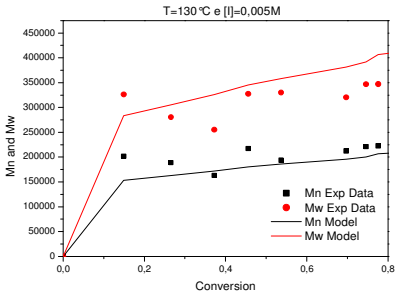


Figure 2. Weight and Number Average Molecular Weights x Conversion.

Figures 4 and 5 show results from a different operating condition; temperature of 125°C and initiator concentration of 0.0029M.

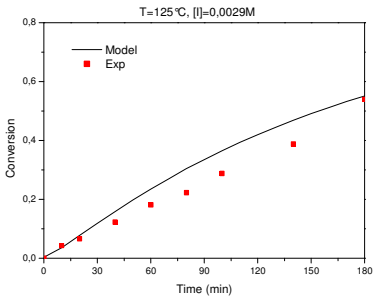


Figure 3. Conversion x Time.

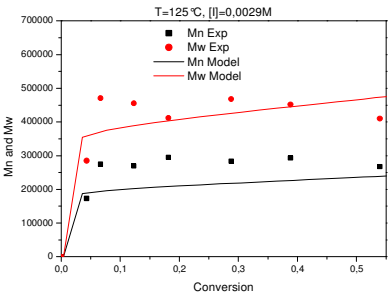


Figure 4. Weight and Number Average Molecular Weights x Conversion.

It is possible to verify from Figures 1 to 4 that the model presents agreement with experimental data. It was possible to explore the model using different operating conditions providing a study of effects of temperature and initiator concentration on the reaction rate and proprieties of the polymer. Figures 5 and 6 show the effect of temperature in conversion and number average molecular weight.

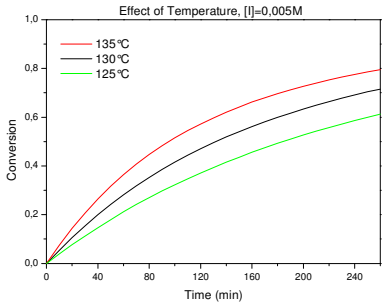


Figure 5. Conversion x Time.

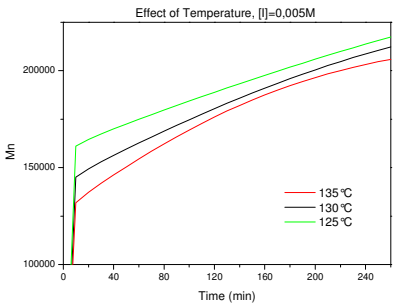


Figure 6. Number Average Molecular Weight x Time.

It is possible to verify in Figure 5 that conversion increases as the temperature increases. That is expected since an increasing in temperature increases the reaction rate generating more radicals to react with the molecules of monomer. Figure 6 shows that Mn decreases as the temperature increases, what was also expected, since a large quantity of radicals implies in a large quantity of chains, but short ones, making the molecular weight decreases.

Figures 7 and 8 show the effect of initiator concentration in conversion and number average molecular weight.

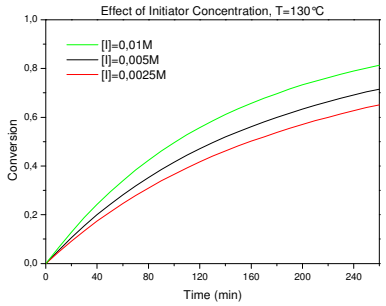


Figure 7. Conversion x Time.

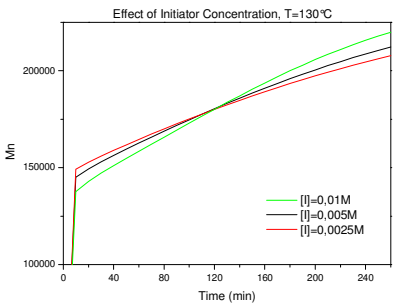


Figure 8. Number Average Molecular Weight x Time.

As expected, Figure 7 shows that an increasing in initiator concentration is also very important to achieve faster conversions, since a large quantity of initiator implies in a large quantity of consumed monomer. Figure 8 shows that, first, Mn decreases as the initiator concentration increases and then this behavior changes and Mn starts to increase with the increasing in the initiator concentration. The first part was expected, since a large quantity of radicals implies in a large quantity of chains, but short ones, making the molecular weight decrease, but the behavior in the second part is really unexpected. Maybe, this can be explained by the fact that with the increase on the initiator concentration, more radicals are formed resulting in the increasing of the number of chains and before 120 minutes, Mn (which depends on the number of chains)

decreases. However, after 120 minutes, as the initiator concentration is smaller (there is less initiator in the reaction), when all initiator is consumed, the polymeric chains that have undecomposed peroxides start to break, making Mn decrease. At this point, the inversion in the behavior begins because the highest initiator concentration has not all initiator consumed, so its chains have not yet their undecomposed peroxide broken, so molecular weights of the highest initiator concentrations keep still growing.

4. Conclusion

Results from model simulation of styrene polymerization using trifunctional initiator TRIGONOX 301 showed agreement with experimental data. These results from simulation are new, since there is no research published about this subject. The existent models are very few and they are not based in the Method of Moments. So far it was possible to verify some aspects of the behavior of a trifunctional initiator through the study of the effects of temperature and initiator concentration and it was possible to see how interest is the behavior of a multifunctional initiator in a polymerization reaction and to verify the benefits of its use.

References

J. R. Cerna, G. Morales, G. N. Eyler, A. I. Canizo, 2002, J. Applied polymer sci., 83, 1-11.

R. Dhib, J. Gao, A. Penlidis, 2000, Polymer Reaction Engineering, v.8, n.4, p. 299-464.

P.F.M.P.B. Machado, L.M.F. Lona, 2005, E. Symp. on Comp. Aided Proc. Eng.-15, 2005, 20 A, 445-450.

M. J. Scorah, 2005, PhD thesis, Dept. of Chem. Eng., University of Waterloo, Waterloo, Ontario, Canada.

M. J. Scorah, R. Dhib, A. Penlidis, 2007, Macrom. React. Eng., 1, 209-221.

W. C., Sheng, J. Y. Wu, G. R. Shan, Z. M. Huang, Z. X. Weng, 2004, J. Appl. Polymer Sci. , 94 , 1035-1042.

M. A. Villalobos, A. E. Hamielec, P. E. Wood, 1991, Journal of Applied Polymer Science, 42, 629-641.

1082

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Fluid dynamics study of the influence of the position of the feed in fabric filters

Sandra M. S. Rocha[a], Luiz G. M. Vieira[a], Mônica L. Aguiar[b], João J. R. Damasceno[a]

[a]*School of Chemical Engineering, Federal University of Uberlândia. Av. João Naves de Ávila, 2121, 38408-100, Uberlândia-MG, Brazil, E-mail: ballueng@gmail.com*
[b]*Department of Chemical Engineering, Federal University of São Carlos. Via Washington Luiz, Km 235, 13565-905, São Carlos-SP, Brazil*

Abstract

The fabric filter is a pioneering piece of equipment used to reduce the emission of small size particulate matter from industrial sources. It is relatively inexpensive and offers easy operation and high efficiency of collection for a wide range of sizes. Aiming at making the filtration operation more efficient and economical, this work investigated the best position of the feed in the filtration box and, consequently, the best way to clean the fabric filter. The cleaning system used in fabric filters is of great importance for the efficiency and durability of the equipment, and it defines the position of the dirty gas inlet in the filtration box. This work used CFD (Computational Fluid Dynamics) techniques to simulate the pressure drop and the fluid dynamics of the gas as a function of the position of the feed, which was varied from the top to bottom of the filtration box. The simulations were carried out using a new polyester filter for each feed position. According to the results, the fluid should be fed next to the bottom of the filtration box, which produces a smaller pressure drop, a better distribution of the gas and less wear of the bags adjacent to the gas feed.

Keywords: Filtration, Fabric Filter, CFD.

1. Introduction

The increasing concern with the quality of the air, the limits of atmospheric emissions of particulate materials required by the current law, as well as the need to recover raw materials or added value products demand a more effective control of the atmospheric emissions on the part of the industries. The filtration is one of the first and most common operations used to clean gases and capture particles from gaseous flows due to its high collection efficiency, low cost of operation and low sensitivity to the fluctuations of the operating conditions. The fabric filter is the equipment most often used in dry filtration for presenting collection efficiency around 99% (Hinds, 1999).

The satisfactory performance of fabric filter for a particular application requires both the right choice of the material that will be used for the filtration (adequate choice of the fiber), which must be compatible with the particulate material that will be collected, and a project that is appropriate to the geometry of the dust collector and to the cleaning conditions (Rodrigues, 2004).

According to Damian *et al.* (2003), the distribution of the air velocity in the filtration box actually does not occur in a uniform way, presenting problems of premature wear of the bags located in regions of high velocity.

In this context, the objective of this work was to evaluate which position of the gas feed could produce a more uniform distribution of the fluid through the bags, as well as a smaller pressure drop during the operation of the filtration system.

2. Method

This work simulated the fluid dynamics of a representative part of an industrial filtration system with air-cloth ratio of 0.04 $m^3/m^2.s$. Figure 1 shows the details of the filtration box with the 13 bags. Bag number one is closest to the gas feed.

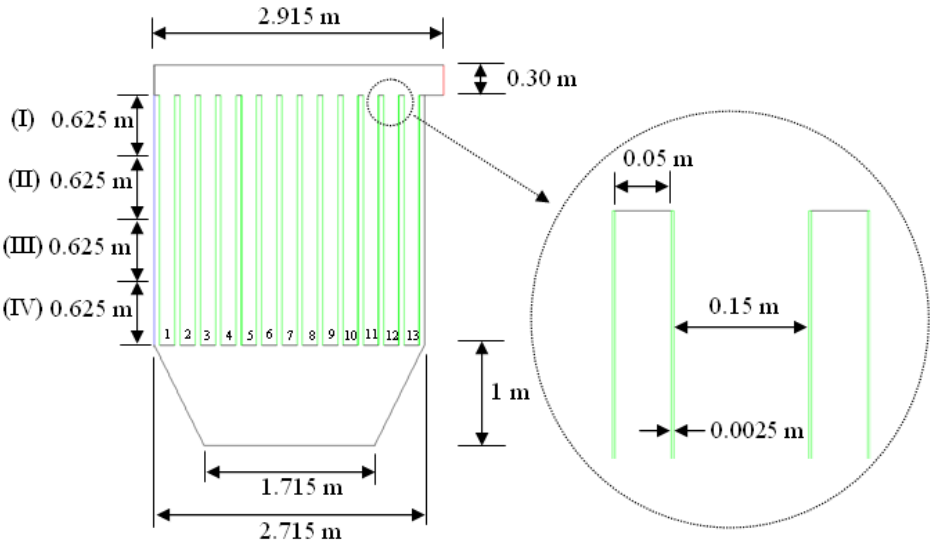


Figure 1 – Dimensions of the filtration system

The was software Fluent was used for the fluid dynamics simulations and the main boundary conditions were: velocity of the air feed 1.56 m/s, permeability $4.69x10^{-10}$ m^2, porosity 82%, thickness $2.5x10^{-3}$ m, relative of humidity of the air 20% and temperature 25ºC, all taken from Rocha *et al*. (2008).

The simulations were carried out bidimensionally with approximately 850,000 computational cells. The turbulence model used was the RSM *(Reynolds Stress Model*). The method *PRESTO*! was used for the interpolation of the pressure, and the algorithm *SIMPLE* was applied for the coupling of the pressure-velocity binomial. *UPWIND* methods of the second order were adopted for the interpolation of the remaining fluid dynamics variables (Vieira *et al*., 2007).

3. Results

The fluid dynamics simulations reveal that the pressure drop in the filtration system studied depends on the position of the gas feed, as shown in Figure 2.

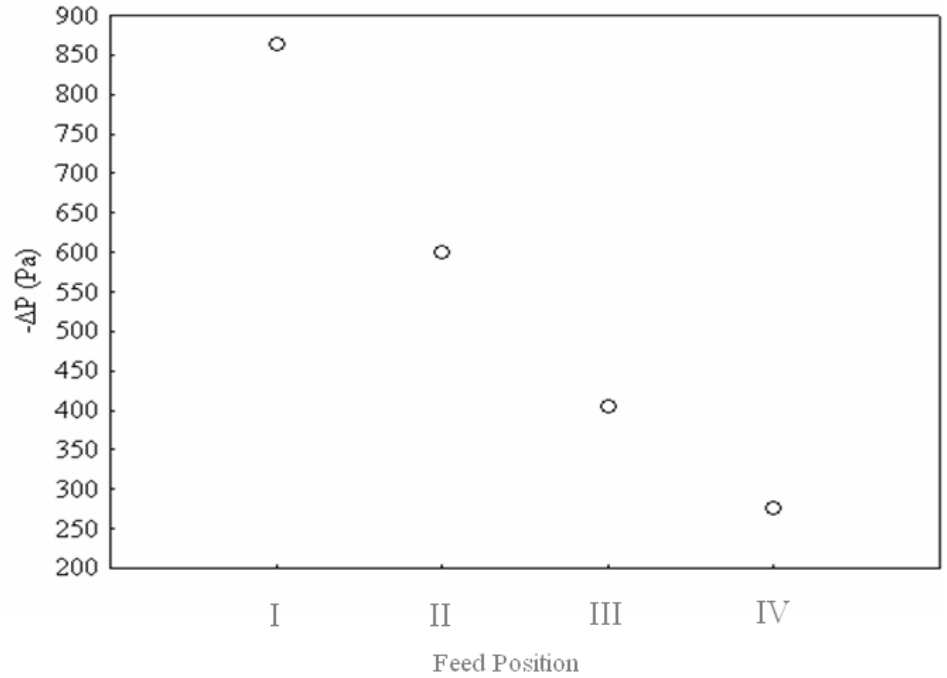


Figure 2 – Influence of the position of the gas feed on the pressure drop

As shown in Figure 2, when the fluid is fed next to the base of the filtration box, the pressure drop in the set of bags tends to diminish. Hence, the position "IV" produced the smallest pressure drop (277 Pa) among the configurations studied. In other words, for a given operational pressure drop (defined by the user), the position "IV" allows a longer filtration time, decreasing the number of cycles. For a given job, the reduction of the number of cycles allows for a smaller number of operations of cleaning of the bags (a procedure more damaging to the structure of the fabric of the filter) and, consequently, a lower wear of the bags. That, in turn, makes the operation more viable, because the bags are the most expensive component of a typical filtration unit.

Figure 3 presents the fluid dynamics profiles for the velocity of the clean air in the interior of the filtration system.

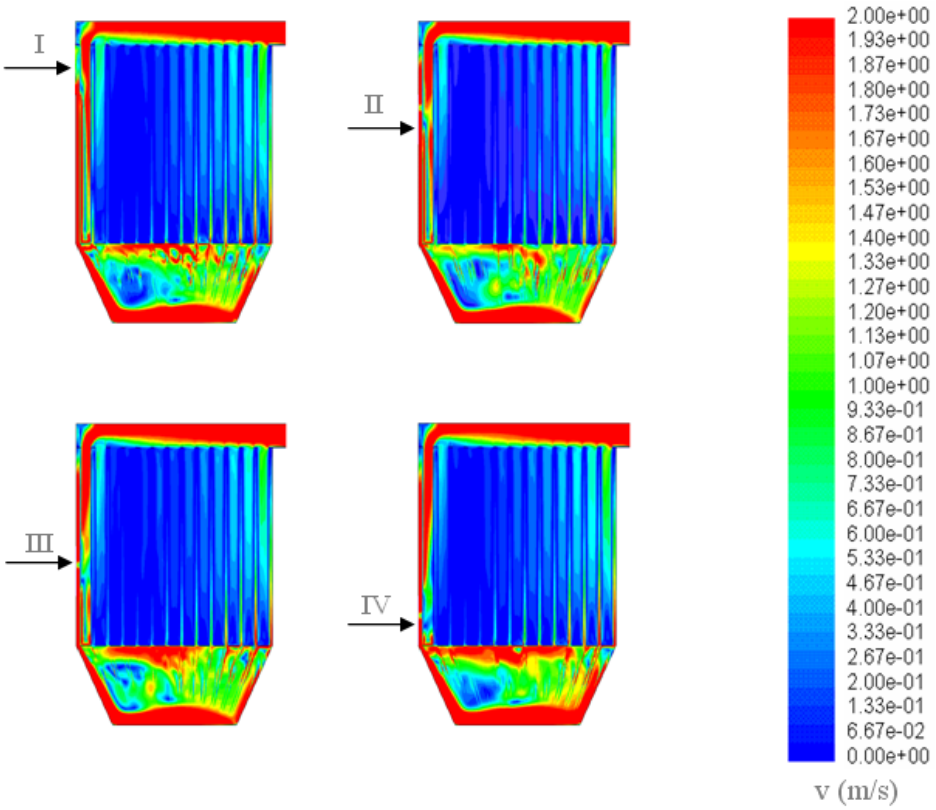


Figure 3 – Fluid dynamics of air in the filtration system.

According to the velocity profiles presented in Figure 3, it was possible to verify that, independently from the position of the air feed, the fluid has a preferential distribution through the bags. The simulations showed that bag "1" is the one that produced the highest rates of filtering amongst all bags. Probably, such behavior is due to the fact that it is subject to the highest gradients of velocity. Consequently, the position of bag "1" will probably make it wear prematurely when compared to the others. The same consideration can be made for bags "2", "12" and "13". On the other hand, it was seen that bags "3", "4" and "5" are those that provided the lowest filtering rates. Finally, the productivity of bags "6" to "11" did not change in the operational conditions used in this work. Figure 4 confirms the previous considerations and presents the rates of filtered mass produced by each one of the bags.

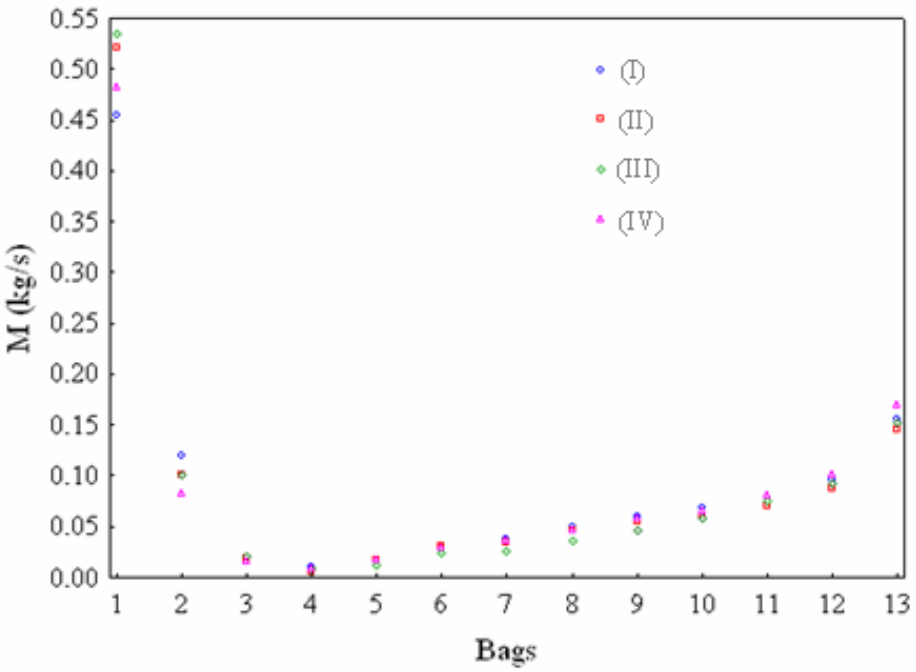


Figure 4 –Distribution of the filtered mass as a function of the positions of the bags and of the gas feed.

As illustrated by Figure 4, the position of the air feed practically did not influence the distribution of the fluid in the bags. Except for position "III", all the others lead to the same rates of filtering. Since the rate of filtered mass in each bag is practically indifferent to the feed positions "I", "II" and "IV", it is convenient to feed the fluids in position "IV", which produces the smallest pressure drop.

4. Conclusions

In the operational and geometric conditions used in this work, and based on the fluid dynamics simulations, the following conclusions can be drawn:

i) The pressure drop in the filtration unit was completely dependent on position of the feed;

ii) Position "IV" (near the base of the unit) provided the smallest pressure drop. This position probably allows longer filtration periods and the need for fewer cleaning operations, increasing the useful life of the bags (the most expensive part of the filtration unit);

iii) The distribution of the fluids through the bags was not uniform. It was seen that the flow showed preference for the bags adjacent to the feed (subject to the highest gradients of velocity, bags 1 and 2);

iv) The non-uniform distribution of the gas through the bags was indifferent to the position of the gas feed in the filtration unit.

Future works are suggested to evaluate the position of the gas feed in the bottom of the equipment, in order to homogenize the distribution of the fluid through bags and to

subject the filter structures to the same rate of wear/deterioration. Another suggestion would be the construction of a prototype to validate the results obtained in this work, both with and without particulate materials.

5**. Simbology**

M – Rate of fluid mass produced in each one of bags (kg/s);
v – Velocity of the fluid (m/s);
-Δp – Pressure drop between the feed and the outlet of the filtration unit (Pa).

References

A.C.M**.** Rodrigues, 2004. Estudo da Formação e Remoção de Tortas de Filtração de Gases em Filtros de Tecidos de Polipropileno e de Algodão. Tese (Doutorado)—Engenharia Química - Universidade Federal de São Carlos.

L.G.M. Vieira; C.A. Silva Jr.; J.J.R. Damasceno; M.A.S. Barrozo, 2007, A Study of the Fluid Dynamic Behaviour of Filtering Hydrocyclone. Filtration & Separation, v. 58, p. 282-287.

R.R. Damian, 2003, Desenvolvimento de um Sistema de Filtragem compacto para Usinas de Asfalto. Monografia – Engenharia Mecânica – Universidade Federal do Rio Grande do Sul.

S.M.S. Rocha; J.J.R. Damasceno; M.L.Aguiar; 2008, Estudo Estatístico e Otimização da Porosidade e Perda de Carga na Filtração Gás Sólido em Filtros de Tecidos In: XVII COBEQ, Recife, PE**.**

W.C. Hinds, 1999 Aerosol technology: Properties, behavior and measurement of airborne particles. Wiley Interscience.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Decomposition Techniques for Multi-Scale Structured Product Design: Subspace Optimization

Charles C. Solvason, Nishanth G. Chemmangattuvalappil,
Susilpa Bommareddy, Mario R. Eden
Department of Chemical Engineering, Auburn University, Auburn, AL 36849, USA, edenmar@auburn.edu

Abstract

Recent developments in the area of integrated process and product design have shown that products can be designed in terms of their properties without committing to any specific components a priori. Although current techniques make use of group contribution methods (GCM) to design molecules, there are many properties, atomic arrangements, and structures that cannot be represented using GCM. One approach to expand the capability of GCM to handle a more diverse range of solutions is to combine property clustering with decomposition techniques in a reverse problem formulation. This approach first utilizes multivariate characterization techniques to describe a set of representative samples, and then uses decomposition techniques such as principal component analysis (PCA) and partial least squares (PLS), to find the underlying latent variable models that describe the molecule's properties.

Keywords: Reverse problem formulation, principal component analysis, multi-scale product design, property clustering, molecular design

1. Introduction

One of the most common methods for designing molecules for specific end uses while minimizing computational expense has been group contribution (Gani et al., 2005, Marrero and Gani, 2001). In product design, the associated properties of concern are most often consumer attributes which do not have group contribution parameters (Hill, 2005). One way to address this concern is to map the consumer attribute data from the mega-scale into a set of properties on the macro scale that can be described by group contribution (Solvason et al., 2009a). This step is often performed via chemometrics which defines an empirical relationship through the use of design-of-experiments (DOE) and multivariate-linear-regression (MLR). Any uncertainty in the relationship between the attributes and properties is handled by increasing the size of the feasibility regions in the property domain and validating the enumerated candidate molecules in the attribute domain (Solvason et al., 2009b). Often the attribute-property relationship is poorly defined, which limits the effectiveness of this type of approach. An improvement is to map the attribute information down to a domain subspace that exhibits a stronger attribute-property relationship. The constraints on this new domain are that it must be linear in the constituent space and it must have the ability to be described by a molecular combinatorial technique. A key difference in this method is that the domain subspace is not required to be one of the known properties described by group contribution.

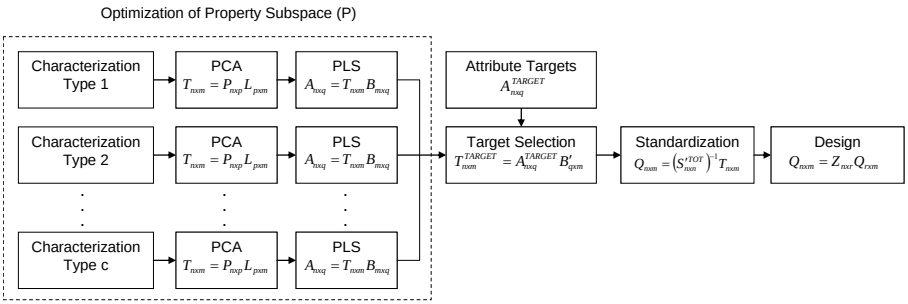


Figure 1: Method Flowchart

2. Methods

The objective of this paper is to enumerate all possible molecules that meet a set of target attributes using a domain subspace mapping function combined with molecular group theory applied to chemometric data. To achieve this objective, a reverse problem formulation is applied where the attributes are mapped down to a domain subspace comprised of properties that have better attribute predictive power than conventional GCM described properties. Several tools, such as characterization and decomposition can be used to find the domain subspace. Molecular combinatorial techniques and property clustering are then applied to find and interpret the solution to the reverse problem formulation. Fig. 1 illustrates the method.

2.1. Characterization Techniques

Characterization is a class of tools associated with the determination of not only chemical constituency or molecular structure, but also of larger structural characteristics describing the orientation and alignment of these molecules often called microstructure at the meso-scale. Some examples of characterization techniques include nuclear magnetic resonance (NMR), x-ray diffraction (XRD), and infrared spectroscopy (IR). The techniques are often applied to a training set of molecules defined by an experimental design used to explore the interesting facets of a set of property attributes. The added structural information available from the characterizations can be used to extend the group contribution method to higher orders as well as discern some orientation specific information. The choice of characterization technique for the specific design can be written into a MINLP, minimizing the predicted variance Q^2_m of the models describing the attribute-property relationships. To significantly reduce experimentation, improve prediction power, and ensure orthogonality of the models decomposition techniques are applied.

2.2. Decomposition Techniques

The most common decomposition is principal component analysis (PCA). By definition, PCA uses the variance-covariance structure to compress the property data to principal component data that contains much of the system variability. This result also improves the interpretation of the data structure by consolidating multiple property effects into single, underlying latent variables which are devoid of colinearity. If the original property data are of the same type, then the eigenvalues can be considered measures of the contrasts, or loadings $\boldsymbol{L_{pxm}}$ of the original variables; and the eigenvectors are referred to as scores $\boldsymbol{T_{nxm}}$ (Johnson and Wichern, 2007).

$$P_{nxp} = T_{nxm} L'_{mxp} \tag{1}$$

In most cases the first 2 or 3 principal components can account for 80% to 90% total variance. The remaining components can be removed without much loss of information (Johnson and Wichern, 2007). The relationship between the principal component scores $\boldsymbol{T_{nxm}}$ and the attribute properties $\boldsymbol{A_{nxq}}$ is then developed using a PLS model of a new DOE factorial design where the scores are varied between their high (+1) and low (-1) levels:

$$A_{nxq} = T_{nxm} B_{mxq} \tag{2}$$

In Eq. 2 $\boldsymbol{B_{mxq}}$ are the regressed coefficients found using PLS. It should be noted that the PLS model uses a separate set of scores and loadings to develop the relationship between $\boldsymbol{A_{nxq}}$ and $\boldsymbol{T_{nxm}}$.. The overall predicted variance of this model is then estimated using the following:

$$R^2_{predM} = \prod_j^q \left(R^2_{predj}\right)^{\frac{1}{q}},\ j \in q \tag{3}$$

Choosing the characterization that maximizes the prediction power in Eqn. 3 for a given number of experiments is a procedure that can be written as a MINLP. Care must be taken to ensure that an appropriate model fitness, as measured by R^2, is not sacrificed to improve the prediction power, R^2_{predM}. Also, although this method can select the appropriate characterization, it is experimentally costly since it requires that each design point be tested by each characterization technique until the appropriate one for the molecular design is chosen. To reduce the required experimentation, molecular design techniques are applied.

2.3. Molecular Design

In order to use characterization to extend GCM to handle complex structures and orientations in molecular design, the dissemination of characterization data must follow a set of rules designed to shadow those developed by Marrero and Gani (2001). First, the characterization must be able to completely quantify each individual molecular group used in the design. Second, in some characterizations, progressively larger groups completely contain the information of the smaller groups, but also contain corrections for 2nd order, 3rd order, structural, and orientation effects. This hierarchal nature is handled by specifying that all groups should be combined such that the largest functional group is specified first, then the second largest, and so on. Third, in some situations only partial overlaps may occur, for which the method will fail to specify any corrections to the first order combination at the overlap interface. To minimize this impact, it is specified that the groups be built such that the combinations occur across the C-C bond which carries the smallest amount of information (Marrero and Gani, 2001). Unlike conventional GCM, the molecular group specific variables are case specific. For instance, in many places in literature, typical absorbance wavelengths are published for IR and NMR spectroscopy. In order to convert these spectra to the underlying molecular group scores $\boldsymbol{T_{rxm}}$, a case specific set of loadings $\boldsymbol{L_{pxm}}$ are needed, which are derived from the original DOE. The full molecule scores $\boldsymbol{T_{nxm}}$ are then estimated as shown in Eq. 4.

$$T_{nxm} = Z_{nxr} T_{rxm} \tag{4}$$

where $\boldsymbol{Z}_{nxr}$ are the number of structure and orientation specific groups in the molecules being designed.

2.4. Property Clustering

Property clustering is a tool used to improve the interpretation of the subspace properties by deconstructing the design problem into a Euclidean vector in the cluster domain and a scalar called the Augmented Property Index *AUP*. The clusters themselves are conserved surrogate properties described by property operators, which have linear mixing rules, even if the operators themselves are nonlinear. Methods for the application of group contribution methods for molecular design have previously been developed using property clustering by Eljack *et al.* (2007, 2008). To utilize the latent variables in the property clustering algorithm, it is important to recognize that the data structure of Eq. 5 follows a linear mixing rule:

$$T_{nxm} = P_{nxp} L_{pxm} \tag{5}$$

$$\psi^{mix}_{nxm} = X_{nxp} \psi_{pxm} \tag{6}$$

The pure properties $\boldsymbol{\psi}_{\text{pxm}}$ have the same structure as the loadings $\boldsymbol{L}_{\text{pxm}}$, thus the loadings can be thought of as the pure values of the principal components. Likewise, the response $\boldsymbol{\psi}^{\boldsymbol{mix}}{}_{\text{nxm}}$ data are predicted mixture properties and have the same structure as the score $\boldsymbol{T}_{\text{nxm}}$ data. That means that the mixture fraction $\boldsymbol{X}_{\text{nxp}}$ in the property models is related to the multivariate data in $\boldsymbol{P}_{\text{nxp}}$. However, there is a concerning difference between the two methods: the mixture fractions sum to 1 across the properties for each sample and the multivariate data $\boldsymbol{P}_{\text{nxp}}$ do not. In order for latent variable models to be utilized in property clustering, it is necessary to standardize the latent variable structure by dividing $\boldsymbol{X}_{\text{nxp}}$ and $\boldsymbol{T}_{\text{nxm}}$ by $\boldsymbol{S}^{\boldsymbol{T}}{}_{\boldsymbol{nxn}}$, resulting in Eqn. 8.

$$S_k = \sum_{i}^{p} X_{ik} \ , \ k \in n \tag{7}$$

$$Q_{nxm} = R_{nxp} L_{pxm} \tag{8}$$

The new $\boldsymbol{Q}_{\text{nxm}}$ matrix now represents standardized scores or mixtures. The loadings matrix $\boldsymbol{L}_{\text{pxm}}$ remains unchanged and the $\boldsymbol{R}_{\text{nxp}}$ matrix now represents fractions of loadings whose cumulative sum is one for each run. Unfortunately, although the components sum to one, they are sometimes negative due to the mean-centering of the multivariate property data prior to PCA. The constraint that the fractions must be between 0 and 1 is removed with no effect on the associated mathematics, only on their interpretation. At this point, the loadings $\boldsymbol{L}_{\text{pxm}}$ are the underlying latent variable domain subspace. Both full molecules and molecular group subspace properties $\boldsymbol{Q}_{\text{nxm}}$ or $\boldsymbol{Q}_{\text{rxm}}$ can be found by multiplying the latent variables $\boldsymbol{L}_{\text{pxm}}$ by the associated fractions $\boldsymbol{R}_{\text{nxp}}$ or $\boldsymbol{R}_{\text{rxp.}}$ Since the molecular and group subspace property relationships in Eqs. 2 and 4 were derived using a decomposition technique, the constraints imposed by decomposition should also be observed for any new molecules or mixtures created.

$$Q_{nxm} = Z_{nxr} Q_{rxm} \tag{9}$$

The molecular design of Eq. 9 is a representation of a linear mixture of the underlying latent variable subspace properties, all of which are linear in nature; it is essentially a *linear mixture of linear mixtures*. This observation assumes that any nonlinearity in the attribute system is handled by the attribute-latent property relationship and not the molecule-group subspace property relationship (Muteki and MacGregor, 2006).

3. Case Study – Acetaminophen Tablet Design

Three attributes that are important to direct compression tablet manufacturing are disintegration time, crushing strength, and ejection force. These attributes have been notoriously difficult to analyze based on traditional mixing design because of the complex and highly nonlinear nature of pharmaceutical excipients. In order to better control these attributes, they are mapped down to a domain subspace where they can be approximated as linear combinations of molecular group parameters. The domain subspace was found to be characterized by three properties using a training set of 24 excipients $\boldsymbol{P}_{\mathbf{nxp}}$ (Gabrielsson *et al.,* 2003). To reduce the number of parameters in the subspace, decomposition was performed using PCA. Although the number of parameters could be reduced, it was decided to keep all of them for illustrative purposes. Using PLS models developed from the training set, consumer specific set of targets $\boldsymbol{A}_{nxq}$ were mapped to the domain subspace $\boldsymbol{Q}_{\mathbf{nxm}}$ as shown in Table 1.

Table 1. Design Targets

Subspace Targets	Q1	Q2	Q3
UL	2.00	2.00	1.00
LL	1.00	0.00	0.00

The molecular groups identified by the characterization were identified as CH, CH_2, OH, CHO, O, CH_2-O, CHOH, CH_2OH, $CHCH_2OH$, CHCHO, Ocyc, α-pyranose, β-pyranose, and cellulose. Using the molecular design procedure outlined in Fig. 1, the molecular groups $\boldsymbol{Q}_{\mathrm{rxm}}$ were combined to build a complete set of molecules $\boldsymbol{Q}_{\mathrm{nxm}}$ matching the domain targets as shown in Fig. 2 and Table 2. Of note was the selection of α-pyranose over β-pyranose compounds as candidates, effectively removing the traditional microcrystalline cellulose excipient from consideration. Other candidate excipients are also shown in Table 2.

Table 2. Designed Molecules

Candidate Molecules	Q1	Q2	Q3
CH2OH-CH2OH	1.22	1.20	0.61
CH2OH-(α)pyranose-CH2OH	1.85	1.08	0.90
CH2OH-CH(OH)-CH2OH	1.79	1.07	0.43
CH2OH-CH2-O-CH2OH	1.75	1.06	0.15
CH2OH-O-CH2OH	1.75	1.03	0.14
CH2OH-CH(CH2OH)-CH2OH	1.80	1.45	0.57
OH-CH2-OH	1.17	0.47	0.13
OH-(α)pyranose-CH2OH	1.80	0.35	0.42
OH-CH(CH2OH)-CH2OH	1.75	0.72	0.09
OH-CH2-CH2-CH2-OH	1.75	1.29	0.23

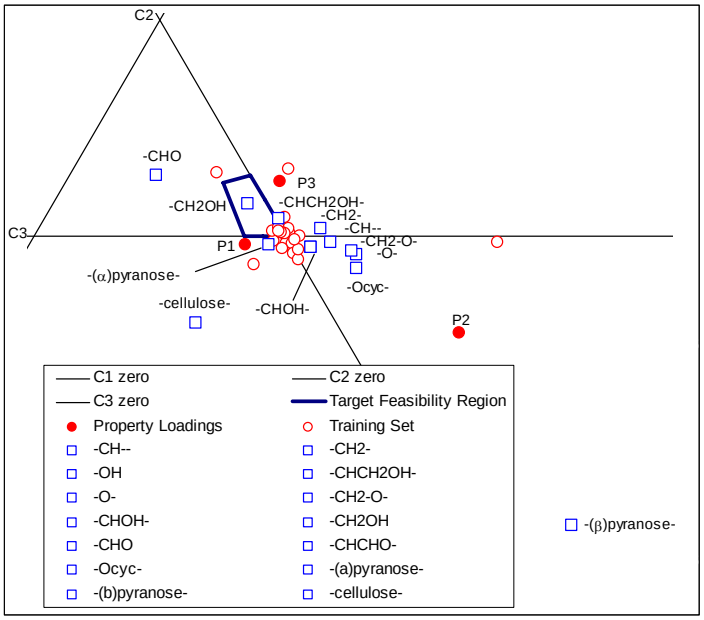


Figure 2: Molecular Design Cluster Diagram

4. Conclusions

In summary, the combination of property clustering and principal component analysis offers many insights and advantages for structured product design. In particular CAMD problems are no longer hindered by a lack of structure information in the molecular design. Rather, the uncertainty in predicting large molecular structures has now been removed from the models and replaced solely with the experimenter's ability to choose appropriate training sets, for which many proven techniques exist. This represents a useful addition to the existing CAMD methodology. Furthermore, the method is universal in nature and can be extended to include many characterization techniques.

References

N.G. Chemmangattuvalappil, F.T. Eljack, C.C. Solvason, M.R. Eden (2008), *Comp. & Chem. Eng.* 33(3), pp. 636-643.

F.T. Eljack, M.R. Eden, V. Kazantzi, M.M. El-Halwagi (2007), *AIChE Journal*, 53(5), 1232-1239.

F.T. Eljack, M.R. Eden (2008), *Comp. & Chem.Eng.* 32(12), 3002-3010.

J. Gabrielsson, N. Lindberg, M. Palsson, F. Nicklasson, M. Sjostrom, and T. Lundstedt (2003), *Drug Development and Industrial Pharmacy*, 29(10), 1053-1075.

R. Gani, P.M. Harper, M. Hostrup (2005), *Ind.Eng.Chem.Res.* 44, 7262-7269

R.A. Johnson and D.W. Wichern (2007), Applied Multivariate Statistical Analysis. *Pearson Prentice Hall,* Upper Saddle River, NJ.

M. Hill (2004), *AIChE Journal*, 50(8), 1656-1661.

J. Marrero, R. Gani (2001), *Fluid Phase Equilibria*, 182-183.

K. Muteki and J.F. MacGregor (2006), *Chemom. & Intell. Lab.* Sys., 85, 186-194

C.C. Solvason, N.G. Chemmangattuvalappil, F.T. Eljack, M.R. Eden (2009a), *Ind. Eng. Chem. Res.*, 48(4), 2245-2256

C.C. Solvason, N.G. Chemmangattuvalappil, M.R. Eden (2009b). *Comp.& Chem. Eng.* (In Press).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Simultaneous Solution of Process and Molecular Design Problems using an Algebraic Approach

Susilpa Bommareddy, Nishanth G. Chemmangattuvalappil, Charles C. Solvason and Mario R. Eden

Department of Chemical Engineering, Auburn University, AL 36849, USA

Abstract

The property integration framework has allowed for simultaneous representation of processes and products and established a link between molecular and process design from a properties perspective. The simultaneous approach involves solving two reverse problems. The first reverse problem identifies the property targets corresponding to the desired process performance. The second reverse problem is the reverse of a property prediction problem, which identifies the molecular structures that match the targets identified in the first problem. Therefore, successful tracking of properties is the key in applying reverse problem formulation for integrated process and product design problems. In this contribution, an algebraic technique has been developed for solving process and molecular design problems simultaneously. The molecules are identified to meet the process performance defined in terms of properties. Since, both process and molecular property operators target the same optimum process performance, the set of inequality expressions can be solved simultaneously to identify the molecules that meet the desired process performance. Since this approach is based on an algebraic algorithm, any number of properties can be tracked simultaneously.

Keywords: Property Operators, Reverse problem formulation, Molecular Design

1. Introduction

The identification of optimal molecule(s) corresponding to optimum process performance is a challenging issue. To achieve this goal, it is necessary to consider the aspects of both process and product design simultaneously. The concept of reverse problem formulation (RPF) has helped to formulate the integrated process-product design problems without leading to MINLP formulations (Eden *et al.*, 2003a). Techniques have been developed by Eden *et al.* (2003a, 2003b) for the identification of property targets corresponding to the optimum process performance using a visual approach. Algorithms to identify the molecules that meet the process targets have been developed by Eljack *et al.* (2007) and Chemmangattuvalappil *et al.* (2009). However, there is a need for a simultaneous algorithm that can be used to identify the suitable molecules corresponding to specific process performance. Since the process performance may depend upon several properties, the algorithm should be able to solve for any number of properties.

2. Normalized Process Property Operators

Property operators are functions of the original properties tailored to obey linear mixing rules (Shelley and El-Halwagi, 2000) so that they can be used to add the property

contributions from different streams linearly to obtain the property operator corresponding to the product stream:

$$\psi_j(P_{jm}) = \sum_{s=1}^{N_s} x_s \psi_j(P_{js}) \tag{1}$$

Where $\psi_j(P_{js})$ is the property operator of the j^{th} property P_{js} of stream s, x_s is the fractional contribution and N_s is the total number of streams. To allow comparison of different property values on a single platform, property operators are made dimensionless by dividing it by an appropriately chosen property reference. The general dimensionless property operator mixing rule can be estimated as follows:

$$\Omega_j(P_{jm}) = \sum_{s=1}^{N_s} x_s \Omega_j(P_{js}) \tag{2}$$

3. Group Contribution Method for Property Estimation

The group contribution method (GCM) provides a convenient way to relate molecular structures to properties. In GCM, the property function $f(Y)$ of a compound is estimated as the sum of the property contributions of all the molecular groups present in the structure (Marrero and Gani, 2001):

$$f(Y) = \sum_i N_i C_i + \sum_s N_s C_s + \sum_t N_t C_t \tag{3}$$

N_i, N_s and N_t are the numbers of first, second and third order groups and C_i, C_s, C_t are their respective property contributions.

4. Molecular Property Operators

The property models in GCM can be used for molecular design. Molecular property operator ψ^M is defined as (Eljack *et al.*, 2007):

$$\psi_j^M(P_{jm}) = \sum_{g=1}^{N_g} n_g P_{jg} \tag{4}$$

Where, P_{jg} is the contribution of property j from group g, n_g is the total number of that group in the molecule.

5. Simultaneous Process and Molecular Design

Since, both process and molecular property operators target the same optimum process performance, the set of inequality expressions can be solved simultaneously to identify the molecules that meet the ideal process performance. Since the approach is based on an algebraic algorithm, any number of properties can be tracked.

5.1. Estimation of number of first order groups

The mathematical expression for the Free Bond Number (*FBN*), which is the number of free bonds in each molecular string, is (Eljack *et al.*, 2007):

$$FBN = \sum_{g=1}^{N_g} n_g FBN_g - 2\left(\sum_{g=1}^{N_g} n_g FBN_g - 1\right) - 2N_r \tag{5}$$

Where, N_r is the number of rings in the final molecule and FBN_g is the number of free bonds in each group. The following expressions can be developed to ensure the existence of a meaningful molecule.

$$n_g \geq 0 \quad , \quad \sum n_{gr} \geq 3 \text{ or } 0 \quad , \quad FBN = 0 \tag{6}$$

The property targets to be satisfied by the sink can be stated as:

$$P_j^{lower} \leq P_j \leq P_j^{upper} \tag{7}$$

The normalized property operators corresponding to these targets can be written as:

$$\Omega_j^{lower} \leq \Omega_j \leq \Omega_j^{upper} \tag{8}$$

It is evident that, for each property, there will be two inequality expressions; one for the minimum value and the other for the maximum value for a given sink (Qin *et al.*, 2004). The Normalized process property operator is expressed as a function of normalized molecular property operators to allow the integration of process and molecular design problems. Equations 4-9 are solved simultaneously to find the maximum values of n_g.

$$\Omega_j = f(\psi_j^M) \tag{9}$$

5.2. Molecular Design

In order to increase the accuracy of property based molecular design techniques, the effects of higher order molecular groups are to be considered while designing molecules. The following methodology has been developed to estimate the contributions of higher order groups (Chemmangattuvalappil *et al,* 2008):

Rule 1. Higher order groups can only be formed from complete molecular fragments.
Rule 2. If any of the higher order groups completely overlap some other higher order group, only the larger group must be chosen in order to prevent redundant description of the same molecular fragment.

So, if (k: n) is the set of first order groups that are the building blocks of one second order group, s, (n_{gk}:n_{gn}) is the number of those first order groups present in the molecule, η is the number of occurrences of one particular first order group in a selected second order group, n_{gs} is the number of possible second order groups from those first order groups, then:

$$n_{gs} = Min\left(\frac{n_{gk}}{\eta_k} : \frac{n_{gn}}{\eta_n}\right) \tag{10}$$

n_{gs} must be rounded down to the nearest integer number according to Rule 1. If Ω_{jg2} is the property contribution from the second order groups, the normalized property operator for the property contributions from second order groups, Ω_{ijs} is calculated as:

$$\Omega_{ijs} = \sum_{s=1}^{N_S} n_{gs}\Omega_{jg2} \tag{11}$$

If (n_{gk}:n_{gn}) has subsets of smaller second order groups (n_{gl}:n_{gm}), the number of those small second order groups, n^*_{gs} is found by:

$$n_{gs}^{*} = Min\left(\frac{n_{gl}}{\eta_l} : \frac{n_{gm}}{\eta_m}\right) - Min\left(\frac{n_{gk}}{\eta_k} : \frac{n_{gn}}{\eta_n}\right) \quad (12)$$

According to Rule 1, this must be rounded down to the nearest integer. The third order effects can be calculated using the same approach. The normalized property operator for molecule *i* can now be estimated as (Chemmangattuvalappil *et al,* 2009):

$$\Omega_{ij} = \Omega_{ijf} + \Omega_{ijs} + \Omega_{ijs}^{*} + \Omega_{ijt} + \Omega_{ijt}^{*} \quad (13)$$

Where Ω^{*}_{ij} corresponds to the property value from overlapped small groups. Now, utilizing the above rules the maximum number of second and third order groups are found for each first order group combination without invoking Rule 2. All possible combinations of the numbers $[0,n_g]$, $[0,n_{gs}]$ and $[0,n_{gt}]$ are generated subject to following constraints based on Rule 2.

$$n_{ga}^{x} + n_{ga}^{y} \leq \max(n_{ga}^{x}) \quad (14)$$

Where *a* is *s* or *t* indicating second or third order group and *x* is a group that is completely overlapped by *y*. The above method of generation enables the identification of isomers as the possibility of nonexistence of each second and third order groups is considered. The possible molecules are screened out by checking if the combination of all the groups satisfies Eqs. 4 and 7.

6. Case Study

A current gas treatment process uses fresh methyl diethanol amine, MDEA, ($HO-(CH_2)_4-CH_3N-OH$) and two other recycled process sources (*S*1, and *S*2) as a feed into the acid gas removal unit. Another process stream, *S*3, currently a waste stream could be recycled as a feed if mixed with a fresh source to allow the mixed stream properties to match the sink (Kazantzi *et al.*, 2007). Design objectives and requirements are to find a solvent that will replace MDEA as a fresh source and that will maximize the flow rate of all available sources (*S*1, *S*2 and *S*3). The following three properties are considered: critical volume (V_c), heat of vaporization (H_v) and heat of fusion (H_{fus}). Additionally, two thermal constraints are imposed on the synthesized molecules to make sure that the designed molecule will remain in liquid state at the process conditions and to prevent excessive solvent losses via evaporation. The process property targets and flow rate data for all streams (*S*1, *S*2 and S3) and the sink along with the property operators corresponding to each property are summarized in Table 1.

Table 1. Property data for gas purification

Pj	Property Operators	Property Bounds on Sink	S1	S2	S3
V_c	$V_{cM} = \sum_{s=1}^{Ns} x_s \cdot V_{cs}$	530-610	754	730	790
H_v	$H_{vM} = \sum_{s=1}^{Ns} x_s \cdot H_{vs}$	100-115	113	125	70
H_{fus}	$H_{fusM} = \sum_{s=1}^{Ns} x_s \cdot H_{fuss}$	20-40	15	15	20
Flowrate		300	50	70	30

The property operators for the target properties are given in table 2 and the property contributions of all the candidate groups are obtained from Marrero and Gani (2001).

Table 2. Property targets

P_j	ψ_j	Adjustable Parameter	Property Bounds
T_b	$\exp\left(\frac{T}{t_{b0}}\right)$	222.543	532-547
T_m	$\exp\left(\frac{T}{t_{m0}}\right)$	147.45	<380
V_c	$V_c - V_{c0}$	7.95	
H_v	$H_v - h_{v0}$	11.733	
H_{fus}	$H_{fus} - h_{fus0}$	-2.806	

Fourteen first order groups have been considered for molecular design. To insure water solubility and to reduce vapor pressure, the amine must have two or more –OH groups. To limit the extent of corrosion, only one amino group is allowed to be in the amine (N in the amino group either connects to H or C). Finally, to limit detrimental effects of direct exposure to the solvent, tertiary amines are ruled out in this case study.

The functions to link property operators found on process and molecular side are formed based on Eqs. 1, 2, 4 and 9. The combined property operator (Eq. 13) of the preselected first order groups in terms of unknown maximum number of each first order group and their molecular property data (Marrero and Gani, 2001) is found. For example, the generated expressions for critical volume are shown below:

$$V_{CM} = x_f * V_{Cf} + x_{S1} * 754 + x_{S2} * 730 + x_{S3} * 790$$

$$V_{CM} = x_f * (V_C + 7.95) + x_{S1} * 754 + x_{S2} * 730 + x_{S3} * 790$$

$$V_C - 7.95 = n_1 * 68.35 + n_2 * 56.28 + n_3 * 37.5 + n_4 * 30.61 + n_5 * 88.2 + n_6 * 74.03 + n_7 * 60.06$$
$$+ n_8 * 117.62 + n_9 * 76.36 + n_{10} * 77.04 + n_{11} * 95.15 + n_{12} * 99.16 + n_{13} * 16.01 + n_{14} * 52.96$$

Similar equations are written for all properties and the maximum number of each first order group is then found solving the equations subject to constraints given by Eqs. 5-8.

CH_3:2 CH_2:1 CH:1 C:1 OH:3 CH_3O:2 CH_2O:1 CHO:1 CH_2NH_2:1 $CHNH_2$:1 CH_3NH:1 CH_2NH:1 CHNH:1 CO:1

Now, different combinations of these first order groups are generated. Possible second and third order groups are screened out and their maximum number is calculated using Eq. 10 and their various possible combinations are generated for each combination of first order groups subject to Eq. 14. Equations 4 and 7 along with the thermal constraints are used to screen the molecules and the structural feasibility of the designed molecules is checked using Eqs. 5 and 6. The final solution is given in Table 3 below.

The generated molecular structures have the properties consistent with the property limits of the sink and also since the second step of the simultaneous design considers the elimination of each possible second and third order groups, successful detection of isomers is reached.

7. Conclusions

In this work, an algorithm has been developed that solves process and molecular design problems simultaneously to find the molecules that would fit the process. Here, an algebraic approach has been followed which ensured that any number of properties can be tracked. In future works, the algorithm will be extended to identify the best molecule considering cost and environmental impact while also concentrating on optimizing the fractional recycle of different feeds.

Table 3. Valid Molecules and their Properties

No.	Molecule	V_c	H_v	H_{fus}
1	2,6,8-trioxa-3-azanonane-4,7-diol	445.27	91.8	26.73
2	2-(hydroxymethoxy)-2-methoxy-1-(methoxyamino)ethanol	448.31	92	27.33
3	2-(2-amino-1-hydroxyethoxy)propan-2-ol	393.05	80.99	31.57
4	1-[(hydroxymethyl)amino]-1-(methoxymethoxy)ethanol	402.27	89.98	27.99
5	1-(hydroxyamino)-2-(hydroxymethoxy)propan-1-ol	375.68	106.44	25.98
6	{[1-(hydroxyamino)-2-methylpropan-2-yl]oxy}(methoxy)methanol	457.34	93.06	29.22
7	[hydroxy(methoxy)methoxy](methylamino)methanol	325.46	79.97	25.71
8	[2-(hydroxyamino)ethoxy](dimethoxy)methanol	436.79	92.21	30.66
9	[2-(hydroxyamino)-1-methoxyethoxy](methoxy)methanol	427.29	95.44	31.76
10	[2-(hydroxyamino)-1,2-dimethoxyethoxy]methanol	453.12	93.57	29.71
11	[(hydroxymethyl)amino](methoxy)(propan-2-yloxy)methanol	456.65	94.82	29.52
12	[(1-amino-2-methoxypropan-2-yl)oxy]methanediol	411.46	95.21	33.33

References

Chemmangattuvalappil N.G, Solvason C.C., Eljack F.T., Eden M.R. (2009). A Novel Algorithm for Molecular Synthesis Using Enhanced Property Operators. Computers & Chemical Engineering, 33(3), 636-643.

Eden M.R., Jørgensen S.B., Gani R., El-Halwagi M.M. (2003a). Reverse Problem Formulation based Techniques for Process and Product Design. Computer Aided Chemical Engineering 15A, 451-456.

Eden M.R., Jørgensen S.B., Gani R., El-Halwagi M.M. (2003b). Property Cluster Based Visual Technique for Synthesis and Design of Formulations. Computer Aided Chemical Engineering 15B, 1175-1180.

Eljack F.T., Eden M.R., Kazantzi V., El-Halwagi M.M. (2007). Simultaneous Process and Molecular Design - A Property Based Approach. AIChE Journal, 53(5), 1232-1239.

Kazantzi V., Qin X., El-Halwagi M.M., Eljack F.T., Eden M.R. (2007): "Simultaneous Process and Molecular Design through Property Clustering Techniques", Industrial & Engineering Chemistry 46(10), 3400-3409

Marrero J., Gani R. (2001). Group Contribution Based Estimation of Pure Component Properties. Fluid Phase Equilibria, 183-184, 183-208.

Qin X., Gabriel F., Harell D., El-Halwagi M.M. (2004). Algebraic Techniques for Property Integration via Componentless Design. Ind. Eng. Chem. Res., 43, 3792-3798.

Shelley M.D., El-Halwagi M.M. (2000). Component-less Design of Recovery and Allocation systems: A Functionality–based Clustering Approach. Computers & Chemical Engineering, 24, 2081-2091

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Multiobjective multiproduct batch plant design under uncertainty: Application to protein production

Alberto Aguilar-Lasserre,[a] Rubén Posada-Gómez,[a] Giner Alor-Hernández,[a] Guilllermo Cortés-Robles,[a] Constantino Moras-Sánchez,[a] Catherine Azzaro-Pantel,[b] Luc Pibouleau,[b]

[a] *Division of Research and Postgraduate Studies, Instituto Tecnológico de Orizaba. Av. Instituto Tecnológico 852, Col. Emiliano Zapata. 09340 Orizaba, Veracruz, México*
[b] *Laboratoire de Génie Chimique-UMR 5503 CNRS/INP/UPS 5, Rue Paulin Talabot, BP1301, 31106 Toulouse Cedex1, France.*

Abstract

The design of batch plants necessary involves how equipment may be utilized, which means that plant scheduling and production must form an integral part of the design problem. The market demand for such products is usually changeable, and at the stage of conceptual design of a batch plant, it is almost impossible to obtain the precise information on the future product demand over the lifetime of the plant. This paper addresses the problem of the optimal design of batch plants with imprecise demands and proposes an alternative treatment of the imprecision by using concepts of fuzzy logic. For this purpose, we extended a Multi-Objective Genetic Algorithm (MOGA) developed in previous works, taking into account simultaneously three criteria, i.e. minimization of the investment cost, the operation cost and the total production time. The case of study is a multiproduct batch plant for the production of proteins taken from the literature. The methodology provides a set of scenarios that are helpful to the decision's maker and constitutes a very promising framework for taken imprecision into account in new product development stage.

Keywords: Multi-objective optimization, genetic algorithm, fuzzy arithmetic.

1. Introduction

In recent years, there has been an increased interest in the design of batch processes due to the growth of specialty chemical, pharmaceutical, and related industries, because they are a preferred operating method for manufacturing small volumes of high-value products. The market demand for such products is usually changeable, and at the stage of conceptual design of a batch plant, it is almost impossible to obtain the precise information on the future product demand over the lifetime of the plant. However, decisions must be made on the plant capacity. This capacity should be able to balance the product demand satisfaction and extra plant capacity in order to reduce the loss on the excessive investment cost or that on market share due to the varying demands on products. Consequently, the mission of the designer, assisted by traditional tools, may prove to be hazardous and makes essential the resort to a more robust approach.
The design of multiproduct batch plants has been an active area of research over the past decade. Most of the work has been yet limited to deterministic approaches, wherein the problem parameters are assumed to be known with certainty. However, in reality

there can be uncertainty in a number of factors such as processing times, costs, demands, and not all the requirements placed by the technology of the process and the properties of the substances are defined. To cope with this, there has been increased interest in the development of different types of probabilistic models that explicitly take into account the various uncertainties (Sahinidis, 2003). For instance, Wellons and Reklaitis proposed an MINLP model for the design of batch plants under uncertainty with staged capacity expansions. Based on the structure of multiproduct batch plants, Straub and Grossmann (1992) developed an efficient procedure to evaluate the expected stochastic flexibility, embedded within an optimization framework for selecting the design (size and number of parallel equipment). Two-stage stochastic programming approaches have also been applied for design under uncertainty (Ierapetritou and Pistikopolous (1996); Harding and Floudas (1997); Petkov and Maranas (1998); Cao and Yuan (2002).

The key point in the optimal design of batch plants under imprecision concerns modeling of demand variations. The most common approaches treated in the dedicated literature represent the demand uncertainty with a probabilistic frame by means of Gaussian distributions. Yet, this assumption does not seem to be a reliable representation of the reality, since in practice the parameters are interdependent, leading to very hard computations of conditional probabilities, and do not follow symmetric distribution rules. In this work, fuzzy concepts and arithmetic constitute an alternative to describe the imprecise nature on product demands.

For this purpose, we extended a multi-objective genetic algorithm, developed in previous works (Dietz et al. 2005). For instance, the optimal design of a multiproduct batch chemical plant is not only to minimize the investment, but also to minimize the operation cost and to minimize the total production time, simultaneously. The paper is organized as follows. Section 2 is devoted to process description and problem formulation. Section 3 presents a brief overview of fuzzy set theory involved in the fuzzy framework within a multi-objective genetic algorithm. The presentation is then illustrated by some typical results in Section 4.

2. Process description and problem formulation

In previous works (Dietz et al., 2005, 2006), batch plant design was carried out minimizing the investment cost and the production system was represented using discrete event simulation techniques in order to take into account different production policies. Two strategies for campaign policies were tested, either monoproduct or multiproduct. In this work, only the monoproduct campaign policy was considered, so that the computation of cycle time can be easily implemented using the classical formulation proposed in (Montagna et al., 2000), involving size and time equations as well as constraints. The model uses the formulation presented in (Modi and Karimi, 1989), then modified in (Xu et al., 1993), for multiproduct batch plant design formulation. It considers not only treatment in batch stages, which usually appears in all kinds of formulation, but also represents semi-continuous units that are part of the whole process (pumps, heat exchangers…). Let us recall that a semi-continuous unit is defined as a continuous unit working by alternating low-activity and normal activity periods. Besides, this formulation takes into account short-term or mid-term intermediate storage tanks. They are used to divide the whole process into sub-processes, in order to store materials corresponding to the difference of each sub-process productivity. This representation mode confers to the plant a major flexibility for numerical resolution, by preventing the whole process production from being

paralysed by one bottleneck stage. Therefore, a batch plant is finally represented by series of batch stages (B), semi-continuous stages (SC) and storage tanks (T).
The model considers the synthesis of I products treated in J batch stages and K semi-continuous stages. Each batch stage consists of m_j out-of-phase parallel items of same size V_j. Each semi-continuous stage consists of n_k out-of-phase parallel items of same processing rate R_k. The item size (continuous variables) and equipment number per stage (discrete variables) are bounded. The *S-1* storage tanks, of size V_s*, divide the whole process into S sub-processes.
For instance, the optimal design of a multiproduct batch chemical plant is not only to minimize the investment, but also to minimize the operation cost and to minimize the total production time, simultaneously. Then, an optimal design problem becomes a MOOP as following:

$$Min\ (Cost) = \sum_{j=1}^{J} (m_j a_j V_j^{\alpha_j}) + \sum_{k=1}^{K} (n_k b_k R_k^{\beta_k}) + \sum_{s=1}^{S} (c_s V_s^{\gamma_s}) \tag{1}$$

$$Min\ (D_p) = \sum_{i=1}^{N} \sum_{j=1}^{M} C_{E_j} \frac{Q_i}{B_{is}} + C_{O_i} Q_i \tag{2}$$

$$Min\ (H) = \sum_{i=1}^{I} \frac{Q_i}{B_{is}} T_{is}^{L} \tag{3}$$

The problem statement involves four forms of different constraints as reported in literature: (i) Dimension constraints: every units has to restrict to its allowable range. (ii) Time constraint: the summation of available production time for all products is inferior to the total production time. (iii) Limiting cycle time for product i. (iv) Volume constraints: the volume V_j has to be able to process all the products i.

3. Overview of fuzzy multiobjective genetic algorithm approach

3.1. Representation of fuzzy demands and time horizon due-date
In the context of engineering design, an imprecise variable is a variable that may potentially assume any value within a possible range because the designer does not know *a priori* the final value that will emerge from the design process. The fuzzy set theory was introduced by Zadeh to deal with problems in which a source of vagueness is involved. The proposed approach involves arithmetic operations on fuzzy numbers and quantifies the imprecision of the demand by means of fuzzy sets (trapezoidal). In this case, the flat line over the interval (q_2,q_3) represents the precise demands with an interval of confidence at level α=1, while the intervals (q_1,q_2) and (q_3,q_4) represent the "more or less possible values" of the demand. The result of the *total production time* $\widetilde{H}_i$ and the operation cost are treated and analyzed through fuzzy numbers. The time horizon $\widetilde{H}$ represented by a fuzzy expression (rectangle) and the production time $\widetilde{H}_i$ (trapezoidal) are fuzzy quantities, three different cases for determination of the criterion may occur, as shown in figure 1. The temporal criterion selected is called "common surface", representing the intersection between the sum of the production time (trapezoid) and the horizon of time to respect (rectangle). The calculation of the criterion depends on each case: for instance, case1 illustrate the solutions which arrive just in time.

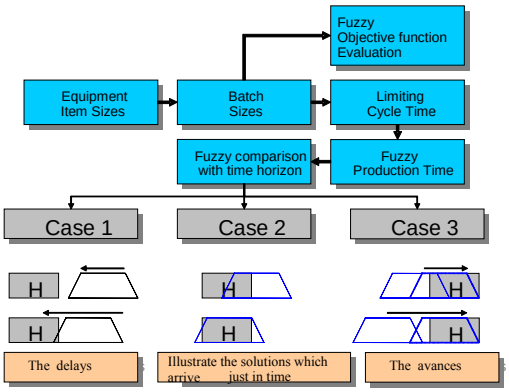


Figure 1 – Fuzzy evaluation procedure in the GA

3.2. Overview of multi-objective genetic algorithm approach

The multi-objective genetic algorithm presented elsewhere (Dietz et al., 2005) was then extended to take into account the fuzzy nature of both demand and horizon time. Let us mention that that the same encoding procedure was adopted since no fuzzy parameter is involved at that stage. The originality of this proposed research is that fuzziness is maintained throughout the computation procedure and no defuzzification is operated so that fuzzy results are proposed to the decision's maker. The tunable parameters of the GA will also not be discussed here. More detail can be found in (Dietz et al., 2006). They involve addition, subtraction, taking the maximum of two fuzzy numbers (mainly at the selection stage and at the Pareto sort procedure), through the extension principle of (Zadeh, 1975). Although there is a large body of literature that deals with the comparison of fuzzy numbers, the approach proposed by (Liou and Wang, 1992) was finally adopted here.

3.3. Treatment of an illustrative example

The case of study is a multiproduct batch plant for the production of proteins taken from the literature (Montagna et al., 2000), (Pinto et al., 2001). The batch plant involves eight stages for producing three recombinant proteins, on one hand two therapeutic proteins, Human insulin (I) and Vaccine for Hepatitis B (V) and, on the other hand, a food grade protein, Chymosin (C). This example is used as a benchmark since short-cut models describing the unit operations involved in the process are available.

The GA parameters are the following ones: Population size 200 individuals, number of generations 400 iterations, crossover probability 40%, mutation probability 30% and the stop criterion considered in this study concerns a maximum number of generations to reach. For the considered example, table 1 shows the values for processing times, size factor for the units, cost data, and the production requirement for each product quantifying the imprecision of the demand by means of fuzzy numbers representing the "more or less possible values".

4. Typical results

The method proposes a sufficiently large range of compromise solutions making it possible to the decision's maker to tackle the problem of the final choice, with relevant

information for his final choice. The result obtained of the multi-objective optimization problem involves 200 non-dominated solutions: among them, 4 represent the case 2, where the production time H_i (trapezoidal) is within of time horizon H (rectangle), 2 represent the first case (the delay) and 194 represent the case 3(the advance).
This study leads to three different scenarios as a fuzzy decision-making approach to preliminary design. Table 2 shows the solution that minimizes the investment cost for case 2. Tables 3 and 4 present the results that minimize the total production time and minimize the investment cost (cases 1 and 3), respectively.

	Processing time $\tau_{i,j}$ (h)						***Size factors (1/kg)***					
	B1	*B2*	*B3*	*B4*	*B5*	*B6*	*B1*	*B2*	*B3*	*B4*	*B5*	*B6*
Human Insulin (I)	1.15	3.98	9.86	5.28	1.2	3.57	8.28	6.92	9.7	2.95	6.57	10.6
Chymosin (C)	5.95	7.52	7.01	7	1.08	5.78	5.58	8.03	8.09	3.27	6.17	6.57
Vaccine for Hepatitis B (V)	3.96	5.07	6.01	5.13	0.66	4.37	2.34	9.19	10.3	5.7	5.98	3.14
γ_i	0.4	0.29	0.33	0.3	0.2	0.35						

Unit price for product i ($/Kg)			***Coefficients $c_{i,j}$***						I=(428260, 432630, 441370, 445740) C=(317520, 320760, 327240, 330480) V=(252840, 255420, 260580, 263160) $\tilde{H}$ =(5950, 5950,6240,6240) Rectangle
	C_P	C_O	*B1*	*B2*	*B3*	*B4*	*B5*	*B6*	Cost of Fermenter=\$250V$^{0.6}$
I	0.70	0.08	0.2	0.36	0.24	0.4	0.5	0.4	Cost of Micro Filter=\$250V$^{0.6}$
C	0.74	0.1	0.15	0.5	0.35	0.7	0.42	0.38	Cost of Homogenizer=\$250V$^{0.6}$
V	0.80	0.07	0.34	0.64	0.5	0.85	0.3	0.22	Cost of Ultra filter=\$250V$^{0.6}$
									Cost of Extractor=\$250V$^{0.6}$
									Cost of Chromatographic=\$250V$^{0.6}$

Minimum size =250 l
Maximum size = 10 000 l

Operating cost factors						
	B1	*B2*	*B3*	*B4*	*B5*	*B6*
C_E	20	30	15	35	37	18

Table 1 – Data used in example

Product	B_{is} kg	T_{Li} h	*Optimal objective function value*	*Storage Tanks*
I	639.6	3.9	*Cost* = 595,266.56 [\$]	V_s = 1737.22 [l]
C	808.6	5.6	$\tilde{D}_p$ = [302914.3 306005.3 312187.2 315278.2] [\$]	
V	576.4	4.8	$\sum \tilde{H}_i$ = [5998.05 6059.25 6181.66 6242.8] [h]	

Table 2 – Fuzzy optimal design of batch plant for case 1 (delay)

Product	B_{is} kg	T_{Li} h	*Optimal objective function value*	*Storage Tanks*
I	636.4	3.9	*Cost* = 601269.4 [\$]	V_s = 1736.2 [l]
C	831.8	5.6	$\tilde{D}_p$ = [299022.5 302073.81 308176.3 311227.5] [\$]	
V	621.4	4.8	$\sum \tilde{H}_i$ = [5962 6022.8 6144.5 6205.3] [h]	

Table 3 – Fuzzy optimal design of batch plant for case 2

Product	B_{is} kg	T_{Li} h	*Optimal objective function value*	*Storage Tanks*
I	642.9	3.9	*Cost* = 598996.9 [\$]	V_s = 1730.6 [l]
C	808.2	5.6	$\tilde{D}_p$ = [300165.7 303228.6 309354.4 312417.3] [\$]	
V	602.1	4.8	$\sum \tilde{H}_i$ = [5937.3 5998.3 6119.4 6180.08] [h]	

Table 4 – Fuzzy optimal design of batch plant for case 3 (advance)

5. Conclusions

In this paper, we have proposed a fuzzy approach to the treatment of imprecise demands in the batch design problem. In real application, designers not only require to minimize

investment, but also to minimize the operation cost and horizon time, simultaneously. An example was used to show the significance of the proposed approach. The results show that a set of compromise solutions is generated to the decision's maker, with an acceptable degree of imprecision affecting the defined criteria, which seems more realistic than a classical crisp approach. This approach will reduce the risk of making design decisions incorrectly. The results obtained on the treated example have shown that three different scenarios were obtained as a fuzzy decision-making approach. Finally, this framework provides an interesting decision-making approach to design multiproduct batch plants under conflicting goals.

References

D.M. Cao, X.G. Yuan, 2002, Optimal design of batch plants with uncertain demands considering switch over of operating modes of parallel units. Industrial Engineering and Chemistry Research, 41, 4616.

A. Dietz, C. Azzaro-Pantel, L. Pibouleau, and S. Domenech, 2005, A Framework for Multiproduct Batch Plant Design with Environmental Consideration: Application To Protein Production, Industrial Engineering and Chemistry Research, 44, pp. 2191-2206

A. Dietz, A. Aguilar-Lasserre, C. Azzaro-Pantel, L. Pibouleau, S. Domenech, 2006, Multiobjective multiproduct batch plant design under uncertainty, Computer Aided Chemical Engineering, Volume 21, Part 2, 2006, pp. 2249-2254

S.T. Harding, CA. Floudas, 1997, Global optimization in multiproduct and multipurposebatch design under uncertainty. Industrial Engineering and Chemistry Research, 36, 1644.

T.S. Liou, M.J. Wang, 1992, Ranking fuzzy numbers with integral value, Fuzzy Sets System, vol. 50, 247.

A. K. Modi, I.A. Karimi, 1989, Design of multiproduct batch processes with finite intermediate storage, Computer and Chemical Engineering, 13, pp 127-138

J.M. Montagna A.R. Vecchietti, O.A. Iribarren, J.M. Pinto J.A. Asenjo, 2000, Optimal design of protein production plants with time and size factor process models. Biotechnol. Prog. 16, 228-237

M.G. Ierapetritou, E.N. Pistikopolous, 1996, Batch plant design and operations under demand uncertainty. Industrial Engineering and Chemistry Research, 35, 772.

S.V. Petkov, C.D. Maranas, 1998, Design of single-product campaign batch plants under demand uncertainty. AIChE Journal, 44, 896.

J.M. Pinto, J.M. Montagna, A.R. Vecchietti, O. A. Iribarren, J. A. Asenjo, 2001, Process performance models in the optimisation of multiproduct protein production plants. Biotechnology and bioengineering, 74 (6), 451- 465

N.V. Sahinidis, 2003, Optimization under uncertainty: state-of-the-art and opportunities. In Proceedings of the Conference on Foundations of Computer Aided Process Operations FOCAPO 2003, Coral Springs, USA, 153.

D.A. Straub, I.E., Grossmann, 1992, Evaluation and optimization of stochastic flexibility in multiproduct batch plants. Computers and Chemical Engineering, 16, 69.

H.S. Wellons, G.V. Reklaitis, 1989, The design of multiproduct batch plants under uncertainty with staged expansion. Computers and Chemical Engineering, 13, 115.

X. Xu., G. Zheng S. Cheng S., 1993, Optimal design of multiproduct batch chemical process – Mixed simulated annealing. Journal of Chemical Engineering, 44, pp. 442-459

L.A. Zadeh, 1975, The concept of a linguistic variable and its application to approximate reasoning, Inform. Sci., 8, 199.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Computer simulation of hydrolytic polymerization of SLS Nylon 6 to correlate polymer characteristics and operational parameters

Maria Carolina B. Costa, [a] Maria Ingrid R. Barbosa,[a] André L. Jardini,[a] Marcelo Embiruçu,[b] Rubens Maciel Filho[a]

[a]*School of Chemical Engineering, State University of Campinas, UNICAMP, P.O. Box 6066, 13083-970, Campinas/SP, Brazil. E-mail address: mcarol@feq.unicamp.br*
[b]*Polytechnical School of Federal University of Bahia, Prof. Aristides Novis Street, 02 – Federação, Salvador/BA, Postcode: 40210-910, Brazil*

Abstract

In recent years, numerical simulation and optimization of chemical reactors in general, and polymerization reactors in particular, have become increasingly popular as a means of understanding the behavior of reactors and improving their performance. The commercial importance of Nylon 6 has stimulated a considerable amount of research directed towards the modeling and simulation of its polymerization through the hydrolytic route in various industrial reactors. Computer simulations based on the mathematical modeling of the polymerization process are of paramount importance for quality control and operational optimization of the existing plant, as well as in the design of similar new plants. The major objective in this work will be to apply modeling and simulation tools for development of new application for Nylon 6. The intention is to be able to correlate reaction parameters and operational parameters (as monomer concentration, water and additives, operational temperature), to final desired properties of the obtained polymer – especially to molecular weight (aiming the polymer production matching specifications to a desired application). In addition, that correlation analysis allows evaluating the variation on polymers production rate – (higher production that minimizes production costs). To make possible to meet such objectives its coupled experimental design procedure together with the nylon reactor mathematical modeling and literature information. The developed procedure showed to be an interesting and important tool to develop products with the required specification whereas keeping lower production costs.

Keywords: simulation, hydrolytic polymerization, Nylon 6

1. Introduction

With the recent trend of building small flexible plants that are close to the market of consumption, there has been renewed interest in batch processes. Since commonly the production volumes are low, batch plants are often multiproduct facilities in which the various products share the same pieces of equipment. This requires that the production in these plants be scheduled.

Considering the economic scale of Nylon 6 production, it is not surprising that a number of polymerization models based on first principles have appeared [1-4]. These models attempt to apply knowledge of kinetics, physical and thermodynamic properties, mass and energy transport, and phase equilibrium to simulate a manufacturing process.

Several examples may be mentioned in which the use of an accurate model allows to improve the product quality, increase the production rate, and reduce production costs. The hydrolytic polymerization of ε-caprolactam is an important commercial process and has drawn the attention of many researches. It will be considered in this work as step required for the development of nylon as an engineering plastic.

1.1. Commercial Nylon 6 Manufacturing Process

The reactor feed typically contains fresh caprolactam monomer, recycled monomer, water and desired additives [5], such as chain-terminating agents (monofunctional amines and carboxylic acids). The process covers the monomer hydrolyzes, the conversion to a polymer, and the building up the polymer molecular weight as a function of the polymer application.

Because the molten polymer is in equilibrium with water and monomer, the byproduct water must be removed by vaporization to increase the molecular weight. However, vaporization of the water may also remove a significant amount of caprolactam, which is recovered and recycled.

In conventional processes, the polymer is then pelletized and leached with hot water to remove low-molecular-weight extractables, such as residual monomer and cyclic oligomers. The polymer is then dried and shipped elsewhere. Figure 1 depicts a conventional Nylon 6 production scheme.

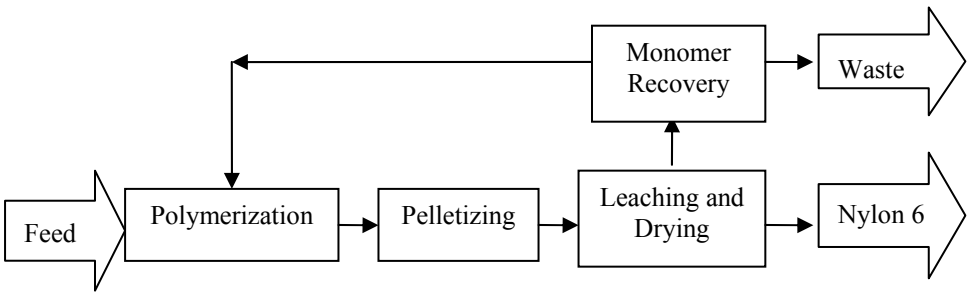


Figure 1 – Block diagram of a conventional Nylon 6 polymerization process.

1.2. Nylon 6 polymerization kinetics

The initial stage for the kinetic model development of reaction consists of the definition of chemical species that will be considered. In general, and also in this work, these species include water (W), ε-caprolactam (CL), cyclic dimmer (CD), monofunctional acid termination, such as acetic acid (AA), monofunctional amin (cyclohexilamin) and polymer.

For all equilibrium reactions considered here that do not involve chain terminator, it were adopted scheme and kinetic proposed by Arai *et al.* (1981), largely utilized in the majority of Nylon 6 kinetic studies published. For the monofunctional acid termination reaction, the scheme proposed by Gupta and Kumar (1987) was adopted. For the termination reaction scheme of the monofunctional amin, a similar behavior to the amino terminal groups was assumed (two reactions, without considering the reactions with oligomers).

The main studies of polymerization reaction modeling of Nylon 6 that exist in literature ignore the formation of superior cyclic oligomers, only considering the dimmers formation. These studies are according to the assumption that the majority of cyclic oligomers exists as dimmer.

Arai *et al.*[6] have presented the accepted standard regarding the chemistry and kinetics of the hydrolytic polymerization of ε-caprolactam. Their reaction mechanism includes

the ring opening of caprolactam (CL), polycondensation, polyaddition of CL, ring opening of a cyclic dimmer (CD), and polyaddition of a CD, as given in Table I.

Table I –Nylon 6 hydrolytic polymerization reaction mechanisms[a]

Reaction name	Equilibrium reaction
Ring opening of caprolactam	$W + CL \underset{k_1'}{\overset{k_1}{\leftrightarrow}} P_1$
Polycondensation	$P_m + P_n \underset{k_2'}{\overset{k_2}{\leftrightarrow}} P_{m+n} + W$
Polyaddition of caprolactam	$CL + P_n \underset{k_3'}{\overset{k_3}{\leftrightarrow}} P_{n+1}$
Ring opening of cyclic dimmer	$W + CD \underset{k_4'}{\overset{k_4}{\leftrightarrow}} P_2$
Polyaddition of cyclic dimmer	$CD + P_n \underset{k_5'}{\overset{k_5}{\leftrightarrow}} P_{n+2}$

[a]P_1 is aminocaproic acid. P_n is a Nylon 6 molecule with a degree of polymerization *n*.

In this work, one more equilibrium reaction to this kinetic scheme was added: the termination reaction with a monofunctional acid, such as acetic acid:

$$AA + P_n \underset{k_2}{\overset{k_2'}{\leftrightarrow}} P_{nx} + W$$

It was assumed that this reaction follows the same kinetics as the polycondensation reaction, as in reference 2.
The analysis of cyclic oligomers higher then dimers was ignored for the sake of simplicity. Experimental data are required to identify the kinetic parameters are shown in the next section.

2. Experimental

The theoretical objective of this work is studying an Experimental Unit of batch polymerization, that comprehends a polymerization reactor (autoclave), a nitrogen supplement system (cylinders), a vacuum system (vacuum pump), a water supplement system for refrigeration and cleaning, a polymer granulation system (granulator), and a drying system to polymer (vacuum greenhouse). The processing recipe is described below.

Researches for development of new applications for Nylon 6 intend to correlate reaction parameters and operational parameters to final properties of the obtained polymer. To make possible that analysis was made an experimental design, after an ample literature examination. As a result, the selected variables to be controlled (since they have a considerable influence on the molecular weight of the obtained polymer and also on its production rate) were: monomer concentration (CL), water concentration (W), acetic acid concentration (AA) and temperature (T).
Table II presents the controlled variables (factors) and their variation levels used in the experimental design. This experimental design made possible an analysis of controlled

variables influence on polymer characteristics, in special on molecular weight, as well as on polymerization reaction rate.

Table II- Experimental design 2^4 with central point

Factor	**Levels**		
	-1	**0**	**+1**
CL	1700g	2000g	2300g
W	19,6g	37,1g	54,6g
AA	1,87g	3,57g	5,27g
T	240°C	260°C	280°C

3. Results and Discussion

In the experimental design obtained, the confidence limit chosen was 95%. Tables III and IV depict the variance analysis (ANOVA) for production rate and molecular weight, obtained through experimental design and simulations made with deterministic software of polymerization process modeling studied here.

Figures 2 and 4 show Pareto charts of the controlled variables effects. Analyzing these diagrams, the effect of the manipulated operational variables on production rate and polymer molecular weight can be evaluated. On the other hand, in Figures 3 and 5, the deviations between observed values and predicted values (calculated by models obtained through experimental design) can be observed.

3.1. Polymerization Rate

Through Figure 2, it can be observed that the production rate is strongly influenced by temperature and variation of water concentration. Production rate increases as temperature and water concentration increase. There is a synergic effect between these controlled variables. The variation effect of monomer (CL) and acetic acid (AA) concentrations were not significant on values range utilized. Figure 3 shows the diagram of predict values (by model) versus observed values. Through ANOVA and Figure 3, it can be observed that the model is satisfactory.

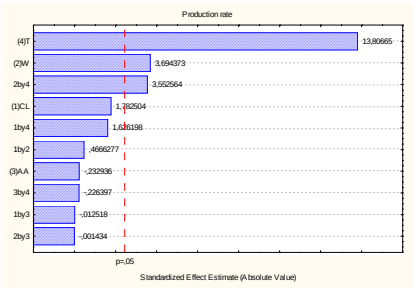


Figure 2 – Pareto chart for production rate

Table III – ANOVA in the form to analyze of the production rate model

	SS	**Freedom Degree**	**MS**	**$F_{calc.}$**	**$F_{tab.}$**	**$F_{calc./tab.}$**
Regression	0,000163	10	0,0000163	24,44	4,06	6,02
Residue	0,000004	6	$6,67*10^{-7}$			
Total	0,000167	16				

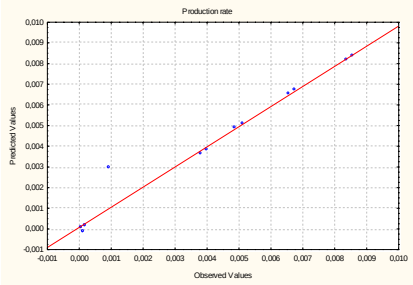


Figure 3 – Predicted values x observed values for production rate.

3.2. Molecular Weight

Through Figure 4, it can be seen that molecular weight is influenced by all controlled variables. This property decreases as water and acetic acid concentrations increase and it increases as the monomer concentration and temperature increase. There is a weak antagonist effect between water and monomer concentration. Figure 5 shows the diagram of predict values (by model) versus observed values. Through ANOVA and Figure 5, it can be concluded that the model is satisfactory.

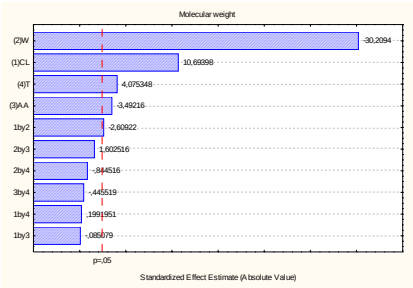


Figure 4 – Pareto chart for molecular weight.

Table IV – ANOVA in the form to analyze of the molecular weight model

	SS	**Freedom Degree**	**MS**	**$F_{calc.}$**	**$F_{tab.}$**	**$F_{calc./tab.}$**
Regression	557344250	10	55734425	106,61	4,06	26,26
Residue	3136709	6	522784,83			
Total	560480959	16				

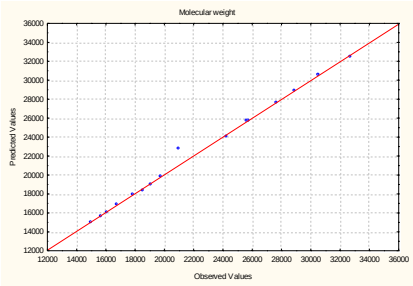


Figure 4 – Predicted values x observed values for molecular weight.

Figure 5 shows the response surfaces for production rate and molecular weight, considering the controlled variables more significant.

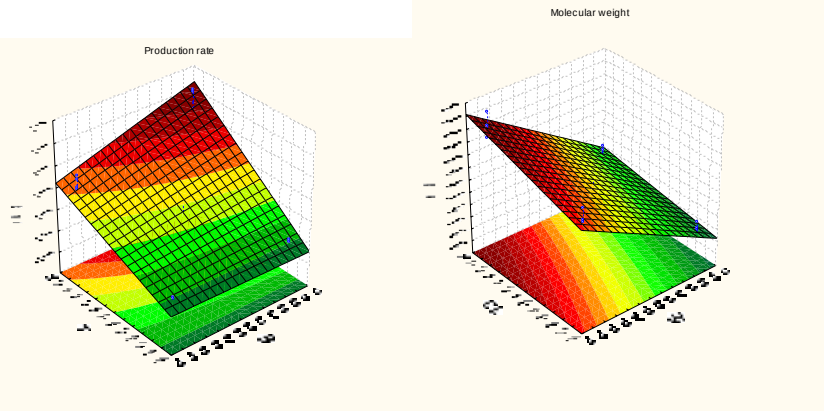


Figure 5 – Response surfaces for production rate and molecular weight.

4. Conclusions

A deterministic mathematic model was developed, enabling an influence analysis (of some controlled variables) on polymerization rate and on molecular weight of obtained polymers. This model can be used to propose optimization strategies aiming high conversions, assuming as restrictions the desired polymer properties for specific applications. This is specially required when rapid prototyping process are considered. The use of desired properties as restrictions is an interesting approach in the formulation of the optimization problem that, together a reliable mathematical model, allow the production of polymers with very tight specific properties.

References

1. Giori, C.; Hayes, B. T., 1970, Hydrolytic Polymerization of Caprolactam. I. Hydrolysis-Polycondensation Kinetics. ***J. Polym. Sci., Polym. Chem. Ed.***, v.8, p.335-349.
2. Agrawal, A. K.; Devika, K.; Manabe, T., 2001, Simulation of hydrolytic polymerization of nylon-6 in industrial reactors: Part I. Mono-acid-stabilized systems in VK tube reactors. ***Ind. Eng. Chem. Res.*** v.40, p.2563-2572.
3. Gupta, S. K.; Kunzru, D.; Kumar, A.; Agrawal, K. K., 1983, Simulation of nylon 6 polymerization in tubular reactors with recycle. ***J. Appl. Polym. Sci.***, v.28, p.1625-1640.
4. Gupta, A.; Gupta, S. K.; Ghandi, K. S.; Mehta, M. H.; Padh, M. R.; Soni, A. V.; Ankleswaria, B. V., 1992, Modeling of hydrolytic polymerization in a semi-batch nylon-6 reactor. ***Chem. Eng. Commun.***, v.113, p.63-89.
5. Seavey, K. C.; Khare, N. P.; Liu, Y. A., 2003, A new phase-equilibrium model for simulating industrial nylon-6 production trains. ***Ind. Eng. Chem. Res.***, v.42, p.3900-3913.
6. Arai, Y.; Tai, K.; Teranishi, H.; Tagawa, T., 1981, Kinetics of hydrolytic polymerization of ε-caprolactam: 3. Formation of cyclic dimmer. ***Polymer***, v.22, p.273-277.
7. Gupta, S. K.; Kumar, A., 1987, *Reaction engineering of step-growth polymerization.* Plenum Press: New York, p.190.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Mathematical Model of Carbon Dioxide Absorption into Mixed Aqueous Solutions

Patricia Mores[a], Nicolas Scenna[a,b] and Sergio Mussati[a,b]
[a]*CAIMI, UTN Facultad Regional Rosario, Zeballos 1341, S2000BQA, Rosario, Argentina.*
[b]*INGAR/CONICET Instituto de Desarrollo y Diseño, Avellaneda 3657 C.P. 3000 Santa Fe, Argentina. E-mail: mussati@santafe-conicet.gov.ar*

Abstract

In this paper, a mathematical model of CO_2 chemical absorption system using MDEA (MethylDiEthanolAmine) and PZ (Piperazine) aqueous solutions is investigated. Precisely, the complex reactive absorption behavior is modeled by an NLP mathematical model. The resulting mathematical model is implemented in GAMS and CONOPT is used as NLP solver. The proposed model will allow to optimize the operating conditions to maximize the CO_2 capture. The model is successfully validated using data from the literature.

Keywords: CO_2 Alkaline Absorption, Optimization of Chemical Absorption Process.

1. Introduction

In recent years, the removal of CO_2 from industrial gas streams has become important. This has resulted from the environmental concern for reduction of greenhouse gas emissions from industrial sources. CO_2 is considered to be the largest contributor to the global warning problem, and is thus the major target for reduction.
Recent studies on the costs of different processes for CO_2 capture from power plants show that post-combustion CO_2 removal by absorption with chemical solvents has the potential to remain the most important process also in the future.
Extensive research activities are directed towards the optimization of the process and also the search for new environmentally friendly and energy efficient absorbents. Several models are available to analyze the solubility of CO_2 in aqueous solutions of alkanolamines and to correlate the equilibrium CO_2 loading. The electrolyte-NRTL, Deshmukh-Mather and the Kent-Eisenberg models are widely used.
Benamor et al. (2005) measured the CO_2 solubility and carbamate concentration in MDEA, DEA and their mixtures at various temperatures (303-323 K) and CO_2 partial pressure (0.09-100 kPa). The experimental equilibrium data for the absorption the CO_2 was analyzed using Deshmukh- Mather model.
Liu et al. (1999) measured the CO_2 solubility in aqueous mixtures of MDEA and PZ for temperatures and CO_2 partial pressures ranging from 303 to 363 K and 13.16 to 935.3 kPa. Experimental data were correlated using a modified Deshmukh–Mather thermodynamic model, in which the effect of salts on Henry's constant, is taken into consideration. They not consider PZ - carbamate formation because they have worked at high loading CO_2, so only protonation amine is considered.
Bishnoi et al. (2002) studied vapor-liquid equilibrium, speciation and amine solubility for the PZ/ MDEA/ H2O/ CO_2 system and modeled their results with electrolyte NRTL model. They take into account the formation of PZ - carbamate and the protonation of

PZ.
Kundu et al. (2006a, 2006b) presented experimental results about solubility of CO_2 in DEA/AMP/H_2O system and DEA/MDEA/H_2O. A modified Clegg-Pitzzer equation to develop a model for both quaternary systems was used by the authors.
In this paper, the complex behavior of the CO_2 absorption CO_2 chemical absorption into MDEA (MethylDiEthanolAmine) and PZ (Piperazine) aqueous solutions is modeled by using mathematical programming. The modeling of the system leads to a Non Linear NLP model. The resulting mathematical model is implemented in GAMS and CONOPT is used as NLP solver.

2. Thermodynamic model

The CO_2-MDEA-PZ-H_2O system includes the chemical reaction equilibria in the liquid phase, and also the gas-liquid equilibrium between two phases.
The modified Deshmukh- Mather thermodynamic model (Deshmukh et al.) is used for the liquid phase and the Peng-Robinson equation of state to describe the vapor phase (Peng et al. 1976).

The following main chemical reactions take place in the aqueous phase for the CO_2-MDEA-PZ-H_2O system:

$$PZH^+ \overset{K_1}{\rightleftarrows} PZ + H^+ \tag{1}$$

$$MDEAH^+ + \overset{K_2}{\rightleftarrows} MDEA + H^+ \tag{2}$$

$$H_2O + CO_2 \overset{K_3}{\rightleftarrows} H^+ + HCO_3^- \tag{3}$$

$$H_2O \overset{K_4}{\rightleftarrows} H^+ + OH^- \tag{4}$$

$$HCO_3^- \overset{K_5}{\rightleftarrows} H^+ + CO_3^{2-} \tag{5}$$

Due to the chemical reactions in the amine-acid-gases-water systems, there is a strong deviation from the ideality in the liquid phase. This behavior is included in the equilibrium constants expressions by activity coefficients (γ_i), which are calculated with the extended Debye-Huckel equation (E. Guggenheim et al). Interaction parameters (β_{ij}) are obtained from Liu et al (1999).

$$\ln(\gamma_i) = -\left(\frac{A \times z_i^2 \times I^{0.5}}{1 + I^{0.5}} \right) + 2 \times \beta_{ij} \times m_j \tag{6}$$

The corresponding thermodynamic equilibrium constants expressions are computed as follows:

$$K_1 = \left(\frac{m_{PZ} m_{H^+}}{m_{PZH^+}} \right) \left(\frac{g_{PZ} g_{H^+}}{g_{PZH^+}} \right) \tag{7}$$

$$K_2 = \left(\frac{m_{MDEA} m_{H^+}}{m_{MDEAH^+}} \right) \left(\frac{g_{MDEA} g_{H^+}}{g_{MDEAH^+}} \right) \tag{8}$$

$$K_3 = \left(\frac{m_{HCO_3^-} m_{H^+}}{m_{CO_2}} \right) \left(\frac{g_{HCO_3^-} g_{H^+}}{g_{CO_2}} \right) \left(\frac{1}{a_w} \right) \tag{9}$$

$$K_4 = \left(m_{H^+} m_{OH^-} \right) \left(g_{H^+} g_{OH^-} \right) \left(\frac{1}{a_w} \right) \tag{10}$$

$$K_5 = \left(\frac{m_{CO_3^{2-}} m_{H^+}}{m_{HCO_3^-}} \right) \left(\frac{g_{CO_3^{2-}} g_{H^+}}{g_{HCO_3^-}} \right) \tag{11}$$

The condition of electro-neutrality is given by:

$$m_{PZ} + m_{PZH^+} = \frac{c_1}{r} \tag{12}$$

$$m_{MDEA} + m_{MDEAH^+} = \frac{c_2}{r} \tag{13}$$

$$m_{CO_2} + m_{HCO_3^{2-}} + m_{CO_3^{2-}} = \frac{\left(c_1 + c_2 \right) a}{r} \tag{14}$$

$$m_{H^+} + m_{PZH^+} + m_{MDEAH^+} = m_{OH^-} + m_{HCO_3^-} + 2m_{CO_3^{2-}} \tag{15}$$

The equilibrium constants and Henry's constant used in the calculation are listed in Table 1.

The resulting mathematical model involves 69 variables and 60 constraints. The model was implemented in General Algebraic Modeling System GAMS. CONOPT was used as the solver for the resulting NLP model.

Finally, it is important to notice that the mathematical model can lead to multiple and/or local optimal solutions due to the non-convex constraints involved in the mathematical model.

Table 1

Expression	Reference
$\ln K_1 = -22.086 + \frac{459.37}{T}$	Liu et al. (1999)
$\log K_2 = -14.01 + 0.0184T$	Barth et al. (1984)
$\ln K_3 = 235.482 - \frac{12092.1}{T} - 36.7816 \ln T$	Edwards et al. (1978)
$\ln K_4 = 140.932 - \frac{13445.9}{T} - 22.4773 \ln T$	Edwards et al. (1978)
$\ln K_5 = 220.067 - \frac{12431.7}{T} - 35.4819 \ln T$	Edwards et al. (1978)
$\ln H_{CO_2} = a_1 + \frac{a_2}{T} + \frac{a_3}{T^2}$	
$a_1 = 2.01874 - 2.83179\left(\frac{c_2}{r}\right) + 4.11932\left(\frac{c_2}{r}\right)^2 - 0.81256\left(\frac{c_2}{r}\right)^3$	Al –Ghawas et al. (1989)
$a_2 = 3135.49 + 1846.22\left(\frac{c_2}{r}\right) - 2612.63\left(\frac{c_2}{r}\right)^2 + 508.592\left(\frac{c_2}{r}\right)^3$	
$a_3 = -813702 - 295623\left(\frac{c_2}{r}\right) + 414660\left(\frac{c_2}{r}\right)^2 - 79674.6\left(\frac{c_2}{r}\right)^3$	

3. Study case, results and discussions

In this section, the proposed model is validated with the literature data and computational performance is presented. The following parameter set was used for the model validation: free PZ concentration: 0.7 mol/m^3, free MDEA concentration: 2.8 mol/m^3, temperature: 323 K.

Design data reported by Liu et al. (1999) are used for comparison purposes. For model testing, the proposed model was used as a "simulator" than an optimizer.

Figure 1 compares the model outputs with data taken from the literature. From the results shown in Fig. 1 it can be concluded that the values of all chemical species obtained from the model agree satisfactorily with those reported by Liu et al. (1999).

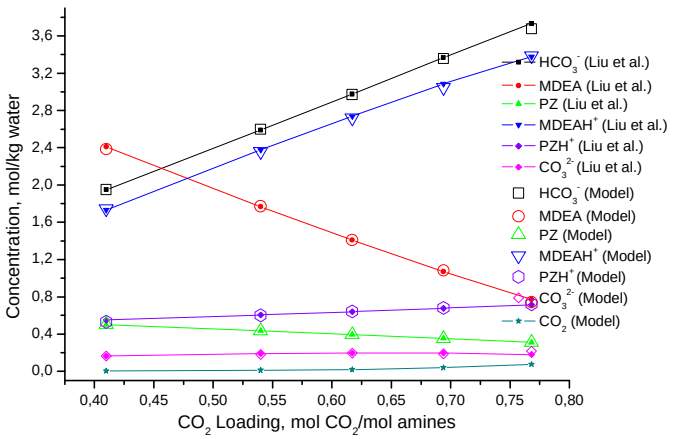


Figure 1. Model outputs vs. data taken from Liu et al (1999)

Computational aspects of the proposed model

Despite the simplicity of the model, the non-convex constraints involved by the mathematical model such as logarithms to compute the activity coefficients and bilinear terms can lead to computational problems and local solutions as well.

The convergence of the proposed model depends strongly on the initial values.

In order to guarantee the convergence of the proposed model, solutions from the simpler model are used as initial values to solve the detailed model. Thus, this initialization procedure improves the model convergence considerably.

The simplified model is based on the following assumptions:

- Effects on the mass and charge balance equations of the concentrations of H^+ and OH^- are neglected. This assumption is reasonable since the concentrations are low.
- All forms of the absorbed carbon dioxide are regarded as bicarbonate since the contents of CO_3 and CO_2 are also very low.

Figure 2 compares the values obtained from the simplified and detailed models.

As is shown in Fig. 2 it is possible to conclude that the simplified model can be effectively used to initialize the detailed model because the first model approximates well to the second one.

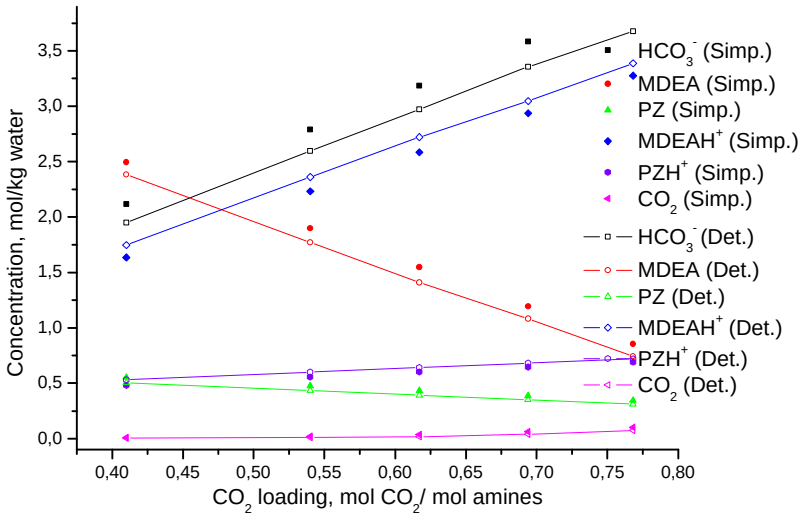


Figure 2. Comparison between simplied and detailed models

Conclusions and Future Work

A thermodynamic mathematical model for the CO_2 Capture with amines was developed. The model is successfully validated using data from the literature. The model proposed in this paper will be properly extended in order to develop a more detailed mathematical model which includes for example the energy balance.

The aim is to develop a model which allows to optimize the configuration and operating conditions of absorption and desorption columns. The assumption of simultaneous chemical and phase equilibria is not good enough for realistic process designs. Therefore, it is necessary the development of detailed and rate based models for packed or plate columns which allow to determine temperature and composition profiles along the columns. Those aspects will be addressed in future works and will be introduced gradually into the model here presented.

ACKNOWLEDGEMENTS.
The authors acknowledge financial supports from Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Agencia Nacional de Promoción Científica y Tecnológica (ANPCyT) and Technological National University (UTN).

Nomenclature

c_1	Fresh PZ concentration [kmol/m^3]
C_2	Fresh MDEA concentration [kmol/m^3]
p_{CO2}	CO_2 partial pressure [kPa]
A	Debye – Huckel coefficient
I	Ionic strength
z_i	Ionic charges
a_w	Water activity
H_{CO2}	Henry's constant in amine solution [kPa.m^3/kmol]
K_i	Equilibrium constant
T	Temperature [K]
m_i	Molality [mol/kg of water]
b_{ij}	Specific interaction coefficients [kg/mol]
g_i	Molal activity coefficients
r	Water Density in solution [kg/lt]

References

M. Kundu, S. Bandyopadhyay, 2006, Solubility of CO2 in water + diethanolamine+N-methyldiethanolamine, Fluid Phase Equilibria, 248, 158-167.

M. Kundu, S. Bandyopadhyay, 2006, Solubility of CO2 in water + diethanolamine+2-Amino-2-Methil-1- Propanol, J. Chem. Eng. Data, 51, 398-405.

A. Benamor, M.K. Aroua, 2005, Modeling of CO2 solubility and carbamate concentration in DEA, MDEA and their mixtures using the Deshmukh-Mather model, Fluid Phase Equilibria, 231, 150-162.

S. Bishnoi, G. Rochelle, 2002, Thermodynamics of Water/Methyldiethenolamine/Piperazine/ Carbon Dioxide, Ind. Eng. Chem. Res., V41, 604-612.

H. Liu, C. Zhang, G. Xu, 1999, A Study on Equilibrium Solubility for Carbon Dioxide in Methylendiethanolamine-Piperazine-Water Solution, Ind. Eng. Chem. Res., 38, 4032-4036.

H. Al-Ghawas, D. Hagewiesche, G. Ruiz Ibañez, O. Sandall, 1989,Physicochemical Propierties Important for Carbon Dioxide Absortion in Aqueous Methildiethanolamine, J. Chem. Eng. Data,39,817.

D. Barth, C. Tondre, J. Delpuech, 1984, Kinetics and Mechanims of the Reactions of Carbon Dioxide with Alkanolamines: A Discussion Concerning the Cases of MDEA and DEA, Chem. Eng. Sci.,39,1753.

R. Deshmuhk, A. Mather, 1981, A Mathematical Model for Equilibrium Solubility of Hydrogen Sulfide and Carbon Dioxide in Aqueous Alkanolamine Solutions, Chem. Eng. Sci, 36, 355.

T. Edwards, G. Maurer, J. Newman, J. Prausnits, 1978, Vapor-Liquid Equilibria in Multicomponent Aqueous Solutions of Volatile Weak Electrolytes, AIChE,24,966.

D. Peng, D. Robinson, 1976, New Two Constant Equation of State, Ind. Eng. Chem. Fundam., 15,59.

E. Guggenheim, 1935, The Specific Thermodinamic Propierties of Aqueous Solutions of Strong Electrolytes, Philos Mag.,19,588.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Heterogeneous batch distillation processes for waste solvent recovery in pharmaceutical industry

Ivonne Rodriguez-Donis[a], Vincent Gerbaud[b], Alien Arias-Barreto[a], Xavier Joulia[b]

[a]*Instituto Superior de Tecnología y Ciencias Aplicada (InSTEC). Quinta de Los Molinos, Ave. Salvador Allende y Luaces. AP6163. Ciudad Habana. Cuba.*
[b]*Chemical Engineering Laboratory. 5 rue Paulin Talabot - BP 1301. 3110. Toulouse. France*

Abstract

A summary about our experiences in the introduction of heterogeneous entrainers in azeotropic and extractive batch distillation is presented in this work. Essential advantages of the application of heterogeneous entrainers are showed by rigorous simulation and experimental verification in a bench batch distillation column for separating several azeotropic mixtures such as acetonitrile – water, n hexane – ethyl acetate and chloroform – methanol, commonly found in pharmaceutical industry.

Keywords: heterogeneous entrainer, batch column, azeotropic and extractive distillation

1. Introduction

Batch distillation is widely used in pharmaceutical and specialty chemical industry to recover valuable components from waste of solvent mixtures. The regular presence of azeotropes restricts severely the number of feasible separations. Azeotropic and extractive distillations are the most common alternatives encountered in the industry, requiring the addition of a third component to enhance the volatility of the components.
Synthesis and design of a new batch distillation process is done in two steps: (1) selection of an adequate entrainer along with the column configuration and the corresponding product cut sequence. It can be done using a recent computer tool RegSolExpert® (ProSim SA) including an exhaustive set of rules for selecting both, homogeneous and heterogeneous entrainers and (2) the process design for determining the main operating parameters by using a batch process simulator such as ProSim Batch (ProSim SA). In general, heterogeneous entrainers have a more privileged position than homogeneous because a higher number of ternary diagrams are suitable. Indeed, more than 70% of the feasibility rules included in RegSolExpert® match up to heterogeneous entrainers. In this paper, we present a general procedure to systematize the search for an appropriate azeotropic or extractive distillation process enabling the separation of non-ideal binary mixtures by introducing heterogeneous entrainers.
Primary criteria for selecting feasible heterogeneous entrainer comprise the non pollutant component and the market availability. For the separation of aqueous-organic mixtures, many feasible entrainers could be found among those inducing a liquid-liquid phase split with water and vice versa, water seems to be a promising candidate for organic wastes. Essential features of the application of heterogeneous entrainers are showed by rigorous simulation and experimental verification in a bench batch distillation column for separating three commonly found mixtures: acetonitrile – water, n hexane – ethyl acetate and chloroform – methanol.

2. Heterogeneous entrainers in batch azeotropic distillation (HABD)

Azeotropic distillation in a batch rectifying column is based on the complete addition of an adequate amount of entrainer together with the azeotropic mixture as initial charge in the boiler. One unstable node given by a component or mixture can be drawn as a top product. The general feasibility criterion for selecting a heterogeneous entrainer can be summarized as follows: choose a low, intermediate or high boiling entrainer (E) giving rise to a ternary system in which the heterogeneous azeotrope is an unstable node of the residue curve map. The separation of all components is better if: (a) the heterogeneous azeotrope is binary instead of ternary, (b) the heteroazeotrope composition defines an approx. equal liquid phase split in the decanter and (c) the immiscibility gap is large leading to a high purity of the distillate and the entrainer-rich phase. These features will be analyzed considering two azeotropic mixtures: acetonitrile – water and n hexane – ethyl acetate in a batch rectifier and a quaternary mixture in a middle vessel column.

2.1. Rectifying column configuration

The mixture acetonitrile - water displays a minimum boiling azeotrope with $x_{acetonitrile}$ = 0.676 at 76.8°C. The initial entrainer candidate list included 91 entrainers spanning a wide spectrum of chemical families. According to RegSolExpert®, only 21 entrainers matched with some feasibility criterion (Gerbaud et al., 2006). Four homogeneous entrainers were rejected because the separation was possible due to the curvature of the distillation boundaries and led to complex sequences of operating tasks. From the 17 remains heterogeneous compounds only three light entrainers did not form ternary azeotrope: dichloromethane (39.4°C), chloroform (61°C) and acrylonitrile (77.3°C). Only chloroform and acrylonitrile were tested. Acrylonitrile has been found earlier by manual search (Rodriguez et al. 2002) in an experimental database (Gmehling, J., 1994). Figure 1a and 1b display the residue curve map for chloroform and acrylonitrile, respectively.

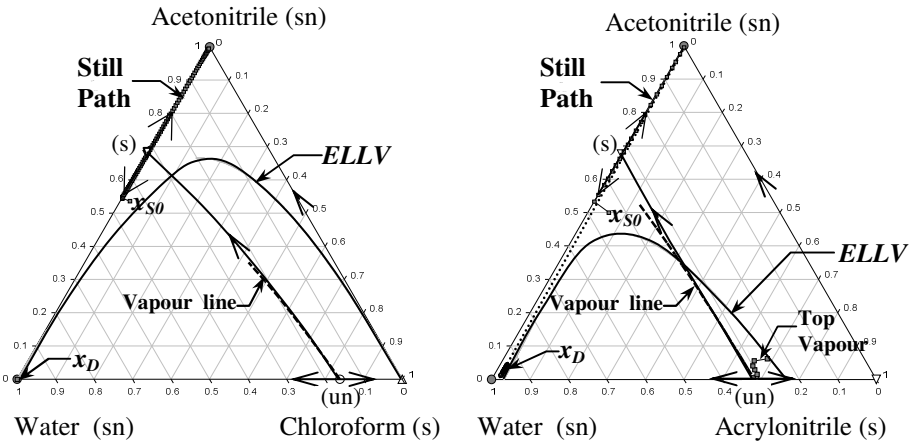


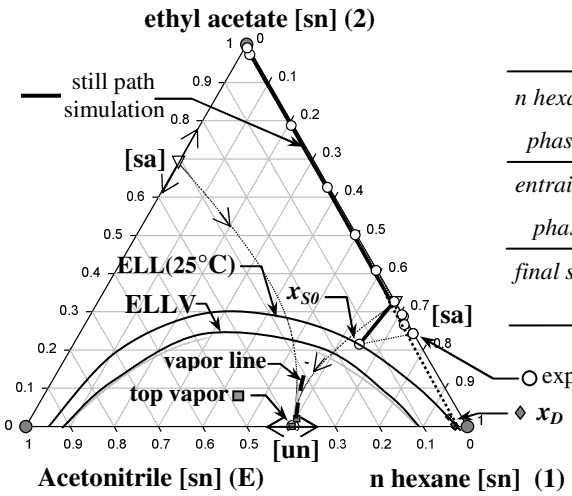
Figure 1. Residue curve map and simulation results of HABD.

Simulation results for both heterogeneous entrainers are reported in Table 1 considering only the reflux of the entrainer-rich phase and showing the superiority of chloroform in recovery and purity of all components. However, the operating times were very similar (about 6.5 hours) because both entrainers lead to a similar phase split ratio into the decanter from the condensed top vapour (about 0.16 in water-rich phase).

Table 1. Simulation results chloroform vs acrylonitrile

	Chloroform	Acrylonitrile (* experimental values)	
Aqueous	x_{Water}=0.9996	x_{Water}=0.947	*x_{Water}=0.946
Phase	$R_{ecovery}$=97.8%	$R_{ecovery}$=92.6%	*$R_{ecovery}$=89.0%
Entrainer	x_E=0.9990	x_E=0.712	*x_E=0.6916
Phase	$R_{ecovery}$=81.0%	$R_{ecovery}$=72.0%	*$R_{ecovery}$=69.0%
Final	$x_{acetonitrile}$= 0.9945	$x_{acetonitrile}$= 0.992	*$x_{acetonitrile}$= 0.995
Still	$R_{ecovery}$= 99.2%	$R_{ecovery}$= 95.6%	*$R_{ecovery}$= 91.8%

Experimental validation of HABD with acrylonitrile were carried out in a SHOTT batch distillation column available in a bench scale (Rodriguez et al., 2002; 2005) considering two reflux policies: case (1) only the reflux of acrylonitrile-rich phase from the decanter for the mixture acetonitrile – water and case (2) an additional reflux given by a portion of the condensed vapour by manipulating the open/close time of the solenoid valve for the separation of n hexane – ethyl acetate using acetonitrile. Experimental values are shown in Figure 1b and Table 1 for case (1) and Figure 2 displays the simulation and experimental results for the case (2) demonstrating the acceptable performance of the HABD process.



	simulation vs experiments	
n hexane phase	x_1=0.951 $R_{ecovery}$=95.0%	x_2=0.968 $R_{ecovery}$=94.5 %
entrainer phase	x_E=0.736 $R_{ecovery}$=54.5%	x_E=0.728 $R_{ecovery}$=48.4%
final still	x_2= 0.992 $R_{ecovery}$= 81.0%	x_2= 0.991 $R_{ecovery}$= 75.1%

Figure 2. Simulation and experimental results for n hexane – ethyl acetate – acetonitrile.

Those two examples bring important insights on HABD:

(a) less entrainer amount is required in the initial charge, compared to azeotropic distillation with a homogeneous entrainer.

(b) the liquid – liquid splitting of the top condensed vapour allows the straightforward withdrawal of one component (water or n hexane) as distillate,

(c) reflux policy can be versatile given by only the reflux of the entrainer-rich phase or by a combined reflux of both decanted phases (Rodriguez-Donis et al., 2002, Skouras et al., 2005),

(d) the adequate reflux policy displaces the typical distillation boundary of the residue curve map close to the binary side of azeotropic components. Therefore, the feasible region is not limited by the typical distillation boundary of the residue curve map anymore (Lang, 2006).

In HABD, the still path moves in a bigger region wherein it is connected to the binary heterogeneous vapour at the column top and the boiler can reach the stable node of acetonitrile at the end of the process (see still path in Figure 1 and 2 starting in x_{S0}). In the case of the sole reflux of the decanted entrainer-rich phase, the boundary of HABD is given by a hypothetical line joining the homogeneous azeotrope to be separated and the composition of the product-rich phase (dotted line in Figure 1b and 2). Although in both cases good agreement was obtained between simulation and experimental results, the reflux policy established in case (2) provided a more stable amount proportion of both liquid phases inside the decanter during the whole process. Theoretically as computed by simulation in Figure 1b and Figure 2, the composition of the top vapour stayed closer to the binary heteroazeotrope composition for case (2). As in continuous heterogeneous azeotropic distillation, liquid – liquid splitting occurred inside the top section in the SCHOTT column and experimental validations went well.

2.2. Middle vessel column configuration (MVC) with light or intermediate entrainer

Analysis of the heterogeneous azeotropic distillation in this column configuration has been made only by using rigorous simulation with ProSim Batch. MVC enable to split quaternary mixtures by a single task like the ternary mixtures using a batch rectifier. The separation of acrylonitrile (light entrainer) – acetonitrile – water – acetic acid mixture was studied elsewhere involving a single basic distillation region in the quaternary composition space (Rodriguez et al., 2001). The binary heteroazeotrope reaches the column top and a suitable reflux policy keep acetonitrile and acetic acid stay into the middle vessel and the boiler at the end of the process, respectively. Now, the use of an intermediate boiling heterogeneous entrainer (isobutyl acetate) is viable for separating a methanol – water – propanoic acid mixture (residue curve map in figure 3).

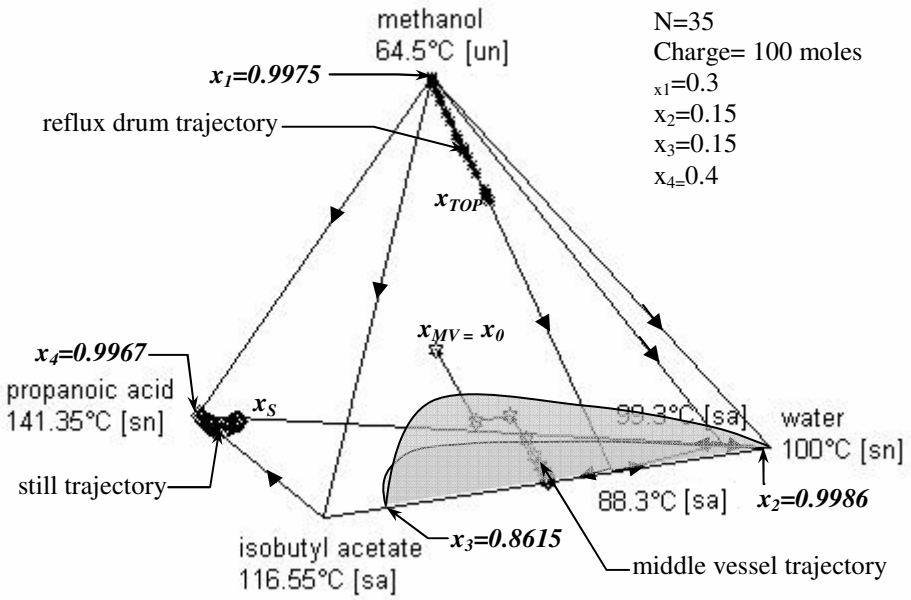


Figure 3. Separation of a quaternary mixture for HABD in a middle vessel column.

The homogeneous mixture methanol – water – propanoic acid exhibits a saddle azeotropic point and it can not be separated by conventional distillation process.

However, separation is feasible in a MVC by HABD with isobutyl acetate. Two binary azeotropes occur but in this case, both are saddle points and the composition space is divided by an unstable distillation boundary in two basic distillation regions. Methanol is the sole unstable node while water and propanoic acid are the stable nodes of their respective regions. The still initial charge into the still and the middle vessel (x_0) has to be located inside the region wherein propanoic acid is the stable node. Figure 3 displays the composition into the boiler (x_S) and the reflux drum (x_{top}) after the heat and liquid filling operation of the whole column. An adequate definition of the liquid holdup inside the reflux drum and the middle vessel allows the separation of four components by one single operating step under infinite reflux ratio. Methanol goes to the top reflux drum (x_1), propanoic acid concentrates into the boiler (x_4) and the saddle binary heteroazeotrope water (x_2) - isobutyl acetate (x_3) is kept into the middle vessel at the end of the process. Figure 3 shows the final composition and the composition path inside the still, the condenser and the middle vessel along with the column specifications operated under reflux total conditions until steady state is reached (after about 2 hours). Distillate removal or closed operation can start afterwards.

3. Heterogeneous entrainers in batch extractive distillation (HBED)

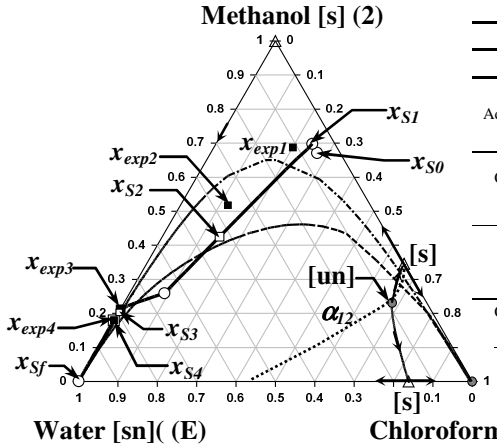
Extractive distillation is the most used process in the industry due to its operating simplicity. But, the selection of the entrainer is based mainly on a criterion: High boiling homogeneous entrainer forming no additional azeotropes. Despite the criterion simplicity, designing a new extractive distillation process is often complicated. Indeed, many suitable entrainers enhancing significantly the relative volatility of azeotropic components usually give rise to new binary and even ternary heterogeneous azeotropes, that are often saddles. Therefore, these entrainers were always considered as unfeasible although it is not the case using HEBD process (Rodriguez et al., 2003) because a more general criterion rules extractive distillation process feasibility (Rodriguez-Donis et al., 2009a, 2009b). Unlike to azeotropic distillation, the entrainer is fed continuously at some column position, leading to different column sections.

Our precedent works concerned three ternary mixtures: acetonitrile – water with butyl acetate, ethanol – water with ethyl acetate and chloroform – methanol with water (Rodriguez et al., 2003, Van Kaam et al. 2008). HEBD brings a few more advantages:

(1) HEBD in a batch rectifier is simpler than extractive batch distillation using a homogeneous entrainer fed at some intermediate tray, as is usual. Indeed, the heterogeneous entrainer can be fed at the column top leading only to an extractive column section and allowing the withdrawal of the saddle binary heteroazeotrope as top vapor. The further condensation of the top vapor generates two liquid phases similar to the heteroazeotropic distillation process,
(2) Water is often a suitable non pollutant heterogeneous entrainer for organic wastes treatment. And, organic compounds are suited for aqueous-organic wastes split.
(3) HEBD offers more ternary diagrams alternatives than the numerous HABD ones.

Experimental validation in the SCHOTT distillation column was recently performed for the separation of chloroform – methanol – water waste using the inner component, water, as heterogeneous entrainer. Figure 4 displays the simulation results (x_S) and the experimental verification (x_{exp}) of the process. Even if the performance of the process was good and adequate agreement between experiments and prediction was also obtained, extreme care must be taken with the regulation of the heat duty in the boiler. Accumulation of water into the boiler required a continuous increasing of the oil bath

temperature wherein the boiler was immersed. However, pollution of methanol in the chloroform-rich phase occurs rapidly if the correct ratio of the water flow rate and the vapor overflow inside the column is surpassed. Indeed, a significant amount of heterogeneous entrainer was required for the separation of all studied ternary system by HEBD given by the interception of the univolatility line α_{12} in the heterogeneous edge close to the entrainer vertex. A better column configuration must be considered such as middle vessel column or those having two boilers as proposed by Hua et al. (2007).



	Simulation Results		Experimental Results	
Products	Mol	x_{MOLAR}	Mol	x_{MOLAR}
Aqueous Phase Decanter	**0.045**	x_1 0.0016 x_2 0.0027 **x_3 0.9957**	**1.07**	x_1 0.0035 x_2 0.0211 **x_3 0.9754**
Chloroform Distillate	**5.2**	**x_1 0.9906** x_2 0.0048 x_3 0.0046	**4.86**	**x_1 0.9900** x_2 0.0044 x_3 0.0056
Final Still x_{Sf}	**72.3**	**x_1 0.0001** x_2 0.1795 x_3 0.8204	**69.2**	**x_1 0.0006** x_2 0.1800 x_3 0.8194
Chloroform Recovery		***95.4***		***89.0***

Figure 4. Simulation results and experimental validation of HEBD.

4. Conclusions

Application of heterogeneous entrainer for separating non-ideal mixtures using azeotropic and extractive batch distillation demonstrated important advantages having a straight effect over the economy of the process: more potential candidates reducing the time of the preliminary process synthesis, less amount of entrainer is usually required, heterogeneous entrainer neutralizes the effect of the distillation boundaries of the residue curve map giving a bigger feasible region and allowing the separation of all components in one single operating task. For the separation of aqueous-organic mixtures, many feasible entrainers could be found among those inducing a liquid-liquid phase split with water and vice versa, water seems to be a promising candidate for separating organic wastes. Possibility of the use of non pollutant entrainers is increased.

References

Gerbaud V., X. Joulia, I. Rodriguez-Donis, et al., 2006, Chem. Eng. Proc, 45(8), 672-683.
Gmehling, J., 1994, Azeotropic Data. VCH Editor: Weinheim.
Hua C., L.I. Xingang, X.U. Shimin, B.A.I. Peng, 2007, Chin. J. Chem. Eng., 15, 286-290.
Lang P. and Modla G., 2006, Chem. Eng Sci. 61, 4262 – 4270.
ProSim S.A. http://www.prosim.net
Rodriguez Donis I., V. Gerbaud and X. Joulia, 2002, AIChE J. 1168 -1178
Rodriguez Donis I., V. Gerbaud and X. Joulia, 2001, ESCAPE'11 Proceedings. 499-505.
Rodríguez-Donis I, J. Acosta-Esquijarosa, V. Gerbaud et X. Joulia, 2003, AIChE J. 3074 - 3083
Rodriguez Donis I., V. Gerbaud and X. Joulia, 2009, Ind. Eng. Chem. Res., 48(7), 3544–3559.
Rodriguez Donis I., V. Gerbaud and X. Joulia, 2009, Ind. Eng. Chem. Res., 48(7), 3560–3572.
Skouras S. V. Kiva and S. Skogestad, 2005, Chem. Eng Sci. 60, 2895-2909
Van Kaam R., I. Rodriguez-Donis, V. Gerbaud, 2008, Chem. Eng. Sci. 78 – 94.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

In Situ Incorporation of Recycled Polymer in Suspension Polymerizations

Caio K. Melo, Matheus Soares, Príamo A. Melo and José Carlos Pinto

Programa de Engenharia Química / COPPE – Universidade Federal do Rio de Janeiro Cidade Universitária, CP: 68502, Rio de Janeiro, 21941-914 RJ, Brazil.

Abstract

Polymer materials find widespread use because of their characteristic low costs and wide range of end-use properties. However, the continuous accumulation of polymer materials in the environment constitutes a significant environmental problem. In order to promote the reutilization of used polymer materials, it is shown in this work that the *in situ* incorporation of recycled polystyrene into virgin polystyrene beads can be performed successfully during suspension polymerizations. The influence of increasing recycled polystyrene contents on the course of the polymerization is investigated. It is shown that the recycled polystyrene content does not affect the kinetics of the polymerization significantly, indicating that operation policies can be adapted to increasing amounts of recycled material with the help of common suspension polymerization models.

Keywords: Recycling, polymer, polystyrene, styrene, suspension polymerization.

1. Introduction

Polymer materials find widespread use because of their characteristic low costs and wide range of end-use properties. However, the continuous accumulation of polymer materials in the environment constitutes a significant environmental problem. However, polymer wastes should be regarded as raw materials for the manufacture of new products through recycling processes. Several techniques can be applied to reintroduce recycled products into the market with competitive performance, such as mechanical recycling, chemical recycling and energy production. Mechanical recycling usually requires the combination of high temperatures and shear stresses (energy consumption) in order to mix used polymers with other materials. Typical operation conditions can cause the thermo-oxidation and the degradation of the polymer product, hindering the final properties of the material (Luzuriaga, 2005). Chemical recycling usually requires depolymerization of the recycled material through solvolysis (e.g. hydrolysis and alcoholysis) or themal-catalytic methods (e.g. pyrolysis). In both cases, the obtained product is normally a complex mixture of monomer, oligomers and residues, requiring complex posterior purification. Energy production is usually achieved through combustion, producing toxic gases that must be removed from the gas exhaust before its final release in the atmosphere (Spinace, 2005).

Polystyrene is among the most popular polymer materials, used mainly for packing and production of disposable materials because of its low cost and good processability. However, final pieces are produced in large scale and most times present short lifecycle. For this reason, the waste management of polystyrene pieces constitutes an important environmental problem. Waste management procedures should certainly consider the

development of new technologies for waste reuse and recycle, instead of enhancement of polymer degradation in landfills (Vilaplana, 2008).
In order to promote the reutilization of used polymer materials, it is shown in this work that the *in situ* incorporation of recycled polystyrene into virgin polystyrene beads can be performed successfully during suspension polymerizations. The influence of increasing recycled polystyrene contents on the course of the polymerization is investigated. It is shown that the recycled polystyrene content does not affect the kinetics of the polymerization significantly, indicating that operation policies can be adapted to increasing amounts of recycled material with the help of common suspension polymerization models

2. In situ incorporation of additives during suspension polymerizations

It is known that additives can be incorporated *in situ* during suspension polymerizations in order to modify the final properties of the final material and develop specific applications. For example, x-ray contrast has been incorporated *in situ* into poly(methyl methacrylate) resins used for production of bone cements (Lemos, 2006). High impact polystyrene (HIPS) can also be produced through *in situ* incorporation of polybutadiene during the styrene suspension polymerization (Casis, 2005). The incorporation of additives during the polymerization process is advantageous, because improved homogeneity (and, consequently, enhanced mechanical properties) can be obtained. Despite that, recycling of used polymer materials through suspension polymerizations have seldomly been reported.

3. Experiments

3.1 Chemicals and Experimental Procedures

Styrene was provided by Nitriflex Resinas S/A, with a minimum purity of 99%. The suspending agent [poly(vinyl alcohol), PVA] and the initiator (benzoyl peroxide, BPO) with a minimum purity of 75% and moisture content of 25%, was supplied by Vetec Química Fina. Hydroquinone was supplied by Vetec Química Fina and used as a polymerization inhibitor to interrupt the polymerization of collected samples. Tetrahydrofuran (THF), used for GPC analyses, was supplied by Tedia Brasil with a minimum purity of 99.9%. Recycled polystyrene materials were obtained as clean disposable cups.

Table 1 – Basic polymerization recipe.

Component	*Mass*
Initiator BPO	4g
Stabilizer PVA	2.2g
Styrene	100g
Water	400g

Suspension polymerizations were carried out in a 1-L jacketed glass reactor (FGG Equipamentos Científicos Ltda) at 80 °C, with a total organic hold up of 20 wt %, unless stated otherwise. Initially, the reactor was fed with distilled water, containing the specified amount of suspending agent (PVA). When the desired temperature was

reached, the reagents and the initiator (BPO) were added. The system was kept under isothermal conditions with a constant agitation of 1000 rpm. The experimental setup is presented in Figure 1. The basic polymerization recipe is presented in Table 1. The specified amount of recycled polymer was dissolved in fresh monomer before monomer feeding.

The reactor was equipped with a reflux condenser, which was fed by a refrigeration bath (3). A heating bath (2) was used to keep the temperature of the reaction medium at the desired setpoint value. The mechanical agitator (5) was equipped with a helix-type impeller. The reactor temperature was monitored in-line with the help of a J-type thermocouple connected to a microcomputer (4).

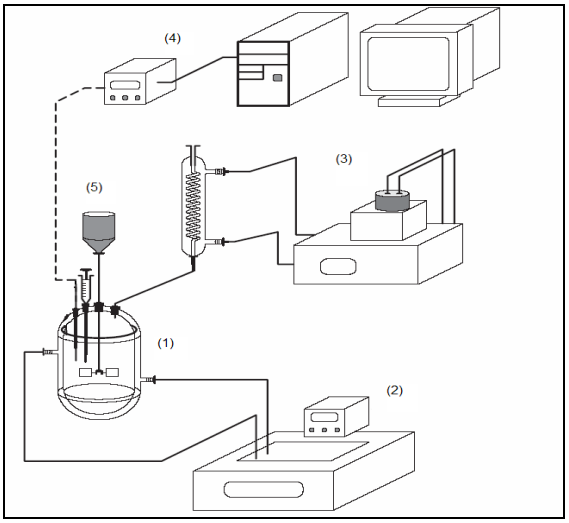


Figure 1 - Experimental Module.

4. Characterization

Samples of approximately 5.0 mL were withdrawn from the system at regular intervals of time for analyses. About 100 ppm of hydroquinone was used to halt the reaction just after the withdrawal of the samples. The total monomer conversion was determined through gravimetry by drying polymer samples until constant weight in a vacuum oven at 40 °C.

The weight-average molecular weight was measured through gel permeation chromatography (GPC). The system was composed of three linear columns with gel porosities ranging from 10^3 to 10^6 Å. A refractometer was used as the detector. The refractometer and the pumping system were connected to a personal computer for data acquisition and data handling. The calibration curve was built using samples of polystyrene with known molecular weight and a polydispersity index smaller than 1.05. Approximately 10-15 mg of dry polymer were dissolved into 2 mL of solvent and filtered through porous membrane filters prior to analysis. After filtration, 200 µL of the polymer solution was injected into the GPC system. Analyses were performed at 35 °C.

5. Model

5.1 Kinetic Model

The kinetic mechanism used to describe the styrene suspension polymerization comprises the usual initiation, propagation, transfer to monomer, inhibition and termination by combination steps, as shown in Table 2. Assuming that the long-chain and quasi-steady-state hypotheses are valid for the polymer radicals, it is possible to develop the following set of mass balance equations for the styrene polymerization process, as shown in Table 3. The well-known method of moments (Ray, 1972) was used for computation of molecular weight data. The kinetic parameters used on the model are presented in Table 4.

Table 2 – Kinetic mechanism.

Initiation by initiator decomposition	Chain transfer to monomer
$I \xrightarrow{K_d} 2R_1$	$R_i + M \xrightarrow{K_{tm}} \Theta_i + R_1$
Spontaneous thermal initiation	Termination by combination
$3M \xrightarrow{K_{th}} R_1(+D)$	$R_m + R_n \xrightarrow{K_{tc}} \Theta_{m+n}$
Monomer propagation	Inhibition
$R_i + M \xrightarrow{K_p} R_{i+1}$	$R_i + H \xrightarrow{K_{tS}} \Theta_i$

Table 3 – Mass balance equations for polymerization.

Initiator Mass Balance

$$\frac{dI}{dt} = -K_d I$$

Inhibitor Mass Balance

$$\frac{dH}{dt} = F_H - K_{tS}\left(\frac{H}{V_O}\right)\left(\frac{m_0}{V_O}\right)V_O$$

Monomer Mass Balance

$$\frac{dM}{dt} = -K_{th}\left(\frac{M}{V_O}\right)^3 V_O - (K_P + K_{tM})\left(\frac{M}{V_O}\right)\left(\frac{m_0}{V_O}\right)V_O$$

Live Chain Mass Balance

$$\frac{dR_i}{dt} = K_P\left(\frac{M}{V_O}\right)\left(\frac{R_{i-1}}{V_O}\right) - K_P\left(\frac{M}{V_O}\right)\left(\frac{R_i}{V_O}\right) - K_{tM}\left(\frac{M}{V_O}\right)\left(\frac{R_i}{V_O}\right) - K_{tS}\left(\frac{H}{V_O}\right)\left(\frac{R_i}{V_O}\right) - K_{tc}\left(\frac{m_0}{V_O}\right)\left(\frac{R_i}{V_O}\right)$$

Dead Chain Mass Balance

$$\frac{d\Theta_i}{dt} = K_{tM}\left(\frac{M}{V_O}\right)\left(\frac{R_i}{V_O}\right)V_O + K_{tS}\left(\frac{H}{V_O}\right)\left(\frac{R_i}{V_O}\right)V_O + \frac{K_{tc}}{2}\sum_{j=1}^{i-1}\left(\frac{R_{i-j}}{V_O}\right)\left(\frac{R_j}{V_O}\right)V_O$$

Moments and Molecular Weights

$$m_k = \sum i^k R_i \quad , \quad x_k = \sum i^k \Theta_i \quad , \quad M_n = \frac{x_1 + m_1}{x_0 + m_0} PM \quad , \quad M_w = \frac{x_2 + m_2}{x_1 + m_1} PM$$

Table 4 – Kinetic parameters for styrene polymerization.

Parameters	References
$k_P = 1.09\text{x}10^7 \exp(-7050/RT)$ *L/(mol.s)*	KALFAS *et al.* (1993)
$k_{tc} = 1.7\ \text{x}10^9 \exp(-2268/RT)$ *L/(mol.s)*	KALFAS *et al.* (1993)
$k_d = 5.7\times10^{14} \exp(-153000/RT)$ s^{-1}	KALFAS *et al.* (1993)
$k_{tM} = 2.31\times10^6 \exp(-12670/\text{RT})$ *L/(mol.s)*	OLIVEIRA *et al.* (1998)
$k_{th} = 2.19x10^5 \exp(-27440/RT)$ $L^2/(mol^2.s)$	ASTEASUAIN *et al.* (2007)

6. Results

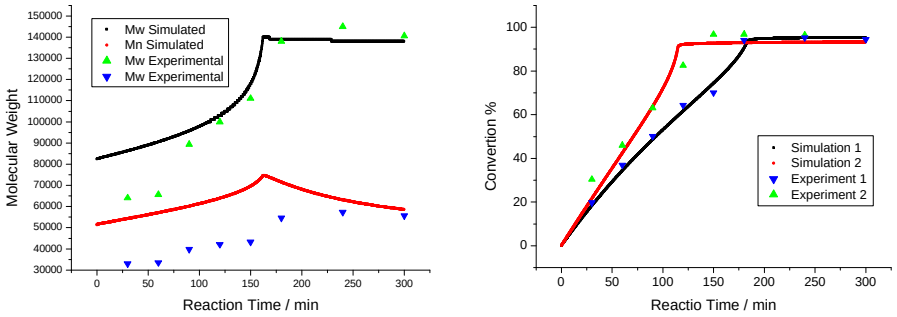


Figure 2 – Validation of model simulation with experimental data.

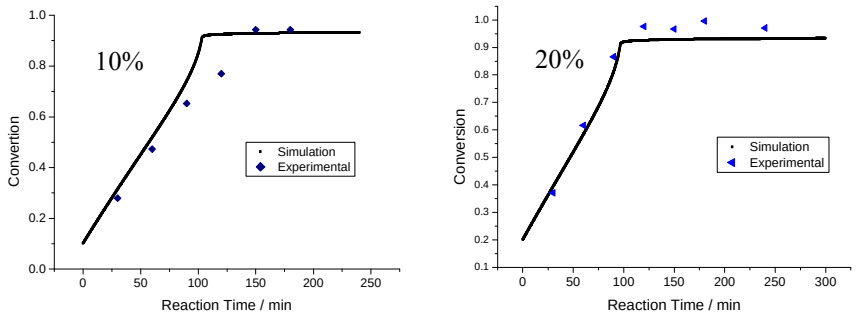


Figure 3 – Simulation and experimental results for reactions containing 10 and 20% of recycled PS.

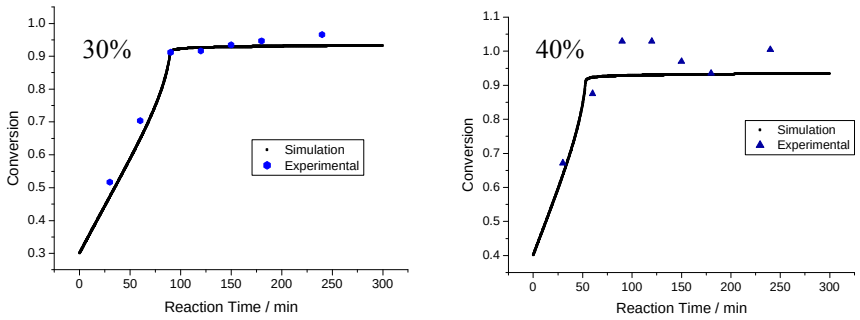


Figure 4 – Simulation and experimental results for reactions containing 30 and 40% of recycled PS.

The model was validated for different polymerization runs. Results can be seen in Figure 2. When recycled polystyrene is incorporated into virgin polystyrene beads, it can be assumed that the initial monomer conversion is equal to the recycled polystyrene content (as a mass fraction). As it can be seen in Figures 2 to 4, the model describes the course of monomer conversion very well. This is very interesting because it shows that the gel-effect does not depend on the polymer quality significantly, as one might expect. (As observed experimentally, the average molecular weights of recycled polymer materials were much larger than the ones produced through suspension polymerizations

and approximately equal to $3.0x10^5$ g/mol.) Additionally, as the recycled polymer material certainly contained unknown additives, it was shown that possible inhibitory effects (Rabelo, 2000) might be disregarded. Despite that, Table 5 shows that the obtained molecular weight averages (M_n and M_w) were smaller than predicted by the model (M_n^{cal} and M_w^{cal}), indicating that additives might promote chain transfer reactions. (Results not shown here for lack of space confirmed this hypothesis.)

Table 5 – Comparison between experimental and calculated average molecular weights.

	M_n	M_w	M_n^{cal}	M_w^{cal}
Experiment - 10% PS incorporated	11347	23753	20708	82637
Experiment - 20% PS incorporated	15278	67481	19551	102923
Experiment - 30% PS incorporated	13731	85439	19392	109153
Experiment - 40% PS incorporated	11550	93153	18855	123479

7. Conclusion

The influence of increasing recycled polystyrene contents on the course of suspension styrene polymerizations was investigated. It was shown that the recycled polystyrene content does not affect the kinetics of the polymerization significantly, indicating that operation policies can be adapted to increasing amounts of recycled material with the help of common suspension polymerization models. It was also shown that unknown additives of recycled materials do not cause any significant inhibitory effects, but can promote chain transfer reactions.

References

ASTEASUAIN, M., Soares, M., Lenzi, M. K., et al., Living Free Radical Polymerization in Tubular Reactors. I. Modeling of the Complete Molecular Weight Distribution Using Probability Generation Functions. Macromolecular Reaction Engineering, v. 1, pp. 622-634, 2007.

CASIS, N., Estenoz, D., Gugliotta, L., Oliva, H. Meira, G. Heterogeneous Bulk Polymerization of Styrene in the Presence of Polybutadiene: Calculation of the Macromolecular Structure. Journal of Appl. Pol. Scien., v. 99, p. 3023-3039, 2006

KALFAS, G., Yuan, H., Ray, W. H., "Modeling and Experimental Studies of Aqueous Suspension Polymerization Processes. 2. Experiments in Batch Reactors". Industrial Engineering Chemistry Research, v. 32, pp.1831-1838,1993.

LEMOS, L.; Nele, M.; Melo, P. A.; Pinto, J. C.; Modeling of Bone Cement Production, Macromol. Symp., 243, 1, 13-23, 2006.

LUZURIAGA, S.; Kovárová, J.; Fortelny, I.; Degradation of pre-aged polymers exposed to simulated recycling: Properties and thermal stability. Polymer Degradation and Stability, v. 91, p. 1226-1232, 2006.

OLIVEIRA, A. T. M., Biscaia, E. C., Pinto, J. C., "Optimization of Batch Solution Polymerizations: Simulation Studies Using an Inhibitor and a Chain-Transfer Agent". Journal of Applied Polymer Science, v. 69, pp. 1137-1152, 1998.

RABELO, M.; Aditivação de Polímeros. 1 ed. São Paulo, Artliber, 2000.

RAY, W. H.; On the Mathematical Modeling of Polymerization Reactors. Journal of Macromol. Scien. – Reviews in Macromol. Chem., C08(01), p. 01-56, 1972.

SPINACÉ, M. A. S.; Paoli, M. A.; A Tecnologia da Reciclagem de Polímeros. Química Nova, v. 28, p. 65-72, 2005

VILAPLANA, Francisco; Amparo Ribes-Greus; Karlsson, Sigbritt. Polymer Degradation and Stability, v. 91, n. 2, p. 2163-2170 (2006).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Synthesis and Design of Combined Biological Nitrogen and Phosphorus Removal WWT Plants

Noelia Alasino, Miguel C. Mussati, Nicolás Scenna, Pío Aguirre

INGAR Instituto de Desarrollo y Diseño (CONICET-UTN), Avellaneda 3657, Santa Fe(S3002GJC), Argentina.
E-mail address: (nalasino; mmussati; nscenna; paguir)@santafe-conicet.gov.ar

Abstract

In the present work, a previous superstructure model developed for simultaneous optimization of the process configuration and equipment dimensions- i.e., process synthesis and design- and the operation conditions of activated sludge wastewater treatment plants (Alasino et al., 2007) will be extended to account for phosphorus as well as nitrogen removal. Continuous operation is supposed, and the influent wastewater flowrate and composition are assumed known. The performance criterion selected is to minimize the net present value including investment and operating costs while verifying compliance with the effluent permitted limits.
The Activated Sludge Model No. 3 extended with the Bio-P module for computing biological phosphorus removal is used to model the reaction compartments, and the Takàcs model for representing the secondary settler. In the present work, new variables and equations for components accounting for phosphorus removal processes are incorporated and the superstructure is also enlarged to embed the most widely used combined N and P removal process configurations. Therefore, the complexity of the problem has been increased. The problem is posed as a NLP problem, specifically a nonlinear programming problem with discontinuous derivatives -DNLP-, as it results in a highly nonlinear system with non-smooth functions. The model is implemented and solved using General Algebraic Modelling System (GAMS), and proved to be robust and flexible. Results for case studies are presented and discussed.

Keywords: activated sludge process, combined N and P removal, optimal synthesis and design, ASM3+Bio-P.

1. Introduction

The activated sludge (AS) system is the most used biological process for municipal wastewater treatment (WWT) plants. Along the activated sludge treatment process, the wastewater stream can be exposed to different environmental conditions (anaerobic, anoxic, and aerated zones) in order to facilitate the different microbiological processes such as the release or uptake of phosphorus and the nitrification/denitrification processes.
In Alasino et al. (2007) the simultaneous optimization of the process configuration and equipment dimensions -i.e. process synthesis and design- and the operation conditions of activated sludge wastewater treatment plants for nitrogen (N) removal based on a superstructure model was presented. That work was one of the first for synthesis and design based on a superstructure model for the wastewater treatment process, using rigorous models for biochemical and settling processes, and aiming at costs minimization. Then, in Alasino et al. (2008), the optimal operation conditions resulting

from a superstructure model that embeds the most widely used configurations for combined N and phosphorus (P) removal, aiming at minimizing operating annual costs was investigated for given wastewater specifications. In that work the plant equipment was supposed given, and no investments costs were computed. The plant lay-out used as the departing model was that proposed by Gernaey and Jorgensen (2004), which corresponds to the A2/O process. The other configurations embedded were the UCT process (VIP process), the modified UCT process and the Bardenpho process.
In the present work, a superstructure model for simultaneous optimization of both the process configuration and equipment dimensions- i.e., process synthesis and design- and the operation conditions of activated sludge wastewater treatment plants in continuous operation, and for given wastewater specification and flow rate, is developed to account for P as well as for N removal. Here, the model embeds up to seven reaction compartments and a secondary settler, and allows for flow distribution of the main process streams (i.e. nitrate recycle, sludge recycle and fresh liquid), bypasses and external carbon source dosage along the reaction zone. The objective function is to minimize the Net Present Value -NPV- considering investment and operating costs. The compartments volumes and the operating conditions are to be optimized, while the decanter (with fixed dimensions) is supposed given. Each compartment will operate in aerobic, anoxic, or anaerobic conditions according to the aeration flowrate computed and the streams fed to it. As explained in the previous work (Alasino et al., 2007), all the decision variables are modeled as continuous, and some small lower bounds were used when necessary to avoid numerical problems. When a continuous variable (e.g. reaction compartment volume and flowrate of aeration, fresh wastewater, recycles, bypasses, and external carbon source to each reaction compartment) take a zero value at a solution point, the corresponding unit and/or stream is removed from the superstructure, meaning that it is not part of the final process configuration (flowsheet), and not in the sense that the mathematical representation of this unit and/or stream is eliminated in the mathematical model. In future works, new mathematical models based on MINLP or Generalized Disjunctive Programming –GDP– methodologies will be presented.

2. Problem Definition

The problem addressed is the simultaneous optimization of the process configuration (process synthesis), equipment dimensions (design) and the operating conditions (flow rates of aeration, recycles and fresh feed to each reaction compartment and external carbon source dosage) of ASWWTPs for combined biological N and P removal, aiming at minimizing the net present value. The influent wastewater specifications, effluent permitted limits, and the cost model computing operation and investment costs, are assumed given.

3. Process Description

In ASPs, the WW stream is exposed to different environmental conditions (anaerobic, anoxic and aerated zones) to facilitate the different microbiological processes such as the release or uptake of P, nitrification and denitrification. Reduction of carbonaceous matter and nitrification (ammonium is converted to nitrate by autotrophs) are favored by aerobic conditions; while denitrification (nitrate is converted to N gas by heterotrophs) is favored by anoxic ones, if readily biodegradable C is available. Biological P removal relies on P uptake by aerobic heterotrophs (known as phosphate-accumulating organisms PAOs) capable of storing orthophosphate in excess of their biological growth

requirements. Under anaerobic conditions, PAOs convert readily available C (e.g., VFAs) to C compounds called polyhydroxyalkanoates PHAs. PAOs use energy generated through the breakdown of polyphosphate molecules to create PHAs. This breakdown results in P release. Under subsequent aerobic or anoxic conditions, PAOs use the stored PHAs as energy to take up the P that was released in the anaerobic zone, as well as any additional phosphate present in the WW.

4. Process Optimization Model

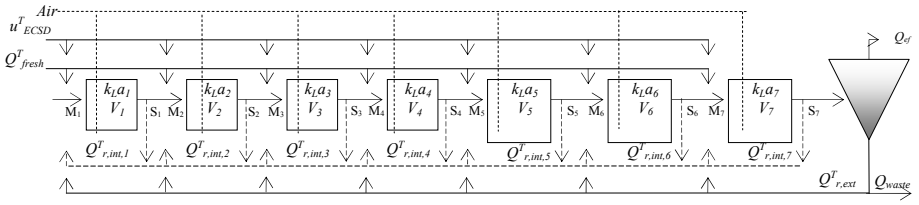


Figure 1. Representation of the WWTP Superstructure.

The considered process superstructure is shown in Fig. 1. The superstructure embeds the more widely used process alternatives for combined N and P removal. More specifically, the model embeds a chain of up to seven reaction compartments followed by a secondary settler (the latter with fixed dimensions), and allows for flow distribution of the main process streams i.e. nitrate and sludge recycle, bypasses and fresh feed streams, and external carbon source dosage along the reaction zone. The compartments volumes are to be optimized. Each compartment operates in aerobic, anoxic, or anaerobic conditions according to the aeration flow rate computed and the streams fed to it.

4.1. *Reactor model*

For the aeration tanks, steady state CSTR model is considered. The ASM3 model (Gujer et al, 1999) extended with the Bio-P module (Rieger et al., 2001) is chosen to model the biological processes. The stoichiometric coefficients and kinetic constants are interpolated to 15 ^oC as proposed by Alasino et al. (2008). The volumes of the reaction compartments and their k_La values are to be optimized using upper and (very small) lower bounds as proposed in Alasino et al. (2007).

4.2. Secondary settler model

The secondary settler is modeled as a non-reactive settling tank subdivided into 10 layers of equal thickness, using the double-exponential settling velocity model (Takács et al., 1991). A fixed settler depth of 4 m and a cross area of 1500m^2 are adopted.

4.3. Splitter and mixer mass balances.

Splitters and mixers models are also needed to represent the proposed superstructure.

4.4. Effluent quality limits

These effluent thresholds values were used as specification constraints (Copp, 2002; Gernaey and Jorgensen, 2004): $S_{NH,ef}$: 4 gN m^{-3}; $P_{tot\ ef}$: 1.5 gP m^{-3}; $N_{TOT,ef}$: 8 gN m^{-3}; BOD_{ef}: 10 gCOD m^{-3}; COD_{ef}: 100 gCOD m^{-3}; $X_{SS,ef}$: 30 gSS m^{-3}.

4.5. Maximum values for operation variables

The maximum values for the operation variables taken from Copp (2002) are: $Q_{r,ext}$: 36892m^3d^{-1}; $Q_{r,int}$: 92230m^3d^{-1}; Q_{waste}:1844.6m^3d^{-1}; u_{ECSD}: 2*10^3 kgCODd^{-1}; k_La_i:360 d^{-1}.

4.6. Objective Function

Here, the Net Present Value (NPV) is adopted as the objective function to be minimized, computing the investment (IC) and operating (OC) costs, as in Alasino et al. (2007).

The investment cost includes reaction tanks, aeration systems, secondary settler, influent pumping station, and sludge pump costs. The operation cost computes the cost for pumping and aeration, dosage of an external carbon source, excess sludge treatment for disposal, and fines according to contaminating units discharged.

The investment cost of the plant is computed as follows (Alasino et al., 2007):

$$IC = IC_t + IC_a + IC_{sett} + IC_{ips} + IC_{sr}$$

The investment cost functions used IC_p have the basic structure $IC_Z = b_Z Z^{\alpha_Z}$, where b and α are parameters and Z is an equipment characteristic dimension; for instance, the volume V_i of each reaction compartment R_i; the Oxygen Capacities $OxCa_i$ ($OxCa_i = B\ k_L a_i\ V_i$) for the aeration system. and the cross area $A_{settler}$ for the secondary settler. For this case, two different investment costs are computed: cost for tank construction and electromechanical system. Finally, the influent pumping station (concrete, screws and screening) and the sludge pump cost are considered. For these items, the characteristic dimension are the influent wastewater flow rate (Q_{FreshT}) and the recycled sludge flow rate (Q_{rext}).

The annual operating cost is computed as follows (Vanrolleghem and Gillot, 2002):

$$OC^{T,Annual} = OC_a + OC_{pump} + OC_{EQ} + OC_{SLDGD} + OC_{ECSD}$$

where OC_p is the annual operating cost of unit p. EQ, E_a, E_{pump}, u_{SLDGD} and u_{ECSD} are the effluent quality index, aeration energy demand, pumping energy demand, waste sludge production rate, and external carbon source dosage rate, respectively, whose expressions are given in detail in Alasino et al. (2007) and Gernaey and Jorgensen (2004). The annual unitary operation costs are (Vanrolleghem and Gillot, 2002; Mussati et al., 2002; Gernaey and Jorgensen 2004): α_{EQ}: 50 Euro day (kgPU year)$^{-1}$; α_E: 25 Euro day kWh year)$^{-1}$; α_{SLDGD}: 75 Euro day (kgS_S year)$^{-1}$; α_{ECSD}: 109.5 Euro day (kgCOD year)$^{-1}$. The effluent quality index EQ (kg contaminating unit d^{-1}), which is related to the fines paid due to contaminant discharge, is computed by weighting the compounds loads having influence on the water quality that are usually included in the legislation. Finally, the total operating cost is computed as follows:

$$OC = \Gamma\ OC^{T,Annual}$$

where Γ is used to update costs to the present value. The updating term is:

$$\Gamma = \sum_{j=1}^{n} \frac{1}{(1+i_d)^j} = \frac{1-(1+i_d)^{-n}}{i_d}$$

where i_d is the interest rate (discount rate) and n is the life span of the WWTP. A discount rate (i_d) of 0.05 and a life span (n) of 20 years are used.

4.7. Influent Wastewater Specifications

The influent WW flow rate is set at 18446 m^3 d^{-1}. The influent WW composition used consists of the original flow-weighted average dry weather influent composition for ASM1 proposed in COST, modified to make it compatible with the ASM3+BioP model. The influent $PO4^{=}$ concentration (S_{PO}) has been taken from Gernaey and Jorgensen (2004). The nonzero input concentrations for compounds are: S_I: 30 gCOD

m^{-3}; S_S: 69.5 gCOD m^{-3}; X_I: 51.2 gCOD m^{-3}; X_S: 202.32 gCOD m^{-3}; X_H: 28.17 gCOD m^{-3}; X_{SS}: 215.493 gSS m^{-3}; S_{NH}: 40.60 gN m^{-3}; S_{ALK}: 7 gCOD m^{-3}; S_{PO4}: 9.01 gP m^{-3}.

5. Results and Discussion

A multiple starting point strategy was adopted for model initialization, and several locally optimal solutions were found depending on the initial value set as in the previous work (Alasino et al., 2007). The WWTP configuration showing the minimal NPV value is represented in Figure 2. Table 1 shows the effluent concentrations and costs.

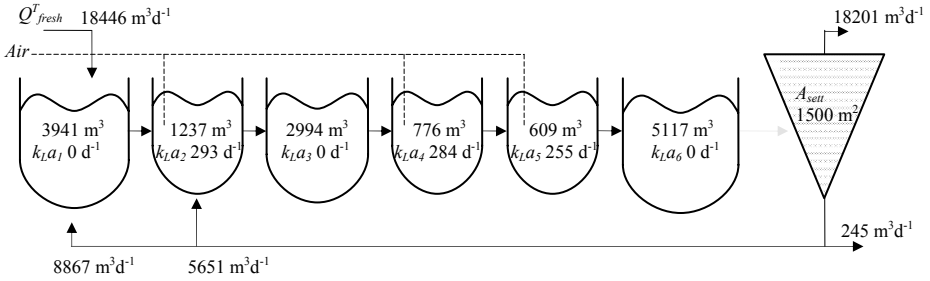


Figure 2. Optimal configuration and main process variable values.

Table 1. Main variables optimal values and costs.

Effluent Contaminant	**Value**	**Effluent Contaminant**	**Value**
$S_{NH,ef}$, g N m^{-3}	3.34	BOD_{ef}, g m^{-3}	1.67
$N_{TOT,ef}$, g N m^{-3}	5.37	COD_{ef}, g CODm^{-3}	45.00
$P_{TOT,ef}$, g P m^{-3}	1.50	$X_{SS,ef}$, g SS m^{-3}	15.05
Costs (Euros)			
OC_{EQ}	196397	IC_t	2382674
OC_{pump}	14763	IC_a	181764
OC_a	156959	IC_{set}	442671
OC_{SLDGD}	193860	IC_{ips}	268985
OC_{ECSD}	0	IC_{sr}	35310
$OC^{T,Annual}$	***561979.8***		
OC	***7002268***	***IC***	***3311404***
NPV	**10313672.3**		

As shown in Fig. 1, the optimization resulted in six reaction compartments with volumes of 3941, 1237, 2994, 776, 609 and 5117 m^3, respectively. That is, the optimal configuration includes six of the seven available reaction compartments. The volumen of the seventh (last) compartment achieves the lower bound, which is sufficiently small and insignificant from a practical point of view; so, that compartment is eliminated and is not represented in the figure. The second, fourth and fifth compartment are aerated;

thus, the WWTP consists of a sequence of compartments with the following characteristics: Ana-Ae-Anox-Ae-Ae-Anox. The optimal solution considers external recycle distribution: around 61% of the external recycle flowrate (8867 m^3 d^{-1}) is fed to the first compartment and the rest to the second one. No feed distribution is necessary and no external carbon source is dosed. As can be appreciated in Table 2.a, only the $P_{TOT,ef}$ concentration achieves the effluent threshold value.

In the previous work for N removal (Alasino et al., 2007), the synthesis and design for the same wastewater specification concentration and flowrate (except for the P content), with a superstructure consisting on 5 reaction compartments, resulted in three reaction compartments with volumes of 1083 m^3 (moderately aerated, k_La: 218 d^{-1}), 9096 m^3 (slightly aerated, k_La: 36 d^{-1}) and 6099 m^3 (slightly aerated, k_La: 27 d^{-1}), respectively, with fresh feed distribution (around 56% of the influent flowrate -10360 m^3 d^{-1}-, fed to the first compartment and the rest to the second one), and an external sludge recycle of 11841 m^3d^{-1} conducted to the first reaction compartment. The total reaction volume (sum of the reaction compartments volume) computed for that case is similar to this, but as P compounds and the organisms for P removal (X_{PAO}) were not considered, only 3 reaction compartments for developing conditions to N removal were needed. When phosphorous removal is targeted, an anaerobic compartment is essential to develop conditions for phosphorous release process, which is a biological step needed for phosphorous removal. As carbonaceous matter is consumed for P as well as for N removal processes, i.e. for P release and uptake process and for the nitrification/denitrification process, there is a shortage of carbonaceous matter and the effluent quality reached is not as good as in the previous work. Taken in conjunction with the fact that the effluent S_{NH} is higher in this case and that the P compounds are here computed, it results in a worse EQ (effluent quality) and in a higher cost due to fines. In addition, the higher amount of sludge for disposal produces higher sludge treatment cost. This results in an increment in the net present value, which is 10313672 Euros against 7473917 Euros in the previous work.

References

N. Alasino, M. C. Mussati, N. Scenna, 2007, Wastewater treatment plant synthesis and design, Ind. Eng. Chem. Res., 46, 23, 7497.

N. Alasino, M. C. Mussati, N. Scenna, P. Aguirre, 2008, Combined nitrogen and phosphorus removal. Model-based process optimization, 18th European Symposium on Computer Aided Process Engineering, Volume 25, 163-168.

J Copp, 2002, The COST Simulation Benchmark: Description and Simulator Manual. Office for Official Publications of the European Community, Luxembourg.

K.V. Gernaey and S.B. Jørgensen, 2004, Benchmarking combined biological phosphorus and nitrogen removal wastewater treatment processes. Contr. Eng. Pract., 12, 357.

W. Gujer, M. Henze, T. Mino and M. van Loosdrecht, 1999, Activated Sludge Model No. 3, Wat. Sci. Techn., 39, 183.

M. Mussati, K. Gernaey, R. Gani and S. Bay Jørgensen, 2002, Performance analysis of a denitrifying wastewater treatment plant. Clean Techn. & Env. Policies, 4, 171.

L. Rieger , G. Koch ,M. Kuhni ,W. Gujer and H. Siegrist, 2001, The EAWAG Bio-P module for activated sludge model No. 3. Wat. Res., 35(16), 3887.

I. Takács, G. Patry, and D. Nolasco, 1991, A Dynamic Model of the Clarification-Thickening Process. Wat. Res., 25, 1263.

P.A. Vanrolleghem and S. Gillot, 2002, Robustness and economic measures as control benchmark performance criteria, Wat. Sci. Techn., 45(4-5), 117.

Acknowledgements. The financial support from CONICET and ANPCyT is acknowledged.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Production of Anhydrous Ethanol by Extractive Distillation of Diluted Alcoholic Solutions with Ionic Liquids

Víctor H. Álvarez[1], Pedro Alijó[2], Diego Serrão[2], Rubens Maciel Filho[1], Martín Aznar[1] and Silvana Mattedi[2]

[1]*School of Chemical Engineering, State University of Campinas (UNICAMP) P.O. Box 6066, 13083-970, Campinas-SP, Brazil*
[2]*Chemical Engineering Department, Polytechnic School, Federal University of Bahia (UFBA), R. Aristides Novis 2, Federação, 40210-630 Salvador-BA, Brazil*

Abstract

In this work, the use of the ionic liquid 1-methyl-3-methylimidazolium dimethylphosphate or ethylene glycol as an separation agent was evaluated in a two-column or three-column configuration extractive distillation for the production of anhydrous ethanol (99.8% weight, 99.5% mol) from diluted solutions (11% weight). The extractive distillation process was simulated and optimized with the HYSYS® software. The vapor liquid equilibrium for the system was calculated by the NRTL thermodynamic model, whose interactions parameters were previously published. Experimental design was used in combination with modeling and simulation to determine the operational conditions that minimize reboiler duty. Simulation results showed that both separating agents present similar reboiler duty to produce high-purity ethanol.

Keywords: anhydrous ethanol, ionic liquids, extractive distillation, simulation.

1. Introduction

The growing demand for anhydrous ethanol and the preservation of the environment need a safe, inexpensive and green process. Many industrial techniques have been proposed to separate of ethanol/water azeotropic mixture, such as crystallization, liquid-liquid extraction, adsorption and azeotropic or extractive distillation (Huang et al. 2008). We have focused our attention on the extractive distillation. Extractive distillation is a partial vaporization process, in the presence of a non-volatile and high boiling point compound, usually called solvent or separating agent, which is added to prevent the azeotrope formation (Perry et al., 1992). In most commercial extractive distillation process, there are three-column or two-column configurations to obtain pure ethanol from a dilute aqueous ethanol solution. The separating agent is fed in the upper part of the extractive column, above the feed stream (Pinto et al., 2000; Ligero and Ravagnani, 2003). In the three-column configuration, the concentrator and stripper are in different columns. In the two-column configuration, both columns could be joined together and the last column is used to recuperate the separation agent, recycling it to the extractive column. Usually, the separating agent is the benzene, but this compound is toxic and shows operating unstable control. Another solvent, as reported by Meirelles et al. (1992), is the ethylene glycol (EG). On other hand, ionic liquids have emerged as possible "green" solvents because they are not volatile and, therefore, do not have

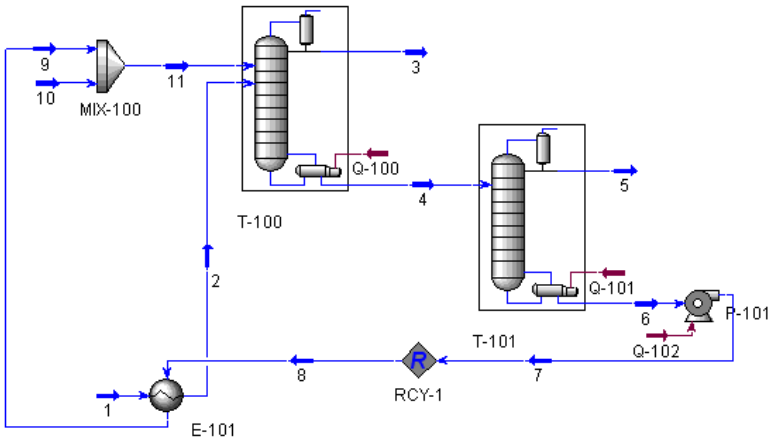
pollutant gas emissions. Ionic liquids are organic salts which are liquid at room temperature, with melting points below 373 K (Marsh et al., 2004). Zhao et al. (2006) reported vapor pressure data for the ionic liquids 1-methyl-3-methylimidazolium dimethylphosphate ([mmim][DMP]) in systems containing ethanol/water at several temperatures and compositions.

The main objective of this work is to model and optimize the two-column and three-column configurations using EG or [mmim][DMP] in order to produce 99.8% weight ethanol from an 11% weight ethanol solution. The results are compared in terms of heat duties and impurities contained in the ethanol product stream.

2. Methodology

The flowsheet of the extractive distillation process for modeling and optimization can be seen in Figure 1. This figures show the recycle point for the simulation (RCY-1).

(a)



(b)

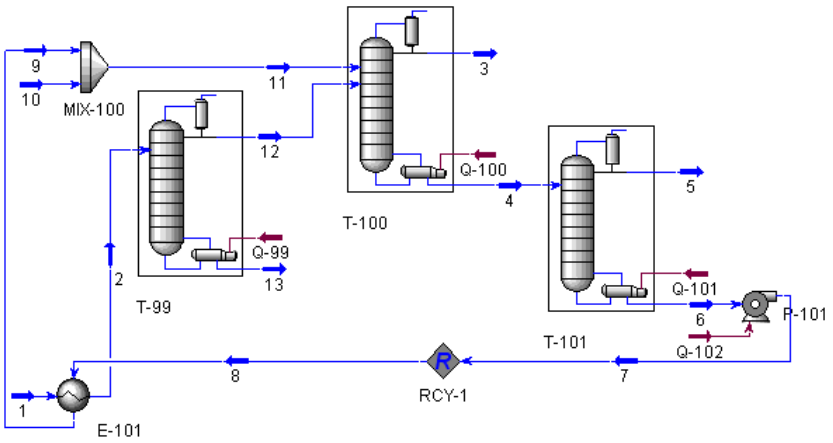


Figure 1. (a) Two-column configuration. (b) Three-column configuration.

2.1. Thermodynamic Model

The ionic liquid was created as a hypothetical component in the HYSYS® software using the liquid density from NIST (2008), pseudo critical properties calculated with the method of Valderrama and Robles (2007) and other properties estimated within HYSYS®. The activity coefficient was described using the NRTL (Renon and Prausnitz, 1968) thermodynamic model. The binary interaction parameters were taken from Zhao et al. (2006) for [mmim][DMP] systems, and from Meirelles et al. (1992) for EG systems. Table 1 displays these NRTL binary interaction parameters. Activity coefficient calculations for the mixture were validated with experimental data reported in previous works (Zhao et al., 2006; Meirelles et al., 1992) using HYSYS® simulator.

Table 1: NRTL binary interaction parameters.

Compound (1) + (2)	α_{12}	Δg_{12o} (J/mol)	Δg_{12t} (J/molK)	Δg_{21o} (J/mol)	Δg_{21t} (J/molK)
water + [mmim][DMP]	0.4116	5065.44	-	-9565.9	-
ethanol + [mmim][DMP]	0.5927	13172.16	-	-6426.83	-
water + EG	0.1859	1383.43	8.0409	-1445.97	-9.1506
ethanol + EG	0.3704	13527.42	-92.7391	-4351.97	53.3769
ethanol + water	0.3008	5612.08	-	-510.81	-

$\Delta g_{12} = \Delta g_{12o} + T\Delta g_{12t}$

2.2. Experimental design

In distillation process with new chemical components, the first step is to know the main operational variables. The second step is selecting the correct interval of work of these main operational variables and, finally their optimization. The software Statistica (Statsoft, v. 7.0) was used to analyze the results for all experimental design.

2.2.1. Finding the main operational variables

Experiments are very time-consuming when many factors have to be considered. In contrast, Plackett–Burman designs (Plackett and Burman, 1946) are very useful for picking up the most important factors from a long list of factors (operational variables) with the less number of experimental trials. In order to perform these steps, the EG separating agent was arbitrarily selected. There were considered as the most important variables those listed in Table 2. Then, the influences of these variables on separation agent loss (X) and energy consumption (Y) was investigated.

2.2.2. Selecting the correct interval of operational variables

In order to select the interval of operational variables, a sensibility analysis was performed, where the column theoretical stage, the separation agent molar flow, and the separation agent feed stage were analyzed. Also, several variables are maintained constants, such as the fermentation broth temperature feed (298 K), the operational pressure of concentrator and extractive column (101.3 kPa), the ethanol molar fraction in the concentrator or extractive column bottom (10^{-6}), the ethanol molar fraction in distillate extractive column (0.995), the water molar fraction in the recovery column bottom (10^{-6}), and the separation agent molar fraction in the distillate recovery column (10^{-6}). Then, the analyzed variables were varied up to obtain constant values of reboiler duty or ethanol molar fraction. The optimal stage for the stream feeds are into the coalescence between the rectification and stripping sections of the distillation column (Walas, 1990). Then, the feed stage was selected as the lowest reboiler duty was obtained for a column.

2.2.3. Optimization of the operational variables:
This step found the operational variables optimal value for the lowest energy consumption (Y). The operational variables selected by Plackett–Burman analysis with limits chosen by the sensibility analysis and the other factor used as constants are the input variables in this optimization. The optimization could be done by either on gradients or complex methods based on local optimization or by direct variation of variables as Taguchi method (Taguchi et al., 1990). In this work, Taguchi method was chosen as it gives a faster, more reliable and more stable solution. Once the interval values of relevant variables were selected, experiments were designed by the Taguchi method with energy consumption as response. The L_9 orthogonal array was selected in the Taguchi method to analyze the variables at different levels.

3. Results

Since the value range of the variables is unknown, a sensibility analysis was performed to estimate it. Thereby, (1) the reboiler duty is maintained constant over 40, 30 and 12 theoretical stages for the concentrator, extractive and recovery columns, respectively, (2) the separation agent molar flow should be over 3 kmol/h to obtain 0.998% w ethanol, (3) lower separation agent molar flow results in a lower reboiler duty, (4) in order to prevent contamination of the distillate, the separation agent feed stage should be 2 and 5 for [mmim][DMP] and EG, respectively, (5) the separating agent temperature feed can be equal to the boiling point of ethanol. In order to know the main operational variables in the process, each operational variable was tested at two levels, as shown in Table 2. The interval values used are in agreement with the previously sensibility analysis.

Table 2. Variables process using ethylene glycol.

N	Variable name	-1	+1
	Two-column configuration		
1	Molar flow of separation agent (kmol/h)	3.00	9.00
2	Number of theoretical stages in extractive column	35.00	55.00
3	Separation agent feed stage	2.00	5.00
4	Pressure in recovery column (kPa)	20.26	101.3
5	Number of theoretical stages in recovery column	15.00	30.00
6	Separation agent temperature (K)	341.44	361.44
	Three-column configuration		
1	Molar flow of separation agent (kmol/h)	3.00	9.00
2	Number of theoretical stages in concentrator column	25.00	50.00
3	Ethanol molar fraction in distillate of concentrator column	0.50	0.90
4	Number of theoretical stages in extractive column	25.00	45.00
5	Separation agent feed stage	2.00	5.00
6	Pressure in recovery column (kPa)	20.26	101.3
7	Number of theoretical stages in recovery column	15.00	30.00
8	Separation agent temperature (K)	341.44	361.44

The Plackett–Burman results are shown as Pareto charts of effects on separation agent lost in distillated (X) and energy consumption (Y) at 95% confidence level. The variables near to the statistical p–value equal 0.05 are significant. The results for the two-column configuration are shown in Figure 2. The energy consumption was mainly affected by the separation agent molar flow (N1), the separation agent feed stage (N3), and the pressure in recovery column (N4), while for solvent loss, the separation agent molar flow (N1), and the separation agent feed stage (N5) are the significant variables.

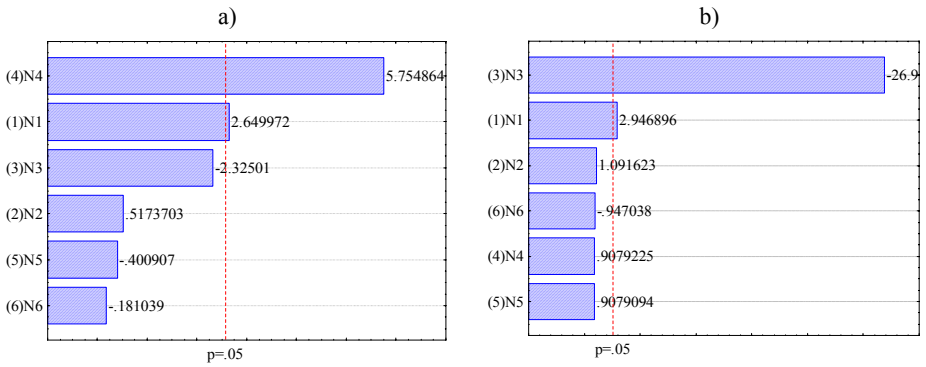


Figure 2. Pareto chart of variables effects at 95% of confidence level for two-column configuration. a) energy consumption, b) agent separation loss.

The results for the three-column configuration are shown in Figure 3. The energy consumption was mainly affected by the separation agent molar flow (N1), the number of theoretical stages in concentrator column (N2), the ethanol molar fraction in distillate concentrator column (N3), the separation agent feed stage (N5), and the separation agent temperature (N8), while for solvent loss, the separation agent molar flow (N1), the number of theoretical stages in concentrator column (N2) and the separation agent feed stage (N5) are the significant variables.

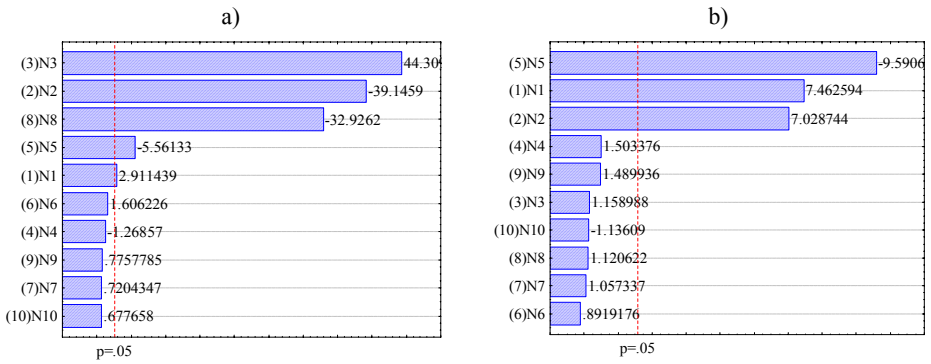


Figure 3. Pareto chart of variables effects at 95% of confidence level for three-column configuration. a) energy consumption, b) agent separation loss.

According the results, the two-column configuration does not need variable optimization, because all variables can be fixed using values from sensibility analysis. The pressure in recovery column and the separation agent molar flow should be the lowest in order to obtain the lowest reboiler duty, 20.26 kPa and 3 kmol/h, respectively. The optimal values for the two-column configuration are displayed in Table 3.

Table 3. Optimum conditions and performance for the two-column configuration.

Separation agent	N1	N2	N3	N4	N5	N6	Energy consumption (kJ/kg ethanol)	Separation agent loss (kmol/h)
[mmim][DMP]	3	40	2	20.26	12	351.44	25819.5	$9 \cdot 10^{-11}$
EG	3	40	5	20.26	18	351.44	25515.5	$1 \cdot 10^{-4}$

The Taguchi method was applied for the three-column configuration variables with significant effect. The factors not marked as statistically important by the Plackett–Burman analysis were fixed as constant (N4=35, N5=5, N6=20.26, N7=18 for EG and N4=35, N5=2, N6=20.26, N7=12 for [mmim][DMP]). Table 4 shows the factors and levels, while the optimal values for operational variables are displayed in Table 5.

Table 4. Factors and their levels for the three-columns configuration.

Level	N1	N2	N3	N8
1	3	25	0.7	341.44
2	6	40	0.8	351.44
3	9	55	0.9	361.44

Table 5. Optimal conditions and performance for the three-column configuration

Separation agent	N1	N2	N3	N8	Energy consumption (kJ/kg ethanol)	Separation agent loss (kmol/h)
[mmim][DMP]	3	55	0.8	361.44	6987.2	$1 \cdot 10^{-6}$
EG	3	55	0.8	361.44	6909.0	$3 \cdot 10^{-5}$

Table 5 shows simulation results obtained using the Taguchi method that, for an efficient use of the energy, the concentrator column should produce an ethanol molar fraction equal to 0.8, a variable not well studied in dehydration ethanol works. The ionic liquid is the separating agent with less contamination of the distillated ethanol.

4. Conclusions

The energy consumption of dehydration of aqueous ethanol by an extractive distillation method was compared for two different configurations and two different separating agents. The calculated results showed that ethylene glycol is more contaminant than [mmim][DMP], but the energy consumption is similar. On the other hand, the three-column configuration process has lower energy consumption than the two-column configuration process.

Acknowledgements

The financial support by FAPESP is gratefully acknowledged. M. Aznar and R. Maciel Filho are recipients of a CNPq fellowship.

References

H.J. Huang, S. Ramaswamy, U.W. Tschirner, B. Ramarao, 2008, Separ. Purif. Tech, 62, 1-21.
E. Ligero, T.M. Ravagnani, 2003, Chem. Eng. and Process, 42, 543-552.
A. Meirelles, S. Weiss, H. Herfurth, 1992, J. Chem. Tech. Biotechnol., 53, 181-188.
K. Marsh, J.A. Boxall, R. Lichtenthaler, 2004, Fluid Phase Equilib., 219, 93-98.
NIST, Ionic Liquids Database-ILThermo, http//ilthermo.boulder.nist.gov/IL (accessed on 2008).
R. Perry, D. Green, J. Maloney, 1997, Perry´s Chemical Engineer's Handbook, 7th ed., McGraw-Hill, New York.
R. Plackett, J.P. Burman, 1946, Biometrika, 33, 305–25.
R. Pinto, M.R. Wolf Maciel, L. Lintomen, 2000, Comput. Chem. Eng., 24, 1689-1694.
H. Renon, J.M. Prausnitz, l968, AIChE J., 14, 135-144.
G. Taguchi, E.A. Elsayed, T.C. Hsiang, 1990, Engenharia da Qualidade em Sistemas de Produção, McGraw-Hill, São Paulo.
J. Valderrama, P.A. Robles, 2007, Ind. Eng. Chem. Res., 46, 1338-1344.
S. Walas, 1990, Chemical Process Equipment. Butterworth-Heinemann, Newton, MA.
J. Zhao, X.C. Jiang, C.X. Li, Z.H. Wang, 2006, Fluid Phase Equilib., 247, 190-198.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Simulation of the azeotropic distillation for anhydrous bioethanol production: study on the formation of a second liquid phase

Tassia L. Junqueira,[a] Marina O. S. Dias,[a] Rubens Maciel Filho,[a] Maria R. W. Maciel,[a] Carlos E. V. Rossell[b]

[a] *School of Chemical Engineering, State University of Campinas – UNICAMP, P.O. Box 6066, Campinas, SP 13083-970, Brazil*
[b] *Interdisciplinary Center for Energy Planning, State University of Campinas, UNICAMP, P.O. Box 6192, 13400-970, Campinas – SP, Brazil*

Abstract

Bioethanol is produced from fermentation of sugars, what produces a dilute solution (around 10 wt% ethanol). Because water and ethanol form an azeotrope with concentration of 95.6 wt% ethanol at 1 atm, an alternative separation process such as azeotropic distillation must be employed to produce anhydrous bioethanol, which can be used in a mixture with gasoline. In this work, simulations of three different configurations of the azeotropic distillation process with cyclohexane for anhydrous bioethanol production were carried out using software Aspen Plus. Process parameters were optimized in order to decrease the formation of a second liquid phase inside the column. Ethanol and entrainer losses as well as energy demand were evaluated.

Keywords: bioethanol, azeotropic distillation, simulation.

1. Introduction

Climate change and the consequent need to diminish greenhouse gases emissions have encouraged the use of bioethanol in a mixture with or as a replacement of gasoline. In order to be used in a mixture with gasoline, ethanol produced from fermentation of sugars, obtained at a concentration of about 10 wt% ethanol, must be concentrated to at least 99.3 %. Since ethanol and water form an azeotrope with 95.6 wt% ethanol at 1 atm, conventional distillation can not achieve the necessary separation that meets product specification. The main processes employed for anhydrous bioethanol production are azeotropic and extractive distillation and adsorption on molecular sieves.
In the conventional configuration of the heterogeneous azeotropic distillation process for anhydrous bioethanol production, two distillation columns are used: the azeotropic one and the recovery unit. Hydrous ethanol and entrainer are fed in the azeotropic column, where anhydrous ethanol is produced on the bottom and a minimum boiling ternary azeotrope formed between ethanol, water and entrainer is produced on the top. The ternary azeotrope is heterogeneous, and the two liquid phases are separated in a decanter. The light phase contains most of the entrainer and is recycled to the azeotropic column, while the aqueous phase is fed to the recovery column where ethanol and eventually entrainer are recovered. The stream containing ethanol recovered may be recycled to the azeotropic column, in order to reduce ethanol losses.
Benzene was the most common entrainer employed in anhydrous bioethanol production, but due to its carcinogenic characteristics cyclohexane is most frequently used

nowadays. The azeotropic distillation process with cyclohexane is the most common method for anhydrous bioethanol production employed in Brazil, which is the second largest ethanol producer in the world. This process is characterized by high steam consumption on column reboilers, as well as high entrainer losses on the product. In other words, the separation step for obtaining anhydrous ethanol has a significant impact on the final product price, so that it is worthwhile to investigate procedures and alternatives to improve the process economy and performance. Simulation is an attractive tool to evaluate the process since it is not suitable to use large scale units to do so.

Simulation of the azeotropic distillation process is often complex and extremely sensitive to project parameters and specifications, mainly because of the formation of a second liquid phase inside the azeotropic column, which influences the mass transfer behaviour: column efficiencies between 25 and 50% are not uncommon when a second liquid phase is present (Higler et al., 2004).

In this work simulations of the azeotropic distillation process with cyclohexane for anhydrous bioethanol production were carried out using software Aspen Plus. Simulation of different process configurations were carried out, and the main parameters evaluated were the formation of the second liquid phase inside the column, the separation of the ternary azeotrope during settling, entrainer losses on the anhydrous bioethanol and energy consumption on column reboilers.

2. Heterogeneous azeotropic distillation

Heterogeneous azeotropic distillation is widely employed in industry as a separation process for azeotropic or close-boiling mixtures, for which conventional distillation can not achieve the necessary separation. In this process, a third component, called entrainer, is added to the binary mixture, producing a heterogeneous azeotrope with one or both components of the original mixture. In the separation of a binary azeotropic mixture, the new azeotrope produces a more favorable azeotropic pattern for the desired separation. Even though the formation of a second liquid phase simplifies entrainer recovery and the recycle process, it may affect column performance and, consequently, decrease efficiency. Different process configurations can lead to a decreased formation of a second liquid phase, without compromising anhydrous bioethanol production.

3. Simulation of the azeotropic distillation process

Three different configurations of the azeotropic distillation process with cyclohexane for anhydrous bioethanol production were simulated using software Aspen Plus. NRTL was the model used for calculation of the activity coefficient on the liquid phase. Approximately 500 m^3/day of anhydrous bioethanol (99.5 wt% ethanol) are produced on each case, from 16,100 kg/h (521 m^3/day) of hydrous bioethanol (93 wt% ethanol).

3.1. Configuration 1

The first configuration of the process is depicted in Figure 1. In this case, hydrous bioethanol is mixed with the recycle stream of the recovery column and fed to the azeotropic column. Entrainer is comprised by the organic phase obtained in the decanter and a solvent make-up stream. Anhydrous bioethanol is produced on the bottom of the azeotropic column, and the ternary azeotrope in the top. After cooling, the ternary azeotrope splits into two liquid phases in the decanter. The aqueous phase is fed to the recovery column, producing pure water on the bottom.

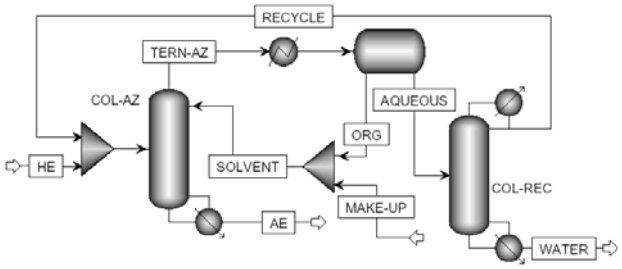


Figure 1. First configuration of the azeotropic distillation process.

3.2. Configuration 2

The main difference between this configuration and the first one is the withdrawn of a fraction of the aqueous phase obtained in the decanter and its mixture with the hydrous ethanol and recycle stream, which are fed to the azeotropic column (Mortaheb and Kosuge, 2004). Configuration 2 of the azeotropic distillation process is displayed in Figure 2.

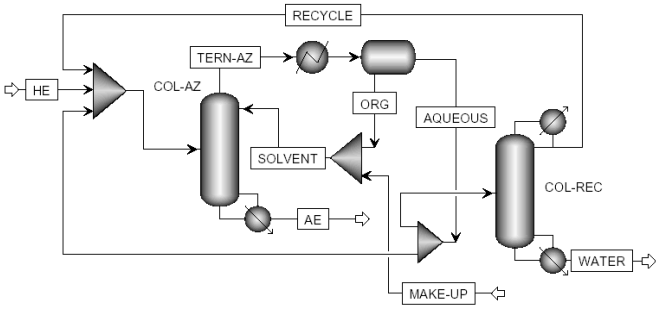


Figure 2. Second configuration of the azeotropic distillation process.

3.3. Configuration 3

In the third configuration part of the ternary azeotrope is condensed and mixed with hydrous ethanol, being subsequently fed to the azeotropic column. In this configuration the top product of the recovery column is not recycled to the azeotropic column, but to the decanter. This configuration is depicted in Figure 3.

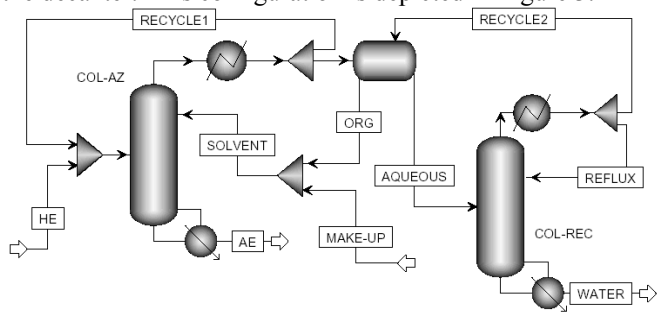


Figure 3. Third configuration of the azeotropic distillation process.

3.4. Simulation parameters

Columns parameters are displayed in Table 1. Stage numbering initiates in the top stage or attached condenser and reboiler is considered the last stage.

Table 1. Process parameters for each configuration.

	Configuration 1	Configuration 2	Configuration 3
COL-AZ number of stages	31	31	31
Feed inlet stage – COL-AZ	18	20	20
Solvent inlet stage	1	1	1
COL-REC number of stages	25	22	17
Feed inlet stage – COL-REC	13	13	13
Decanter temperature (°C)	50	50	50

4. Simulation results and discussion

4.1. Main streams results

The characteristics of process stream differ for each configuration. Results are presented in Tables 2 through 4.

Table 2. Composition of the main streams for the first configuration.

Stream	AE	MAKE-UP	ORG	AQUEOUS	WATER
Temperature (ºC)	78.2	50.0	50.0	50.0	100.0
Flow (kg/h)	15063.6	16.8	42501.6	19524.1	1053.5
Ethanol (wt%)	99.5	0.0	4.7	71.2	0.0
Water (wt%)	0.4	0.0	0.1	10.8	100.0
Cyclohexane (wt%)	0.1	100.0	95.2	18.0	0.0

Table 3. Composition of the main streams for the second configuration.

Stream	AE	MAKE-UP	ORG	AQUEOUS	WATER
Temperature (ºC)	78.2	50.0	50.0	50.0	100.0
Flow (kg/h)	15064.4	21.0	50274.1	22813.3	1057.3
Ethanol (wt%)	99.5	0.0	4.6	71.5	0.0
Water (wt%)	0.4	0.0	0.1	11.3	100.0
Cyclohexane (wt%)	0.1	100.0	95.3	17.2	0.0

Table 4. Composition of the main streams for the third configuration.

Stream	AE	MAKE-UP	ORG	AQUEOUS	WATER
Temperature (ºC)	78.1	50.0	50.0	50.0	78.6
Flow (kg/h)	8395.4	16.8	22877.5	31880.6	7722.0
Ethanol (wt%)	99.5	0.0	5.2	67.6	85.9
Water (wt%)	0.3	0.0	0.1	7.5	14.1
Cyclohexane (wt%)	0.2	100.0	94.7	24.9	0.0

Streams obtained in the decanter (AQUEOUS and ORG) have different compositions on each studied configuration. Configurations 1 and 2 showed similar results, while configuration 3 had a higher concentration of solvent in the aqueous phase. Due to this contamination, recovery column behaviour was significantly affected and "WATER" stream has a high concentration of ethanol (85.9 wt%), so it must be recycled to the previous distillation stage, where hydrous bioethanol is produced.

4.2. Energy consumption and entrainer and ethanol losses

Different process parameters affect ethanol and entrainer losses, as well as energy consumption on column reboilers. The results for these parameters are given in Table 5.

Table 5. Ethanol and entrainer losses and energy consumption on column reboilers for each configuration.

	Configuration 1	Configuration 2	Configuration 3
Ethanol losses (%)*	0.00	0.00	44.26
Entrainer losses (%)**	0.04	0.03	0.07
Reboilers energy (kJ/kg AE)	5588	5828	7235

* Considering the amount of ethanol in the "HE" stream

** Considering the amount of cyclohexane that is fed to the azeotropic column

The "WATER" stream on the third configuration has a very high concentration of ethanol, thus ethanol losses on this configuration are significant. The recycle of this stream to the conventional distillation column would allow ethanol recovery at the expense of an increase in energy consumption of the process, which already has the largest energy consumption on column reboilers. Configurations 1 and 2, on the other hand, present low ethanol and entrainer losses, as well as a relatively low energy consumption.

4.3. Formation of the second liquid phase inside the azeotropic column

Similar liquid flow profiles were observed in the azeotropic column. The profiles for each configuration are represented in Figures 4 through 6.

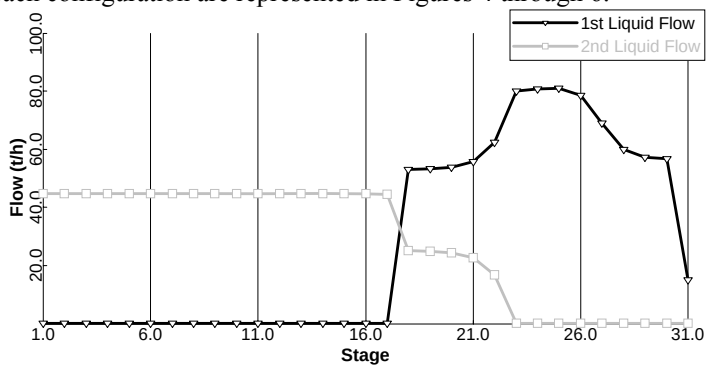


Figure 4. Liquid flow profile in the azeotropic column for configuration 1.

Changes in column parameters, such as feed inlet position in the azeotropic column, greatly influence the formation of a second liquid phase. For the first configuration, the feed inlet stage in the azeotropic column (18) is the one that first presents two liquid phases. The same occurs for the other configurations. The feed inlet stage is the last one that allows column convergence; on the other hand, if the feed stage is located near the

top of the column, two liquid phases are present in this region too. This means that the feed inlet stage in the azeotropic column has to be defined with caution, since it impacts unit behavior and performance.

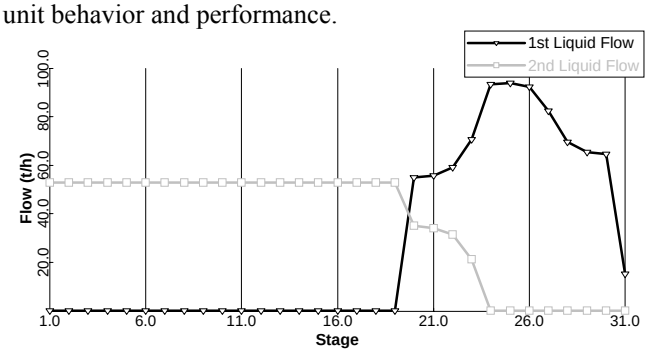


Figure 5. Liquid flow profile in the azeotropic column for configuration 2.

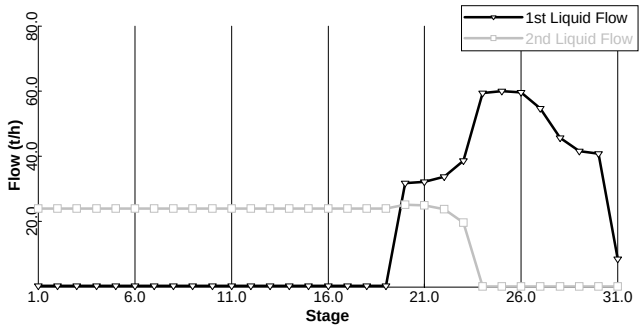


Figure 6. Liquid flow profile in the azeotropic column for configuration 3.

5. Conclusions

Simulations of the azeotropic distillation process with cyclohexane for anhydrous bioethanol production were carried out for three different process configurations. Simulations focus was on decrease of the formation of a second liquid phase inside the azeotropic column, what has significant impacts on process performance. It was observed that changes in process parameters, such as feed inlet stage, greatly influence the formation of the second liquid phase. Configuration 1 presented the best results, considering energy consumption on columns reboilers, ethanol and entrainer losses and product contamination with entrainer.

6. Acknowledgments

The authors acknowledge FAPESP and CNPq for financial support.

References

A. Higler, R. Chande, R. Taylor, R. Baur, R. Krishna, 2004, Nonequilibrium modeling of three-phase distillation, Computers and Chemical Engineering, 28, 2021–2036

H. Mortaheb, H. Kosuge, 2004, Simulation and optimization of heterogeneous azeotropic distillation process with a rate-based model, Chemical Engineering and Processing, 43, 317-326

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Water Source Diagram – An Heuristic Algorithmic Methodology for Reduction of Water Comsumption

Carlos E. P. Siqueira Campos[a], Sara V. Marques[b], Eduardo M. Queiroz[b] and Fernando L. P. Pessoa[b]

[a]*PEQ/COPPE - Federal University of Rio de Janeiro, Av. Athos da Silveira Ramos, 149 - Centro de Tecnologia - Bloco G - Sala 115, Rio de Janeiro, CEP: 21949-900, Brasil*
[b]*EQ/UFRJ - Federal University of Rio de Janeiro, Av. Athos da Silveira Ramos, 149 - Centro de Tecnologia - Bloco E - Sala 207, Rio de Janeiro, CEP: 21949-909, Brasil*

Abstract

Water is a vital natural resource for human life and to the maintaining of natural cycles, but its scarceness is already a reality in some countries. On the other hand, the disposal of wastewater has also been an apprehension, because environmental laws are becoming increasingly rigorous. As a consequence, in case of chemical industries, they are developing campaigns to reduce this consumption, in which the goal is to minimize the use of industrial water and disposal in many segments such as pulp and paper, food (citrus juice), textile, petroleum refinery and petrochemical. So, in order to promote and implement this reduction a method called Water Source Diagram (WSD) is presented in this work. This is an heuristic algorithmic methodology, of simple application, which takes into account the plant restrictions and doesn't make major changes in the existing process, besides it allows manual calculations. The procedure is applied under the same conditions as other existing technologies and also to propose other scenarios not previously discussed. Several flowcharts were developed using WSD procedure which has proved efficiency in the reduction of consumption of this natural resource. In the pulp and paper industry, it was achieved a reduction of water around 89.5% for maximum reuse and 100% for regeneration with reuse, resulting an economy in the total annual costs of 80% (maximum reuse) and 91.6% (regeneration with reuse). In the citric juice and textile industries, the reduction reached 24.5% and 17.9% for maximum reuse respectively, and for regeneration with reuse, the reduction was 31.4% and 50.8% (with reuse and recycle), respectively. In the petrochemical industry, for a simple contaminant, the reduction was 15.7% and 20.4% for maximum reuse and regeneration with reuse, respectively; for multiple contaminants the reduction was 22.1%. Thus, these results show that the presented methodology has an extreme importance in all studied cases, dealing with good signs of economic viability.

Keywords: water source diagram, algorithm, reuse, regeneration.

1. Introduction

Water is a natural resource vital for the maintenance of life on the Planet and to the functioning of other natural cycles, but its scarcity is already a reality in some countries. Industrial segments such as pulp and paper, food (citrus juice), textile, petroleum refinery and petrochemical consume large volume of water and one of its latest trends is the water and wastewater minimization. Several methodologies have been developed to propose mass exchange networks focusing in water and wastewater reuse. Since reuse

causes problems of accumulation of non process elements, the use of regeneration processes becomes necessary in order to reduce contaminants concentration.
Wang and Smith (1994) used the limiting water profile to minimize the water flow rate in systems with single and multiple contaminants, considering only reuse and reuse with regeneration of effluents. In order to overcome the problems concerned with the method proposed by Wang and Smith (1995), Castro *et al*. (1999) developed a procedure in which the water consumption and the synthesis of mass exchange network were respected simultaneously. This procedure was used in systems with single contaminants and considered only the reuse of effluent streams to reduce fresh water consumption. The Water Source Diagram Procedure presented by Gomes *et al.* (2007) and complemented in this work can be considered as an improvement of Castro *et al*. (1999) and Gomez *et al*. (2000) procedures. It is able to take into account a variety of situations, such as (i) reuse, (ii) multiple water sources, (iii) water losses along the process, (iv) flow rate constraints, (v) regeneration with reuse and (vi) regeneration and recycling. Mirre (2007) was the first one to apply this methodology in an industrial case, in which this procedure was used to define the target of minimum consumption of water for an oil refinery process. The study involved the simultaneous consideration of three contaminants in the aqueous streams, where different scenarios were generated using reuse and/or recycle, including regeneration. A scenario was proposed to reuse every effluent stream as a make-up in a cooling tower after its treatment, but with the quality not allowed to disposal directly into the river. Other scenarios considered the presence and absence of an osmosis system to treat the effluent, which led an elevated economic cost to the process. In addition, it was included the scenario involving the possibility of making the effluent distributed treatment, presenting the best cost reduction. The results showed that a reduction about 20% in total costs was achieved. Different studies involving WSD were developed by the Chemicals Processes Integration Group (GIPQ/EQ/UFRJ), but were applied in examples of literature, so remaining a hole in the WSD application in other segments. In this work the Water Source Diagram procedure (WSD) is applied to some industrial processes which have a high consumption of water to generate mass exchange networks with minimum water consumption through reuse, regeneration with reuse and regeneration and recycle. In order to explain the method, one case-segment is completely described and the results of other studies are shortly presented.

2. Methodology - The Water Source Diagram (WSD)

The Water Source Diagram divides the process in concentration intervals and water is allowed to be reused along these intervals. The concentration limits of each interval are considered sources of water and supplied or regenerated water is considered an external water source. These concentrations are ordered and represented into a grid of concentrations. Then, the amount of mass transferred ($\Delta m_{k,i}$) in each operation (k) in each interval of concentration (i) is calculated and indicated between parenthesis.

After this initial construction, heuristics rules are followed: i) external water sources are allowed to be used only when internal water sources are not available, ii) the greatest amount of mass must be transferred into the interval of concentration, iii) when an operation is present in several intervals, its water flow must remain along these intervals until its end. The advantage of this procedure is that the network structure is obtained simultaneously with the minimum fresh water consumption target. The amount of mass transferred in g/h (Δm_{ki}) in each operation k and each interval i is calculated through eq.

(1), where C_{fi} is the final concentration of interval i in ppm, C_{ii} is the initial concentration of this interval in ppm and f_k is the mass flow rate in t/h through operation k (k = 1..., N_{op}). The concentration intervals are identified by the index i (i = 1..., N_{int}), where N_{int} is the total number of intervals.

$$\Delta m_{ki} = f_k (C_{fi} - C_{ii}) \tag{1}$$

The flow rate required from the water source p for operation k, in the interval of concentration i, can be determined by equations (2) and (3), where C_{ij} is the concentration in which internal source j is used in interval i, C^e_p the concentration of the external source p, C_{fi} the final concentration of interval i and $N_{fia,i}$ is the number of available internal sources in interval i.

- External water sources:
$$f^e_{pki} = \frac{\Delta m_{ki} - \sum_{j=1}^{N_{fia},i} (f^i_{jki} x[C_{fi} - C_{ij}])}{C_{fi} - C^e_p} \tag{2}$$

- Internal water sources:
$$f^i_{pki} = \frac{\Delta m_{ki} - \sum_{j=p-1}^{p} (f^i_{jki} x[C_{fi} - C_{ij}])}{C_{fi} - C^i_p} \tag{3}$$

First, the internal source must be used in the interval, and the respective values of f^i_{jki} are calculated before f^e_{pki}. The sum in the eq. (3) represents the amount of contaminant removed by internal sources, in operation k in interval i, which has preference for external sources. The minimum external water source flow rate at 0 ppm can be calculated by eq. (4).

$$f^e_p = \sum_{k=1}^{N_{op}} \sum_{i=1}^{N_i} f^e_{pki} \tag{4}$$

3. Results and Discussions

In order to validate this methodology a petrochemical complex is analyzed. The data were extracted from Mann and Liu (1999), who applied the pinch technology to minimize the water consumption. The initial water consumption for the studied operations was 3037 t/d. Only suspended solids were considered as a single contaminant and it was involved the maximum reuse. Table 1 presents the operational data for the major operations concerning water consumption. For WSD application, operations which present water gain or loss were divided in two parts: one having fixed flow rate and the other representing the water gain or loss, as shown in Table 2. For example, operation 1.2 represents water gain and 2.2 and 3.2 water loss. Table 1 and 2 also illustrate the results for the amount of mass transferred ($\Delta m_{k,i}$) in each operation, assuming operational conditions.

In Table 2, all maximum inlet ($C_{in,max}$) and outlet ($C_{out,max}$) contaminant concentrations are presented, which are directly related to limits of corrosion, plugging, scale, deposit formation etc. From these concentrations, using eq. (1) and also considering that the amount of mass transferred was maintained in each operation, the limit flow rates (f_l) of each operation were recalculated.

Table 1. Petrochemical complex operational data [Mann e Liu (1999)].

Operations	f_{in} (t/h)	f_{out} (t/h)	C_{in} (ppm)	C_{out} (ppm)	Δm^* (kg/h)
1 - Dewatering Filters	360	630	0	200	126
2 - Cooling Tower A	1201	188	0	9.2	1.73
3 - Cooling Tower B	1374	261	0	8.7	2.27
4 - Scrubber	60	60	9	207	11.88
5 - Forward Washing	42	42	0	3	0.13

* $\Delta m = f_{out}C_{out} - f_{in}C_{in}$.

Table 2. Original operational data from the petrochemical complex with water gain and loss and the limiting data (in gray).

Operations	f (t/h)	C_{in} (ppm)	C_{out} (ppm)	Δm (kg/h)	f_l (t/h)	$C_{in,max.}$ (ppm)	$C_{out,max.}$ (ppm)
1.1 - Dewatering Filters	360	0	200	72	227.34	6	322.7
1.2 - Dewatering Filters	270	0	-	54	170.51	6	-
2.1 - Cooling Tower A	188	0	9.2	1.73	196.59	6.4	15.2
2.2 - Cooling Tower A	1013	0	-	9.32	1059.1	6.4	-
3.1 - Cooling Tower B	261	0	8.7	2.27	260.92	2.1	10.8
3.2 - Cooling Tower B	1113	0	-	9.68	1112.64	2.1	-
4 - Scrubber	60	9	207	11.88	60	9	207
5 - Forward Washing	42	0	3	0.13	42	0	3

Figure 1 presents the WSD considering maximum reuse obtained and related to the data of Table 2 and Figure 2 shows the mass exchange network corresponding to this WSD of Figure 1. The minimum of water consumption was obtained by the sum of all quantities of fresh water used and showed a reduction from 3037 to 2561 t/d (15.7 %). The same calculations were followed when the WSD procedure was applied for the situation of regeneration with reuse. The only difference is that the concentration of the regenerated stream was assumed as a new limit of concentration and inserted in WSD as a new source of water.

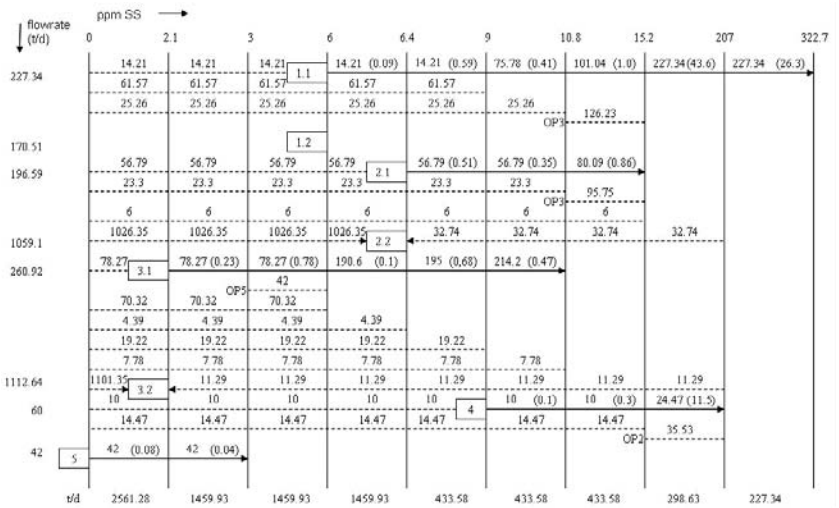


Figure 1. Petrochemical complex WSD with maximum reuse.

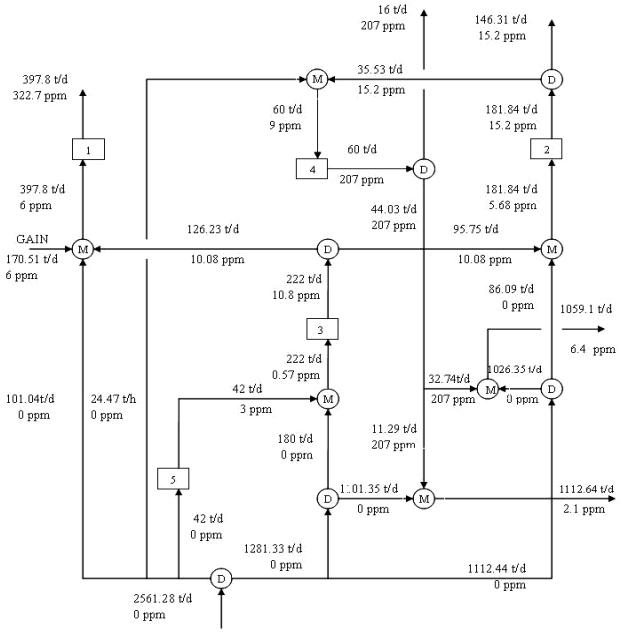


Figure 2. Mass exchange network for maximum reuse.

By the same way, the Water Source Diagram was applied to others industrial segments and it was achieved satisfactory water reduce as observed in the Table 3. The explanation of the WSD procedure for multiple contaminants can be found in Mirre (2007) and Marques (2008).

Table 3. Results for all segments.

INDUSTRIAL SEGMENT	Condition studied	Contaminant	Initial water consumption (t/h)	Water consumption with WSD (t/h)	Water reduce (%)	Water reduce from literature (%)
Pulp and Paper	Maximum reuse	Cl^-	1360	142.1	89.5	-
	Regeneration with reuse			0	100.0	-
Citric Juice	Maximum reuse	COD[1]	240.3	181.3	24.5	21.6[4]
	Regeneration with reuse			164.9	31.4	30.0[4]
Textile	Maximum reuse	COD	43.6	35.8	17.9	17.9[5]
	Regeneration with reuse			21.4	50.8[2]	50.8[5]
Petrochemical	***Simple contaminant***					
	Maximum reuse	SS[3]	126.54	106.72	15.7	14[6]
	Regeneration with reuse		126.54	100.73	20.4	17.5[6]
	Multiple contaminants					
	Maximum reuse	SS	152.79	109.83	22.1	-

[1] Chemical Oxygen Demand; [2] Regeneration with reuse and recycle; [3] Suspended solid; [4] Thevendiraraj *et al.* (2003); [5] Ujang *et al.* (2002); [6] Mann e Liu (1999).

4. Conclusion

The Water Source Diagram procedure applied to four industrial different segments achieved satisfactory water reduces and much better than others procedures from literature. The case studies demonstrated that the WSD application can also be easily used in cases involving maximum reuse and regeneration with reuse through simple calculations.

References

Y.P. Wang and R. Smith, 1994, Wastewater minimization, Chemical Engineering Science, 49, 981-1006.

Y.P. Wang and R. Smith, 1995, Wastewater Minimization with Flowrate Constrains, Trans I Chem, 73(a), 889-904.

P. Castro, H. Matos, M.C. Fernandes and C.D. Nunes, 1999, Improvements for mass-exchange networks design, Chemical Engineering Science, 54, 1649-65.

J.F.S. Gomes, E.M. Queiroz and F.L.P. Pessoa, 2007, Design Procedure for Water/Wastewater Minimization: single contaminant, Journal of Cleaner Production, 15, 474-485.

J. Gomez, M. Salvelski and M. Bagajewicz, 2000, On a systematic design procedure for water utilization systems in refineries and process plants, Chem Eng Commum.

R.C. Mirre, 2007, Water reuse and recovery in the petroleum industry: Synthesis of Mass Exchange Networks, Mastering thesis, 197p. Federal University of Rio de Janeiro (EQ/UFRJ), Brasil.

J.G. Mann, Y.A. Liu, 1999, Industrial Water Reuse and Wastewater Minimization, New York: McGraw-Hill, 523 p.

S.V. Marques, 2008, Water consumption and wastewater minimization – Case Studies, Mastering thesis, 241p. Federal University of Rio de Janeiro (EQ/UFRJ), Brasil.

S. Thevendiraraj, J. Klemes, D. Paz, G. Aso, G.J. Gardenas, 2003, Water and wastewater minimization study of a citrus plant. Resources Conservation & Recycling, 37, 227-250.

Z. Ujang, C.L. Wong, Z.A. Manan, 2002, Industrial wastewater minimization using water pinch analysis: a case study on an old textile plant, Water Science and Technology, 46, 77-84.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Process Integration Methods in Pulp and Paper Industry for Water Usage Reduction

Marcelo Hamaguchi[a], Song Won Park[a]

[a] *Department of Chemical Engineering - USP Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil.* hamaguchi.marcelo@gmail.com

Abstract

The Pulp and Paper industry has struggled to solve problems related to high water consumption. One alternative is to use tools like Process Integration (PI) which includes "Water Pinch", a graphical method, and mathematical programming, which can be used to optimize the water circuit. In a case study in Brazil these two techniques were applied. Data was collected in real time from an integrated thermo mechanical pulping and newsprint mill and then used to develop a simulation that served as the base case model. The analysis is based on residual water reuse in other parts of the process under constraints of concentration limits of dissolved solids and fine fibers. The concentration limits of contaminants were frequently exceeded as shown when recalculated by simulation. Therefore, the results had to be corrected before they became valid. So if the predicted value for reduction of fresh water consumption is 31%, in actual fact is only 23%, which is still a good result.
Keywords: Process Integration, Pinch, Mathematical Programming, Newsprint Mill

1. Introduction

The water circuit analysis by process integration, which can be developed through graphical methods or mathematical programming, is an approach that has its origin in the industrial chemical process. El-Halwagi; Manousiouthakis (1989) introduced the notion of "Mass Exchanger Network Synthesis" (MEN), which initiated a sequence of other publications related to the subject. Alva-Argaéz et al. (1998) used the concept of Process Integration and developed a super structured model, combining graphical methods and mathematical programming. Bagajewicz; Savelski (2001) researched water circuits where only one contaminant is present in the process. Their work showed the possibility of optimizing the water circuit using linear programming. In order to develop graphical methods, Dhole et al. (1996) devised the "composite curve" method where a graph is built using curves of the sources and the demands of water in a process. Other authors like Hallale (2002) have also presented graphical methods to minimize the fresh water consumption. In the pulp and paper sector, Jacob et al. (2002) analyzed, as part of their work, the water circuit using the Pinch method as applied to TMP and newsprint manufacture. In this case, the method was not able to reach multi-objectives targets and the use of mathematic programming was needed to overcome the obstacles. Others authors (Mariano et al. 2007; Jödicke et al 2001) mention that a model has to approach the actual process taking into account environmental and infrastructure aspects. Process simulation in the pulp and paper industry (Syberg; Wild; Simons, 1992) became a useful tool for studies that involve the minimization of water consumption. However, according to Jutila; Leiviska (1981), the problem with simulations of systems in the pulp and paper industry are related to calculation techniques in the model containing recycled streams and it has not yet been entirely solved. Dabros et al (2005) developed a

model after observing the operation of thermo mechanical pulping integrated with four newsprint machines. Simulation was used to analyze sheet breaks and develop alternatives to increase the paper machine runnability.

2. Mill Description

The mill operation consists of two TMP lines and one paper machine responsible for producing 185000 t/y of newsprint paper. The simulation was modeled in Cadsim software.

2.1. TMP production

Stored chips are transferred to the refining system, where they are mechanically processed to form pulp with a consistency of approximately 30-35%. The refining consists of two stages: primary refining, where the chips are refined, followed by secondary refining, where the processed chips from the primary stage are refined to produce the optimal fiber quality. When chips are defibered at high temperatures and high consistence, as occur in the refiner, the fibers are in a high energetic state called latency. The pulp is diluted with a high volume of water and after passing through the latency tank, the pulp is sent to the screens where impurities and residual fibers are removed. The rejects from both lines are diluted before going to a tank which feeds the rejects press. In this stage, the extracted water is reused in the process and the rejects are sent to the refiner located in the line 1. The accepts are sent to the filters, whose function is to concentrate the pulp to a consistency between 7 and 10%. The concentrated pulp is then stored and the residual water returns to the process.

2.2. Mass Preparation and Paper Production

The TMP is mixed with others components necessary to produce newsprint paper, i.e., recovered fiber, refined kraft pulp and paper reject. The mix tank receives the proper proportion of each component. The pulp is then diluted to about 1%, screened and centrifugally cleaned before going to the paper machine headbox.

The paper production starts at the headbox, which has a position of extreme importance in the process. It is responsible for the distribution of fibers in suspension in the forming section. After the headbox, the fiber suspension is dewatered on the forming wire utilizing foils and suction boxes. The next stage in the process is the press section, where the sheet dryness increases from 16-18% to 43%. The press section imparts important properties to the paper including the correct sheet density and is the most economical method of removing the water at this point of the process. After the press section the sheet is heated in the dryer section to increase the dryness to 90% which is its natural moisture level. The white water extracted in the forming section has considerable fiber content. Thus, the fiber recovery system is installed in order to separate two components: clear water for reuse and recovered fiber.

3. Saving Water through PI

Process Integration (PI) enables the development of efficient mass and energy systems for industries. It is based on applications of methods that facilitate the selection and modification of processes.

3.1. Water Pinch

The hypothesis for determining the “water pinch” is that water is used to absorb contaminants. This section presents the graphical method introduced by Dhole *et al* (1996) to deal with conditions where water flows are more important than contaminant load. According to Jacob *et al* (2002), this graphical tool has been applied with success

in the Pulp and Paper area. After defining the sources and demands in the water circuit, the composite curve can be constructed. The curve is used to find the desired consumption of freshwater and to analyze the potential of water reuse with fixed flows. The outflow of each operation that consumes water can be considered a *source* of water. Analogously, the inflows of each operation are the *demands* that have limits for contaminants concentration and can be supplied by a source. The sources and demands flows are indicated in the abscissa, while the respective purities are indicated in the ordinate in decreasing order.

Table 1: Data for source and demand curves

Description	**Order fiber/DS**	**fiber**	**SD**	**Flow**	**Description**	**Order fiber/DS**	**fiber limit**	**DS limit**	**Flow**
		%	**ppm**	**L/min**			**%**	**ppm**	**L/min**
Water extracted from screen 1	S1/S3	0,400	4747	605,7	Dilution of refiners	D1/D1 D2/D2 D5/D4 D6/D5	0,800	5000	94,0
White water from forming section	S2/S9	0,351	2273	54323,0	Water for rejects filter	D3/D14	0,800	3000	299,1
Water extracted from screen 2	S3/S2	0,300	4747	1781,2	Water for TMP filter	D4/D3	0,800	5000	1585,0
Clear filtrate from Save-all	S4/S7	0,050	2487	17450,7	Dilution of rejects from screens	D7/D7 D7/D7	0,800	5000	39,8
Cloudy filtrate from Save-all	S5/S4	0,050	4662	1087,2	Dilution after bleaching	D9/D10	0,800	3000	1483,6
Filtrate from drum filter	S6/S1	0,050	4849	5544,1	Dilution of TMP tank	D10/D11	0,800	3000	1254,9
Clear filtrate from disc filter	S7/S5	0,030	4662	2537,3	Dilution of couch pit	D11/D12	0,800	3000	996,7
Super clear filtrate from Save-all	S8/S6	0,025	2487	2316,2	Water for Save-All filter	D12/D13	0,800	3000	997,4
Filtrate from rejects filter	S9/S8	0,010	2380	1406,8	Water for fiber recovery	D13/D16	0,300	3000	20966
					Water for mixing pump	D14/D15	0,300	3000	43822,8
					Dilution of latency tanks	D15/D6 D16/D9	0,100	5000	8279,3
					Fresh water for TMP	D17/D17	0,018	2500	787,0
					Fresh water for Paper Machine	D18/D18	0,018	2500	3420,0

The advantage of Pinch method is that one source can be splitted into many streams to satisfy different demands, which affect the capital costs. For this reason, the results presented in this paper will focus on the application of this method, although linear programming has also been applied successfully.

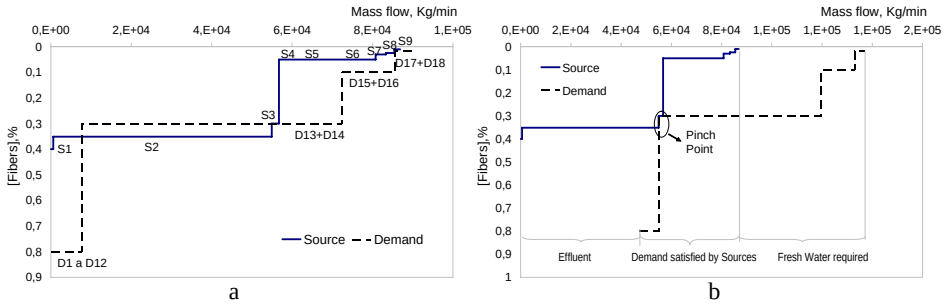


Figure 1 – Construction of sources and demands curves (a). Then the pinch point found after dislocation of demand curve (b).

Figure 1a shows the curves of sources and demands. The objective is to dislocate the demand line to the right and find the pinch point. Figure 1b divides the graphic into three ranges: a) The minimum effluent flow (low purity) is below the pinch point; b) The demands that can be satisfied by available sources; c) Fresh water is used above pinch point to satisfy the demands with high purity;
This system can be improved in order to reduce the consumption of fresh water and effluent flow if "water pinch" point is dislocated. It can be achieved if we increase the purity by mixing the sources. At figure 2a, source flows S2, S3 and part of S4 are mixed to create the stream (S2+S3+S4') to obtain the purity of D3+D14 and thus, dislocate the demands curve to the left in order to find a new pinch point.

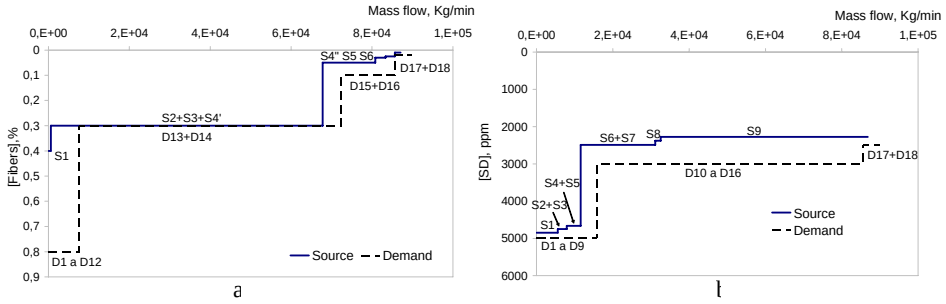


Figure 2 – Mixing of sources to satisfy the limits of demands (a). Mixing of sources to satisfy the limits of DS (b).

According to figure 2b, there is no need to manipulate the curves of water sources and water demands due to the level of dissolved solids. As observed by Jacob et al. (2002), pinch method can be partially automated but is applicable for only one contaminant. On the other hand, linear programming applied to this work for one contaminant added to a concentration of pulp fines, is totally automated and better suited for complex systems.

3.2. Linear Programming

This section explains the concept of mathematic programming, which is an effective method for the analysis, synthesis, and modification of industrial systems aiming for the reuse of water and the minimization of effluent. According to Jacob *et al.* (2002), linear programming is a powerful tool to analyze, simultaneously, the reduction of fresh water consumption and the reduction of and recoverable fibers going to effluent.
Linear programming is able to overcome the disadvantages of Pinch method when applied to this work. It can be used to optimize the water circuit through multi

component analysis, while on the other hand, the Pinch method is restricted to only one component. In order to formulate a case study, Jacob *et al.* (2002) stated that the following constraints have to be defined: (1) The total demand flows fed by a source i cannot be higher than the flow available of this source; (2) Each demand j must be totally satisfied; (3,4) The amount of contaminants provided to demand j cannot be higher than the maximum admissible value defined for this demand.

$$\sum_{j=1}^{n} p_{ij} \le F_i \quad (1) \qquad \sum_{i=0}^{m} p_{ij} = D_j \quad (2) \qquad \sum_{i=1}^{m} c_i p_{ij} \le D_j C_j \quad (3) \qquad \sum_{i=1}^{m} f_i p_{ij} \le D_j F_j \quad (4)$$

The objective function f(x) is the amount that must be minimized under constraints defined by linear equations. Since the objective of the analysis is to minimize fresh water consumption, the objective function is cited as follows:

$$f(x) = \sum_{j=1}^{n} p_{0j}$$

The intention of the method is to reduce the effluent flow and fresh water consumption using qualitative and quantitative constraints imposed by the model. Despite the possibility of analyzing the recovery of fibers in the circuit, this paper will focus on this area. The linear programming in the present work was developed in *MatLab*.

4. Results after Simulation

There is a disadvantage for both methods when the results are transferred to the simulation model. It happens because when original sources are manipulated to satisfy the desired demands, the optimized flows of sources are changed. Since new limits for contaminant concentrations are imposed by the user, the flows defined by the consistency control valve are affected. Besides that, the sources that originated from the filters have fixed consistencies in the model but actually vary with changes in the process. Table 4 shows some differences between the original and optimized values.

Table 4 – Original, simulated, and corrected data after applying Pinch method

	[DS] limit	[DS] Simul.	[DS] after correction	[fibers] limit	[fibers] Simul.	[fibers] after correction	Flow original	Flow Simul.	Flow after correction
	ppm	ppm	ppm	%	%	%	kg/min	kg/min	kg/min
D1	5000	*5294*	4603	0,800	0,400	0,400	94,0	94,0	94,0
D2	5000	*5294*	4603	0,800	0,398	0,398	108,7	110,9	109,0
D3	3000	*5294*	2749	0,800	0,400	0,400	299,1	271,6	272,3
D4	5000	3408	2930	0,800	0,305	0,305	1585,0	1585,0	1585,0
D5	5000	3275	2812	0,800	0,298	0,299	356,2	356,0	356,0
D6	5000	3275	2812	0,800	0,298	0,299	384,1	381,3	373,8
D7	5000	3275	2812	0,800	0,298	0,299	40,0	40,0	40,0
D8	5000	3275	2812	0,800	0,298	0,299	55,0	55,0	55,0
D9	3000	*3275*	2810	0,800	0,298	0,301	1483,3	1538,7	1549,8
D10	3000	*3275*	2806	0,800	0,298	0,303	1254,9	1372,1	1372,5
D11	3000	*3275*	2812	0,800	0,298	0,299	996,7	990,8	991,2
D12	3000	*3275*	2812	0,800	0,298	0,299	997,4	999,9	1000,0
D13	3000	*3275*	2812	0,300	0,298	0,299	20966,0	20966,0	20966
D14	3000	*3282*	2819	0,300	0,272	0,272	43823,0	43824,0	43824
D15	5000	*5298*	4457	0,100	0,050	0,049	8279,0	8262,0	8279,0
D16	5000	3811	3298	0,100	0,028	0,027	4946,0	4886,0	5076,6
D17	2500	*3011*	2500	0,018	0,009	0,009	787,0	62,1	99,3
D18	2500	574	505	0,018	0,002	0,002	3420,0	2824,0	2824,0
Additional Water								**0**	**320,0**
Fresh Water								**4207,0**	**2886,1**

Streams have flows exceeding the concentration limits and therefore appropriate and feasible corrections should be done. Since the software allows the easy manipulation of objects and streams, some new configurations were tested. After testing, the results and their comparison are described also in table 4. The modification requires additional fresh water in some specific points of the process.

5. Conclusion

Both methods, pinch and linear programming, can be used to minimize fresh water consumption with regard to the constraints imposed by the critical contaminants. For the methods applied in the present work, it is very important to know how to define limit values for the concentration of contaminants. Unfortunately, these values are not easily discernable and must be carefully studied and analyzed. Imposing different limit values for the same study case, e.g., can generate significant impacts on the final result, independent of the optimization method utilized. With the integrated paper machine and TMP process, the reduction rate by Pinch and linear programming achieved in the vicinity of 31%. Unfortunately simulation studies demonstrated that many of the streams had unacceptable concentrations of fines and dissolved solids (see table 4). Therefore water usage was increased in the water circuit model in order to reduce the values to acceptable limits, which decreased water savings to 23%, still significant for the industrial unit modeled. The results also show that the application of the Pinch method on an industrial scale is more feasible for industries due to less capital investment.

References

Alva-Argaéz, A.; Kokossis, A.C.; Smith, R. (1998) Wastewater Minimization of Industrial Systems Using an Integrated Approach. Computers and Chemical Engineering. 22 (Suppl) S741-S744.

Bagajewicz,M; Savelski, M. (2001); On the use of linear models for the design of water utilization systems in process plants with a single contaminant. Chemical Engineering Research and. Design 79 (A5) 600-610.

Dabros, M ; Perrier, M ; Forbes, F ; Fairbanks, M ; Stuart, P. (2005) Model-based direct search optimization of the broke recirculation system in a newsprint mill. Journal of Cleaner Production 13 (12) 1416 –1423.

Dhole, V.R.; Ramchandani, N.; Tainsh, R. A.; Wasilewski, M. (1996) Make your process water pay for itself", Chemical Engineering, 103 (1) 100-103.

El-Halwagi, M.M.; Manousiouthakis, V. (1989) Synthesis of mass exchange networks. AIChE J. 35 (8) 1233–1244.

Hallalle, N. (2002) A new graphical targeting method for water minimisation. Advances in Environment Research 6 (3) 377-390.

Jacob, J.; Kaipe, H.; Couderc, F.; Paris, J. (2002) Water Network Analysis in Pulp and Paper Processes by Pinch and Linear Programming Techniques. Chemical Engineering Communication 189 (2) 184 – 206.

Jödicke, G.; Fischer, U.; Hungerbuhler, K. (2001) Wastewater reuse: a new approach to screen for designs with minimal total costs. Computers and Chemical Engineering 25 (2-3) 203-215.

Jutila, L ; Leiviska,K. (1981)The Use of Computer Simulation in the Pulp and Paper Industry. Mathematics and Computers in Simulation 23 (1) 1-11.

Mariano-Romero C.E.; Alcocer-Yamanaka, V.H.; Morales, E.F. (2007) Multi-objective optimization of water-using systems. European Journal of Operational Research 181 (3) 1691-1707.

Syberg, O.; Wild, N.W.; Simons, H.A. (eds.) (1992) Introduction to process Simulation. 2nd edition. Tappi Press, Atlanta.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Exploiting R&D databases for efficient product design: application to brake fluid formulations

Luís H. de Freitas,[a] Galo A. C. Le Roux[b]

[a]*Oxiteno S.A. Indústria e Comércio, Av. das Indústrias, 365, Mauá, 09380-903, SP, Brazil*
[b]*LSCP/CESQ - Department of Chemical Engineering, Polytechnic School - University of São Paulo, Av. Prof. Luciano Gualberto, 380, trav. 3, São Paulo, 05508-900, SP, Brazil*

Abstract

In this work we present a methodology aimed to use the data contained in R&D databases efficiently in order to assist the design of new products by providing models that relate formulations to properties, which are introduced into an optimization problem that represents the design problem together with constraints proposed in order to avoid inaccurate predictions. In this work the methodology is applied to obtain optimal formulations of brake fluids satisfying rigid technical norms and market specifications.
Mixture models are built with Principal Components Regression (PCR) and Partial Least Squares Regression (PLS) which are adequate for systems with incomplete or redundant information, because the informativity of the data with respect to the models is not guaranteed. The models are integrated in an optimization problem, in order to perform the design, which is solved by Mixed Integer Linear Programming (MILP) techniques. Equations that restrain the solution to a set where the information is more trustable are added to the problem. They avoid inaccurate extrapolations that could result in an excessive number of experiments to confirm predictions. The results obtained by this methodology presented good agreement with validation experiments.
The models and optimization tools generated in this work are currently being used at Oxiteno S.A. R&D laboratories and have led to a reduction of time and effort during the development of new formulations.

Keywords: brake fluid, formulations, modeling, optimization

1. Introduction

In this work we present a methodology aimed to use the data contained in R&D databases efficiently. R&D laboratories of chemical industries possess large databases that compile information on a large number of formulations, together with their final properties, of products that were studied once in time. This extremely rich information is in general not exploited as systematically as it could for further product design. In general, the information is in hands of specialists who are in charge of intuitively proposing and devising new formulations. The methodology presented here is aimed to assist the design of new products by providing models that relate formulations to properties, which are optimized with constraints, proposed in order to avoid inaccurate predictions. In this work the methodology is applied to obtain optimal formulations of brake fluids satisfying rigid technical norms and market specifications.

Models are built with adequate techniques for systems with incomplete or redundant information, such as Principal Components Regression (PCR) and Partial Least Square Regression (PLS).

2. Brake Fluids

Brake fluid is the lifeblood of any hydraulic brake system. It is what makes the system operate properly (Scharff, 1989). They are made of mixtures of liquid raw materials employed to transmit pressure within the brake system. In order to do that, brake fluids must perform a variety of functions and must comply with stringent requirements such as Equilibrium Boiling Point, kinematic Viscosity at -40ºC, Evaporation Loss, Wet Boiling Point, Elastomer Swelling, among others. These requirements are defined in various standards of similar content (SAE J1703 - Society of Automotive Engineers J1703, FMVSS 116 - Federal Motor Vehicle Safety Standards 116, ISO 4925 - International Organization for Standardization 4925 and ABNT NBR 9292 - Brazilian Association of Technical Norms). The performance data contained in FMVSS 116, mandatory in the U.S., also works as an international reference (Bosch, 2004). The U.S. Department of Transportation (DOT) has classified the brake fluids into different types such as DOT3, DOT4 and DOT 5.1.

The optimization of brake fluid formulations represents a great competitiveness potential in the automotive segment because it can lead to better quality products with a lower cost. In this work we present the optimization of DOT 3 formulations, used in Brazil mainly in the aftermarket. Besides of that, the raw materials involved in this kind of formulation (mainly glycols and glycol ethers) are available from Ethylene Oxide manufacturers, such as Oxiteno S.A..

The formulations of brake fluids presented here are mixtures of 6 to 8 raw materials chosen out of 18 basic compounds. Each one of the compounds confers different characteristics to the final formulation.

Models are used in order to develop mathematical representations of the design problem that are solved by Linear Programming (LP) techniques. Equations that limit extrapolation problems by restraining the solution to the set where the information is available can be added to the design problem. The problem solved here is similar to the one used in the optimization of formulations of heat stabilizers Poly Vinyl Chloride (Carrillo Le Roux and, da Cruz, , 2006), applied in this case to brake fluid formulations.

In order to validate the methodology, solutions obtained are compared to validation experiments.

3. Materials and methods

3.1. Materials

The development of new products proceeds through steps which target properties are firstly defined and many different “recipes” for obtaining the desired product are tested. The steps that compose the methodology presented here are based on the analysis of historical R&D database, employed in a systematic manner, in order to build models that correlate the final product properties with the compounds in formulations.

This subject is extensively discussed in a recent work (Gani, 2004) where is mentioned that much of the current work in product design is carried out through empirical trial and error approaches, involving time-consuming experiments. Gani (2004) emphasizes the importance to capture the knowledge gained from the past experiments and apply them in a systematic manner in order to reduce the efforts of future trials and, as a consequence, experimentation. According Gani (2004) the major effort is to understand

the molecular structure-property relationships, collect the experimental data, develop the mathematical models and apply the solution techniques.

3.2. Multivariate Calibration

The objective of multivariate calibration techniques is to represent an important quantity of information contained in data that describes a problem by a reduced set of variables, without losing of the relevant information (Carrillo Le Roux and da Cruz, 2006).

In this work, Principal Components Regression (PCR) and Partial Least Squares Regression (PLS) were used in order to build models for mixtures based on historical data and their performances were compared. They are useful to solve collinearity problems (Ferreira et al., 1999; Geladi; Kowalski, 1986; Martens; Naes, 1989).

PCR e PLS generally have similar performances (Beebe; Pell e Seasholtz, 1998; Wise et al., 2003; Rodriguez-Nogales, 2006). According to Esbensen (2006) in the case of modelling only one Y variable, PCR and PLS give similar results, although PLS normally reaches the goal with a less number of factors.

3.3. Optimization

Optimization problems generally can be formulated with mathematical models that involve continuous and/or discrete variables, selected to find equations and/or inequalities, while optimizing the objective function (Grossmann; Biegler, 1995). In the present work the problem was formulated as a Mixed Integer Linear Programming (MILP). Although the *X* variables used here are continuous, binary variables were included in the problem because of the limited availability and/or purity of the raw materials.

4. Application

4.1. Introduction

In this work we present the optimization of DOT 3 formulations, used in Brazil mainly in the aftermarket. Besides of that, the raw materials involved in this kind of formulation (mainly glycols and glycol ethers) are available from Ethylene Oxide manufacturers, such as Oxiteno S.A.. The higher end of the methyl glycol ethers production, such as triethylene glycol methyl ether, is used in hydraulic fluids.

The data set consists of formulations obtained during years of product development, in which 18 different compounds were mixed in different proportions. Formulations are mostly mixtures of 6 to 8 compounds. A more deep description of the formulations will not be presented here because of industrial proprietary reasons. They are grouped into 5 matrices *X* and 5 vectors *Y* that represent respectively the predictors variables (compositions of the formulations) and the chosen properties of DOT3 to be modeled (table 1).

Table 1 - Matrices and vectors dimensions

Variables	Properties of Brake Fluids Formulations				
	BP	VISC	EL	WBP	ESB120
X	170x18	181x18	91x18	63x18	51x18
Y	170x1	181x1	91x1	63x1	51x1

BP - equilibrium reflux boiling point at 760 mmHg (ºC); VISC - kinematic viscosity at -40ºC (cSt); EL - evaporation loss (%w); WBP - wet equilibrium reflux boiling point at 760 mmHg (ºC); ESB120 - effect on styrene butadiene rubber (SBR) at 120°C (mm).

The properties described in table 1 were obtained according to the ABNT NBR 9292 specifications.

4.2. Numerical Tools

Matlab version 6.5 (Mathworks Inc.) was chosen as the framework for the development of the PCR and PLS models, with the use of PLS Toolbox version 3.0 (Eigenvector Vector Inc.).

The problem of optimization was solved by using the Lingo 9.0 solver (Lindo Systems Inc.) that can perform mixed integer linear programming (MILP).

4.2.1. PCR and PLS Linear Mixture Models

In this work we will restrict the analysis to linear mixture models using the equation:

$$\hat{y} = \sum_{i=1}^{18} a_i x_i \qquad (1)$$

In order to transform the data into the most suitable form for the analysis, *X* and *Y* variables were preprocessed, as suggested in the literature (Esbensen, 2006; Ferreira et al., 1999; Geladi; Kowalski, 1986; Martens; Naes, 1989). *X* variables were normalized between -1 and +1 and *Y* variables were autoscaled.

The number of Principal Components (PCs) in the PCR or Latent Variables (LVs) in the PLS were chosen based on the the analysis of the Root-Mean-Square Error of Cross-Validation (RMSECV). This was done using the leave one out methodology, i.e. where each sample is left out of the model formulation and predicted once. In each case, the RMSECV and R^2 was estimated by:

$$RMSECV = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_i)^2}{n}} \qquad (2)$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \qquad (3)$$

4.2.2. Optimization

The MILP model used in this work was formulated in order to minimize the relative cost of the brake fluid formulations and considering that the sum of mixtures compositions must be 100%.

$$min\ f(x) \quad subject\ to:\ h(x,y)=0 \qquad (4)$$

$$lb \leq g(x,y) \leq ub \qquad (5)$$

$x \in R,\ y=\{0,1\}$

where: $f(x)$ - objective function (relative cost of the brake fluid formulations); $h(x,y)=0$ - equations (PCR/PLS models, sum of mixtures compositions must be 100% or decisions variables); $lb \leq g(x) \leq ub$ - inequalities that represent the lower and upper boundaries values for x and y variables.

4.3. PCR and PLS Results

The RMSECV and R^2 statistics for PCR and PLS models are shown in the table 2.

Table 2 - RMSECV and R^2 for PCR e PLS models

Properties	PCs or LVs numbers		RMSECV		R^2	
	PCR	PLS	PCR	PLS	PCR	PLS
BP	16	7	0,282	0,292	0,934	0,930
VISC	18	11	0,259	0,276	0,943	0,938
EL	16	7	0,616	0,588	0,771	0,788
WBP	12	5	0,883	0,873	0,445	0,475
ESB120	13	5	0,597	0,605	0,818	0,810

4.4. Optimization Results

In order to validate the methodology it was used external validation by preparing 5 brake fluids formulations, whose properties were determined according to NBR 9292. They were compared to the previous solutions for the optimization problem. The results for PCR and PLS models are show in the tables 3 and 4 for 2 among the 5 formulations.

Table 3 - Results and relative experimental errors for PCR models

Properties	Formulation 1 - PCR			Formulation 2 - PCR		
	Model	Exp. Value	(% error)	Model	Exp. Value	(% error)
BP (ºC)	234,0	230,9	1,3	250,0	248,6	0,6
VISC (cSt)	1350,0	1366,7	-1,2	1200,0	1185,4	1,2
EL (w%p)	69,8	69,0	1,2	64,6	60,0	7,6
WBP (ºC)	150,2	148,2	1,3	151,3	154,6	-2,1
ESB120 (mm)	0,28	0,33	-15,2	0,29	0,23	26,1
Cost of the formulation	136,36			143,75		

Table 4 - Results and relative experimental errors for PLS models

Properties	Formulation 1 - PLS			Formulation 2 - PLS		
	Model	Exp. Value	(% error)	Model	Exp. Value	(% error)
BP (ºC)	234,0	230,7	1,4	250,0	246,7	1,3
VISC (cSt)	1350,0	1381,4	-2,3	1200,0	1213,2	-1,1
EL (w%p)	70,6	68,8	2,6	63,7	57,2	11,3
WBP (ºC)	149,8	150,2	-0,3	151,1	151,7	-0,4
ESB120 (mm)	0,25	0,20	25,0	0,28	0,37	-24,3
Cost of the formulation	136,23			143,40		

5. Conclusions

PCR and PLS showed no significant differences in their predictive abilities, although PLS required fewer latent variables than PCR. The results obtained by this methodology

presented good agreement with validation experiments. The models and optimization tools generated in this work are currently being used at Oxiteno S.A. R&D laboratories and have led to a reduction in time and effort during the development of new formulations.

References

Beebe, K.R.; Pell, R.J.; Seasholtz, M.B., 1998, Chemometrics: a practical guide. New York: John Wiley & Sons.

Bosch, R., 2004, Automotive handbook, 6th ed. Plochingen: Robert Bosch GmbH, 1232 p.

Cafaggi, S.; Leardi, R.; Parodi, B.; Caviglioli, G.; Bignardi, G., 2003, An example of application of mixture design with constraints to a pharmaceutical formulation, Chemometrics and Intelligent Laboratory Systems 65, 139-147.

Carrillo Le Roux, G.A., da Cruz, U.G., 2006, Development of a mixture design methodology for problems with incomplete information. Application to PVC heat stabilizer design, 16th European Symposium on Computer Aided Process Engineering, v. 21A, p. 1101-1106.

Edgar, T.F.; Himmelblau, D.M.; Lasdon, L.S., 2001, Optimization of chemical processes, 2nd edition, New York: McGraw-Hill.

Esbensen, K.H., 2006, Multivariate data analysis: in practice. 5th edition, Oslo: CAMO Software AS.

Ferreira, M.C.,; Antunes, A.M.; Melgo, M.S.; Volpe, P.L.O., 1999, Quimiometria I: calibração multivariada, um tutorial, Química Nova, v. 22, n. 5, p. 724-731.

Geladi; P.; Kowalski, B.R., 1986, Partial least-squares regression: a tutorial. Analytica Chimica Acta, v. 185, p. 1-17.

Grossmann, I.E.; Biegler, L.T., 1995, Optimizing chemical processes, Chemtech, v. 12, p. 27-35.

Martens, H., Naes, T., 1989, Multivariate Calibration, John Wiley & Sons Ltd., Great Britain.

Pinto, J.M.; Gut, J.A.W., 2005, Otimização de processos químicos.: Apostila do curso PQI5779 - Otimização de Processos Químicos I. Universidade de São Paulo, São Paulo.

Rodriguez-Nogales, J.M., 2006, Approach to the quantification of milk mixtures by partial least-squares, principal component and multiple linear regression techniques. Food Chemistry, v. 98, n. 4, p. 782-789.

Scharff, R., 1989, Complete brake systems, New York: Delmar Publishers, Inc.

Wise, B. M., Gallagher, N.B.; Bro, R.; Shaver, J.M., 2003, PLS toolbox 3.0. Manson: Eigenvector Research, Inc.

Acknowledgement

The authors acknowledge **Oxiteno S.A.**, **FINEP** and **FAPESP** for the financial support.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Prediction of Efficiencies through Simultaneous Momentum, Mass and Energy Transfer Analyses in a Distillation Sieve Tray by CFD Techniques

Dirceu Noriler[a], Henry F. Meier[a], Antonio A. C. Barros[a], Maria Regina Wolf Macel[b]

[a]Regional University of Blumenau, Blumenau, Brasil; [b]State University of Campinas, Campinas, Brasil

Abstract

The main objective of this work is to apply a CFD model under Eulerian-Eulerian framework for gas-liquid flows, with capability to predict the momentum, mass and thermal phenomena of multiphase flows. A two-phase, three-dimensional and transient model with chemical species, energy and momentum conservation balances have been applied for predicting volume fractions, velocity, pressure, temperature and concentration fields, of two-phase flows on sieve distillation tray. The mathematical model was applied in the CFD commercial code for numerical studies, with the construction of a particular numerical grid and with proper sub-routines in FORTRAN language for the closure equations of the model. The results show the profiles as a function of the time and of the position in the distillation sieve tray. The model implemented in this work allows direct application to predict efficiencies in distillation plates, more specifically point, plate, component and global efficiencies.

Keywords: Distillation Sieve Tray, CFD, Efficiency, Mathematical Modeling.

1. Introduction

The equilibrium and nonequilibrium stage models are available to predict the behavior of distillation columns in the open literature [1]. However, they assume perfect mixture in the phases on the plates. However, it has been recognized that the flow pattern on a distillation tray present a large importance on the mass and energy transfer mechanisms. Recent advances and interest in the use of CFD techniques have allowed the study of fluid dynamic in processes and equipments [2]. Contributions have been made in modeling and simulation of gas-solid flow [2-3] and gas-liquid flows, as for example ozonation towers [4-11]. Bubble columns have been firstly considered for modeling bubbling flows [6]. Some works on gas-liquid flow on a distillation tray using CFD techniques can be mentioned [7-10]. The liquid phase flow patterns by solving the time-averaging equations of continuity of mass and momentum only for the liquid phase were studied and the modeling of two-phase flow considering the importance of liquid phase and the gas action with empirical equation were also evaluated [7]. van Baten and Krishna [8] suggested a three-dimensional and multiphase model to represent the hydrodynamics on a distillation sieve tray. The authors use the Eulerian-Eulerian framework and the Reynolds Averaged Navier-Stokes (RANS) equations for modeling turbulent gas-liquid flows assuming that momentum exchange considers only bubble-liquid interaction, i.e., the bubble-bubble interaction was not considered. The two-phase flow behavior considering that the drag force between phases is very high for

distillation cases implies that the gas and the liquid velocities are the same. van Baten and Krishna [8] showed a comparison between heterogeneous and homogeneous models for free surface flow and bubble columns. Gesit et al. [9] applied the heterogeneous model for commercial-scale sieve tray and compared the results with experimental data. Noriler [10] proposed to use the homogeneous model for predicting the hydrodynamic on a distillation sieve tray. Nevertheless, the energy and mass conservations are neglected in most of the works. The model is an improvement of the model proposed by Liu et al. [7] that considers the effects of the gas phase on the liquid phase by empirical equations. The energy equation was available in Noriler et al. [11] that use a two-phase model with empirical equations for momentum and energy inter-phase transfers. The momentum transfer was correlated using the equation proposed by Krishna et al. [8] and the energy inter-phase transfer using the Nusselt number for a sphere particle. The results showed that the temperature fields have an unsteady state behavior, as also occur with volume fraction and velocity fields. The authors demonstrated that the liquid temperature decreases from the inlet up to the weir and the gas temperature increases from the holes up to the top of the tray. In this work, a CFD model under Eulerian - Eulerian framework for gas-liquid flows was implemented, with capability to predict the main phenomenological aspects of multiphase flows. The two-phase, three-dimensional and transient model, with mass continuity, energy, momentum and chemical species conservations have been applied for predicting the volume fractions, velocity, pressure, temperature and composition fields of the two-phase flow on the distillation sieve trays. The model was implemented in the CFD commercial code that uses finite-volume method with variables located in a generalized co-ordinate system for typical operating conditions taken from the literature.

2. Mathematical Modeling

The model considers the flows of gas and liquid in an Eulerian-Eulerian framework, where each phase is treated as inter-penetrating continuum having separated transport equations. A two-phase, three-dimensional and transient model with chemical species, energy and momentum conservation balances have been applied for predicting volume fractions, velocity, pressure, temperature and concentration fields, of two-phase flows on sieve tray distillation. The closure equations for interphase momentum, energy and mass transfers were obtained from van Baten and Krishna [8] methods and the correlation for global coefficient mass transfer from Ranz-Marhan [11]. The turbulence was treated by k-ε standard model.

2.1. Boundary and Initial Conditions

Due to the elliptical characteristics of the partial differential equations of the model, boundary conditions of all frontiers of the physical domain are necessary: at the inlet, uniform profile of velocities and turbulent properties are imposed; no slip conditions on the wall for both phases; and pressure conditions in the outlet were also applied for the two phases.

3. Solution Methodology

Due to the model complexity in a 3-D space-time, it is necessary an efficient strategy to guarantee the stability of the numerical solution. Therefore, it was formulated a methodology to obtain the solution composed by four steps. The first step consists on to solve the model without the energy and chemical species equations. This solution was carried out up to 15 s of real time. Then, this result was used as initial condition to the

second step of the solution, where the energy equation together with the energy interphase transfer were incorporated to the model and the solution was carried out up to 35 s of real time. From this solution, started the next step, which considers the chemical species equations in the model, but without interphase chemical species transfer. In this step, the solution carried out up to 50 s of real time. In the last step, the interphase chemical species transfer was incorporated to the model and the solution was carried out until the system reached the quasi-steady state condition, i. e., about 80 s of real time for all cases analyzed in this work. Noriler [10] presents the solution analyses of the partial differential equations of the model. Based on this, it was applied the Finite Volume method on generalized coordinated and collocated grid to solve the model. A high-order interpolation scheme was used for hydrodynamic equations and upwind for turbulence equations. The pressure-velocity coupling was obtained using SIMPLEC algorithm. The improved RHIE-CHOW algorithm was used to calculate the velocity at the cell faces in order to avoid numerical problems like check-boarding and zigzag. The relaxation factors were not used.

4. Results and Discussions

Based on the physical domain, it was used a distillation sieve tray with 0.30 m diameter, D, and 0.180 m weir length, W. The column has a rectangular entrance for the liquid with 0.180 m x 0.015 m. The weir height, h_w, is 0.080 m, and the fractional holes area to bubbling area, A_h/A_b, is 0.0654. The length between the liquid inlet and the weir is 0.240m with 180 holes in the tray. Figure 1 shows the clear liquid height as a progress variable of the solution.

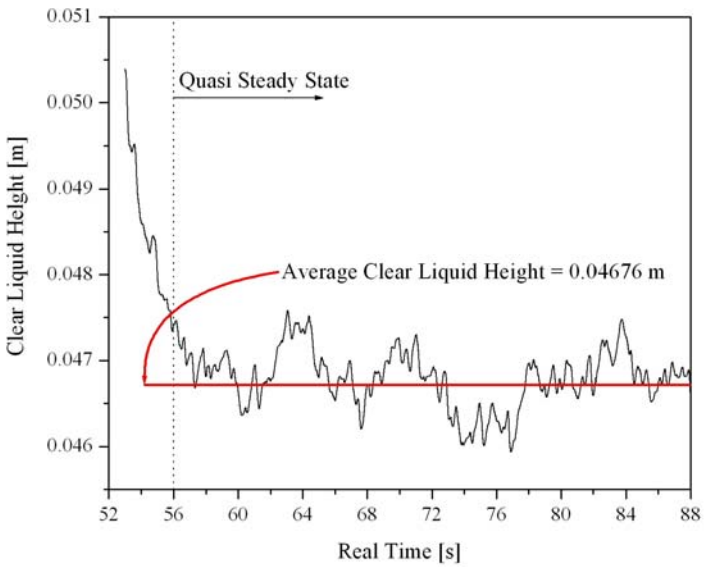


Figure 1: Clear liquid height as a function of real time.

It can be seen in Figure 1 that the flow does not present a steady state well defined, but indeed, present a chaotic regime, characteristic of the quasi steady state, as it was verified in previous works [8-11].

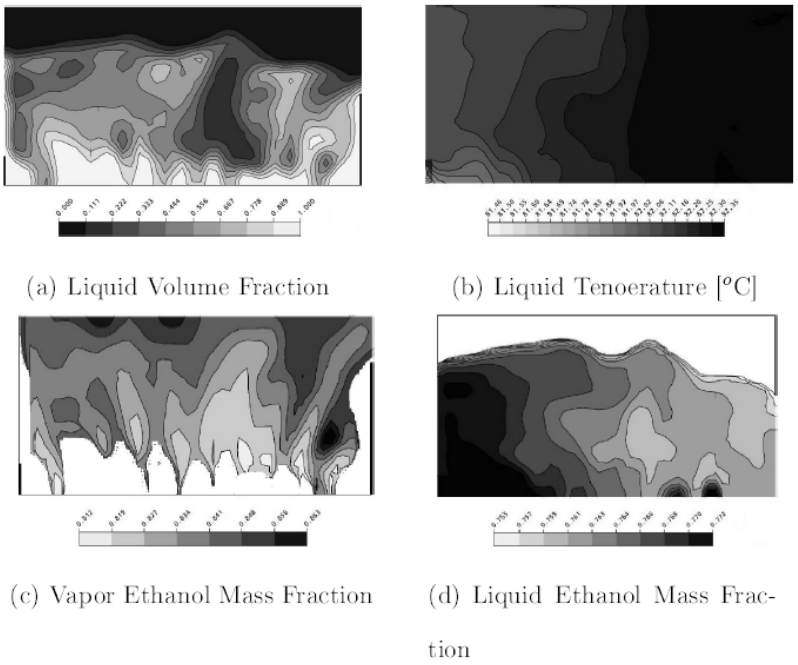


Figure 2: Snapshots of the central plane from the liquid inlet up to the weir of sieve

Figure 2 presents the snapshots of the central plane from the liquid inlet up to the weir of the sieve tray for the liquid volume fraction (Figure 2a), liquid temperature (Figure 2b), ethanol mass fraction in the vapor (Figure 2c), and ethanol mass fraction in the liquid (Figure 2d). It can be seen in Figure 2a, a high liquid volume fraction region near to the bottom, a froth region and a region of the continuous vapor. In this way, Figure 2b evidences that the liquid temperature field is not influenced by the volume fraction field, presenting higher temperature gradients in the liquid flow direction. This fact is due to the high global heat transfer coefficient that results in a high interphase energy transfer rate. In Figures 2c and 2d, it is observed the ethanol mass fraction in the vapor and in the liquid phases, respectively. The white regions in these figures represent absence of the phase. The vapor mass fraction of ethanol increases in the vapor flow direction, from the basis of the tray up to the top of the tray, while the liquid ethanol mass fraction decreases in the liquid flow direction, from inlet liquid up to the weir. These results can be used for developing equations of sieve distillation tray efficiencies based, for example, on the fundaments proposed by Murphree [13] and West [14]. From the results presented in Figure 2, it is possible to determine the global tray efficiency. For the operational and geometric conditions used in this case, the efficiency for this plate and system is 64.16 %. When it is applied the point efficiency concept suggested by West [14], where the tray is divided in regions and then is applied the Murphree [13] efficiency, it can be calculated the efficiency as a function of the position. This information can be used to understand the tray behavior and then to optimize it. The point efficiency is observed in Figure 3. It can be noted that the tray presents high efficiencies near to the inlet and near to the weir. This is due to the recirculation zones that appears together the inlet and weir. Near to the weir, the efficiency reaches the maximum value because the recirculation zones increases the liquid residence time and the ethanol in the vapor phase is the minimum. In the middle of the tray, the efficiency

has small value due to the fact that the liquid phase presents the preferential way in this region.

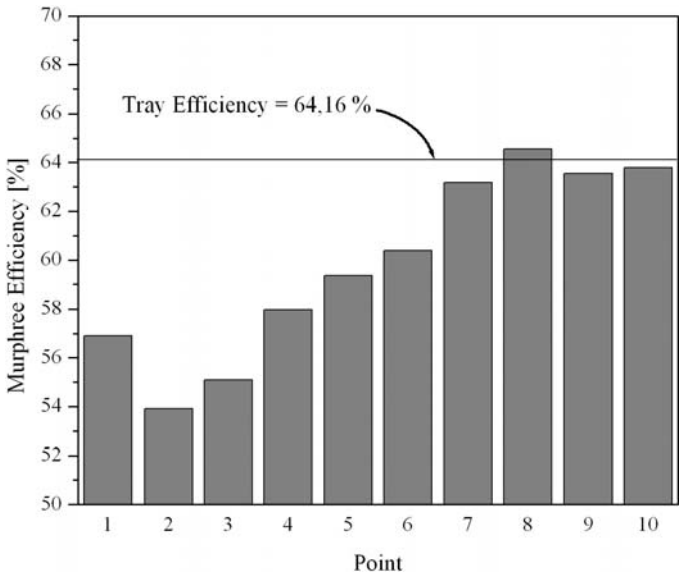


Figure 3: Point efficiency

5. Concluding Remarks

The methodology proposed in this work showed to be suitable, with capability to describe, from CFD techniques, the vapor-liquid flow on a sieve distillation tray. It was possible to calculate the tray efficiency based on models that consider spatial variation in all properties, besides pressure and velocity variations. This is important to design and to optimize distillation trays.

References

[1] A. A. C. Barros, M. H. Pescarini, M. R. Wolf Maciel, 1996, Computer Aided Process Engeneering, 279-284.
[2] H. F. Meier, J. J. N. Alves, M. Mori, 1999, Computers & Chemical Engineering, 23, 247-262.
[3] H .F. Meier, K. Ropelato, K., H. Forster, J. J Iess, M. Mori, 2002a, zur Berechnung und Auslegung von Zyklonen - Teil 1. ZKG International, 55 , 64-75.
[4] Y. Cheng, F. Wei, Y. Guo, Y. Jin, W. Lin, 1998, Powder Technology, 98, 151-156.
[5] Y. Cheng, F. Wei, Y. Guo, Y. Jin, 2001, Chemical Engineering Science, 56, 1687-2001.
[6] R. Krishna, J. M. van Baten, M. I. Urseanu, 2000, Three-phase eulerian simulations of bubble column reactors operating in the churn-turbulent regime: a scale up strategy. Chemical Engineering Science, 55, 3275-3286.
[7] C. J. Liu, X.G. Yuan, K. T. Yu, X. J. Zhu, 2000, Chemical Engineering Science, 55, 2287-2294.
[8] J. M. van Baten, R. Krishna, 2000, Chemical Engineering Journal, 77, 143-151.
[9] D. Noriler, 2003, MSc Thesis, Laboratory of Separation Process Development, State University of Campinas, Campinas, São Paulo,Brazil.
[10] D. Noriler, 2007, PhD. Thesis, Laboratory of Separation Process Development, State University of Campinas, Campinas, São Paulo, Brazil.

[11] D. Noriler, A. A. C. Barros, H. F. Meier, M. R. Wolf Maciel., 2008, Chemical Engeneering Journal. 136, 133-143.
[12] R. Rahimi, M . Rahimi, F. Shahraki, M. Zivdar, 2006, IChenE, 220-229.
[13] E. V. Murphree ,1925, Ind. Eng. Chem, 17, 747.
[14] F. B. West, 1952, Ind. Eng. Chem, 44, 2470-2478.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

On the topological analysis of industrial process data using the SOM

Francesco Corona,[a] Michela Mulas,[b] Roberto Baratti,[c] Jose Romagnoli[c,*]

[a]*Dept. of Information and Computer Science, Helsinki University of Technology TKK, P.O. Box 5400, Espoo FI-02015 TKK, Finland*
[b]*Dept. of Biotechnology and Chemical Technology, Helsinki University of Technology TKK, P.O. Box 6100, Espoo FI-02015 TKK, Finland*
[c]*Dept. of Chemical Engineering and Materials, University of Cagliari, Piazza d'Armi, Cagliari I-09123, Italy*

Abstract

In this paper, we overview and discuss the implementation of some topological approaches to modeling and analyzing industrial process data. The discussed methods are used in visualizing process measurements and extracting information by exploiting the metric structure of the observations. Emphasis is given to modeling with the Self-Organizing Map (SOM). The SOM is a standard method for dimensionality reduction and vector quantization equipped with many displays for visualization. Some of the possibilities of the SOM with process data are discussed by exploring measurements from a full-scale gas treatment plant where the goal is to identify important operational modes and sensitive process variables before developing an alternative control strategy.

Keywords: Process monitoring, Process supervision, the Self-Organizing Map.

1. Introduction

Modern process plants are motivated to maintaining and improving product quality and profit while satisfying more stringent environmental and safety constraints. For efficient operation, any decision-making action related to plant operation requires the knowledge of the actual state of the processes. The availability of easily accessible displays and intuitive knowledge of the states is invaluable with immediate implications for profitability, management planning, environmental responsibility and safety.

Due to the advances in measurement and information technology, process industry is also passing through the era of data explosion; historical data are available in abundance. The data derived approach to process monitoring and supervision seeks to construct a representation of a process from a set of easily accessible measurements that quantify its behavior, without explicitly assuming any *a priori* knowledge of the underlying phenomena. However, the task is daunting and remains a major concern. It is necessary the availability of effective methods that *i*) model the data to extract the structures existing in the measurements, *ii*) automatically identify and reconstruct the most relevant structures for the scope at hand, and *iii*) allow for easily interpretable displays where the states' information is presented to the plant operators. Intuitive knowledge of all visited states is invaluable for a safe plant operation and trustworthy methods are a necessary part in a modern supervision and control strategy.

---

[*] On leave, from the Department of Chemical Engineering, Louisiana State University, USA.

In this paper, we discuss a data derived approach to model, visualize and analyze the information encoded in industrial process data. The approach is based on a classical machine learning method, the Self-Organizing Map, SOM (Kohonen, 2001). The SOM combines many of the properties of other techniques and shares many commonalities with two standard methods for data projection (the Principal Components Analysis, PCA; Jolliffe, 2002) and clustering (the K-means; Hartigan and Wong, 1979). In addition, the SOM is also provided with a set of tools that allow for efficient data visualization in high-dimensional settings. Kaski (1997) and Vesanto (2002) discussed the use of the Self-Organizing Map in exploratory data analysis, and their methods are widely known in many fields, see Oja et al. (2003). The major contributions in using the SOM on process data were proposed by Alhoniemi (2002) and Laine (2003). The SOM is here used as a common framework to identify the process modes for a full-scale gas treatment plant and present the extracted information on intelligible displays.

2. Topological data analysis with the SOM

The Self-Organizing Map is an adaptive formulation of the quantization paradigm on a low dimensional space. Thus, the method singles out from other machine learning techniques because it performs both data clustering and projection. The main advantage of the SOM is in the visualization tools that allow for an intuitive data exploration.

The basic SOM consists of a low-dimensional regular array of K nodes where a prototype vector $\mathbf{m}_k \in \mathbf{R}^p$ is associated with every node k. Each prototype acts as an adaptive model for the observations $\mathbf{v}_i \in \mathbf{R}^p$. During the computation, the data are mapped onto the SOM array and the model vectors adapted according to the expression:

$$\mathbf{m}_k(t+1) = \mathbf{m}_k(t) + \alpha(t) h_{k,c(\mathbf{v}_i)}(\mathbf{v}_i(t) - \mathbf{m}_k(t)). \quad (1)$$

In the learning rule in Eq. 1, t denotes the discrete coordinate of the mapping steps and $\alpha(t) \in (0,1)$ is the decreasing learning rate. The map is computed recursively for each observation. The scalar multiplier $h_{k,c(\mathbf{v}i)}$ indicates the neighborhood kernel that, if chosen in its Gaussian form, as in Eq. 2, acts as a smoothing function

$$c(t) = \alpha(t) \exp\left(-\left\|\mathbf{r}_l - \mathbf{r}_c\right\|^2 / 2\sigma^2(t)\right), \quad (2)$$

centered at the Best Matching Unit, BMU (the model $\mathbf{m}_c(t)$ that best matches with the observation $\mathbf{v}_i$). The vectors $\mathbf{r}_l$ and $\mathbf{r}_c$ (both in $\mathbf{R}^2$, for the 2D map) represent the geometric location of the nodes on the array. Upon performing a parallel comparison algorithm, the subscript c (in Eq. 1 and 2) is assigned to the BMU. The competitive criterion for comparison is usually the Euclidean norm. During calibration, the model vectors learn a nonlinear manifold in the original space such that the relevant topological and metric properties of the observations are preserved on the low-dimensional map. Here, the structures are represented by geometric relationships.

In the case of a slightly modified algorithm with a fixed kernel function, the SOM can be also understood from the optimization of a cost function (Heskes, 1997). The cost closely relates to the objective of the K-means algorithm; the only difference being in the neighbourhood function. In that sense, the SOM operates as the conventional clustering method, if the width of the kernel is constantly set to zero. The SOM has also neat projection properties: the cost also closely resembles the objective optimized by Curvilinear Components Analysis, CCA (Demartines and Herault, 1997).

In the typical case of projections onto 2D arrays, the SOM offers excellent displays for data exploration. In that sense, the approach to topological data analysis with the SOM is mainly visual and intuitive. The visualization techniques used in the experiments are the common component planes and the U-matrix (Vesanto, 2002).

3. Application to an industrial deethanizer

To illustrate the potentialities of topological data analysis using the SOM, a set of measurements from a full-scale process was considered. The monitoring problem consists of modeling and analyzing the operational behavior of an industrial deethanizer, starting from a set of online process measurements. The motivation for choosing this unit is merely illustrative. The column offers a variety of behaviors that reflects the operational usage, hence an interesting groundwork for discussion.
The deethanizer, in Fig. 1, is operated to separate ethane from a feed (a light naphtha) while minimizing the ethane extracted from the bottom (an economical constraint for the subsequent unit in the plant). The operational constraint is in the maximum ethane loss from the bottom; at present, the operational threshold is set to be smaller than 2%.

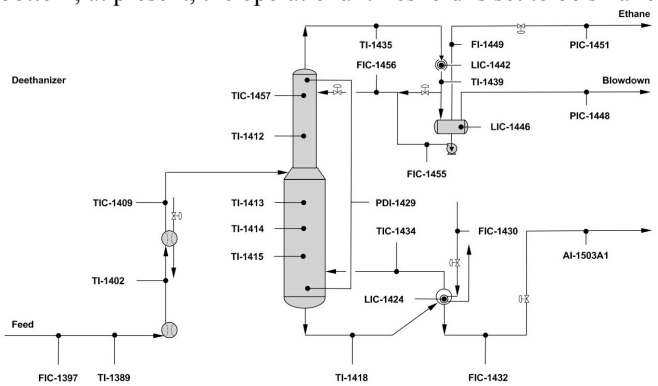


Figure 1. Deethanizer: Simplified flow-sheet.

In order to analyze the behavior of the unit, measurements from a set of process variables was collected from the plant's distributed control system (DCS). The data correspond to 3 weeks of continuous operation in winter asset and 3 weeks in summer asset and are available as 3-minute averages; overall, 27 variables are acquired, for a macroscopic characterization of the unit. In addition, there are several control loops in the process. The column temperature TIC-1457 and the vapour temperature TIC-1434 out of the reboiler is controlled by manipulating the reflux flow-rate FIC-1456 and the steam flow-rate FIC-1430 to the reboiler, respectively; both loops are cascaded to the corresponding flow-rates. The distillate pressure PIC-1451 is controlled by the distillate flow-rate FI-1449 and the reboiler's level LIC-1424 by the bottom flow-rate FIC-1432.
The operational objective of the column is to produce as much ethane as possible (minimizing the propane's concentration from the top of the column) while satisfying the constraint on the impurity from the bottom (ethane's concentration smaller than 2%). Based on the ethane loss from the bottom, three operational modes are identified:

- a *normal* status, corresponding to the operation of the column, where the concentration of ethane is within allowable bounds (in the 1.8% - 2.0% range),
- a *high* status, corresponding to the operation of the column, where the concentration of ethane is exceeding the allowable upper bound (>2%),
- a *low* status, corresponding to the operation of the column, where the concentration of ethane is below the allowable lower bound (<1.8%).

The two abnormal conditions have direct economic implications; when at *low* status, the process is delivering a product out of specifications whereas, when at *high* status the product is within the specifications, but an unnecessary operational cost is observed.

3.1. Exploration and analysis

To understand under which conditions such modes are experienced, the data were analyzed by exploring their clustering structure. Starting from a selection of relevant variables, we augmented this subset with an additional *dummy* indicator, specifically calculated to indicate the status. The new variable was defined so as to take values +1, -1 and 0, according to the operational status. Value 0 is assigned to the *normal* mode, whereas +1 and 1 correspond to *high* and *low* modes, respectively. Notice that the *dummy* variable requires online measurements for the ethane concentration; such a variable, (AI-1503A1), is available in the DCS from a continuous-flow chromatograph.

The set of variables, augmented by the *dummy* indicator, was used to learn a 2D SOM, from the winter observations only. The map consists of a hexagonal array of prototypes initialized in the space spanned by the eigenvectors corresponding to the two largest eigenvalues of the covariance matrix of the data. As usual, the ratio between the two largest eigenvalues was used to calculate the ratio between the two dimensions of the map. On the resulting map, the SOM's component planes and U-matrix were analyzed.

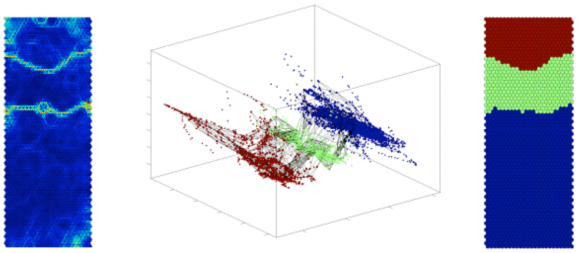


Figure 2. The U-Matrix with three clearly separated clusters in blue (left), the clustered SOM and the winter data after projection onto the principal components space (center) and the map dyed according to the cluster membership obtained from the K-means algorithm (right).

The U-matrix is the display used to visualize the distances between each prototype and its neighbors. The common way to present it consists of an initial projection of the distances onto a color axis and the subsequent display with dyed markers between each node. In the U-matrix, areas with homogeneous coloring (corresponding to small within-cluster distances) are recognized as clusters, whereas cluster borders are identified still as areas with homogeneous coloring but corresponding to large between-cluster distances. The U-matrix is shown in Fig. 2 (left), with a color-coding that assigns blue to close together prototypes (clusters) fading toward red as the distance increases (cluster separations). The visualization permits to recognize 3 distinct taxonomies, as well as several minor substructures. An analogous visualization of the prototypes' grouping is achieved by projecting the nodes onto a low-dimensional subspace, a 3D PCA space (in Fig. 2 (center)). Indeed, also this visualization illustrates the clustering. However, to obtain a quantitative characterization, the prototypes of the SOM can be regarded as a reduced data set and modeled separately with a specific clustering algorithm. We adopted the K-means coupled by the Davier-Bouldin index (as measure of cluster validity) to identify the optimal number K of taxonomies (Milligan and Cooper, 1995). As expected, optimality was found for $K=3$. The operational modes are located in the lower, middle and upper part of the map. The mode-cluster association is straightforward after dying the SOM's region in Fig. 2 (right) with different membership colors (blue, green and red) from the K-means according to a scheme that assigns them to the operational modes (+1, 0 and -1, respectively), and comparing it with the component planes of the *dummy* and ethane (Fig. 3 (left) and (center)).

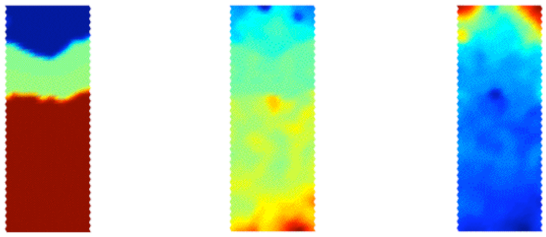


Figure 3. The component planes for the *dummy* variable (left), the ethane concentration (center) and the temperature TI-1414. The coloring ranges from blue (low values) to red (high values).

Fig. 3 also depicts the plane of the column temperature TI-1414 (right); looking for similar patterns in similar positions in the components planes shows a neat dependence between ethane concentrations and temperature measurements. Such pair of variables is characterized by near identical but reversed component planes, thus highlighting the presence of an inherent inverse correlation between the two variables.
To maintain the ethane concentration in normal conditions (within the 1.8-2.2% range, Fig. 4 (left)), the analysis suggests that such temperature should be controlled (within the 52-55°C range, Fig. 4 (center)). The information can be easily retrieved after dyeing the point measurements of the two variables with the corresponding mode membership. A possible variable to manipulate is the steam flow-rate FIC-1430, such a variable is, however, presently unused for control. During the winter period, an overall 85% of off-spec operation, was in fact observed, Fig. 4 (right). The summary is obtained by calculating the number of point measurements falling outside the normality conditions over the total number of observations, and it can be also visually recovered from the area distribution of the clusters in the SOM depicted in Fig 2 (right).

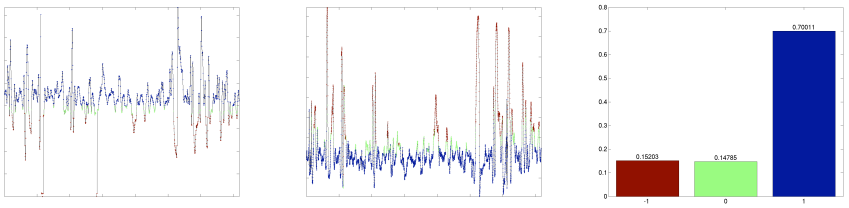


Figure 4. The colored time series for AI-1503A1 (left), TI-1414 (center) and the modes distribution during winter operation (right).

The winter map can also be used to visualize the behavior of the unit in extrapolation; for example, on the summer measurements. The analysis initially projects the new data onto the calibrated SOM, being the mapping based on a nearest neighbor criterion between the new observations and the prototypes. The winter map can be then enhanced by the inclusion of the summer point trajectories followed by the unit. The trajectory, which passes through all the BMUs of each new set of measurements, is shown as red line connecting the visited prototypes, marked as yellow dots. The trajectory makes it possible to indicate the current mode of the process and observe how it was reached.
In Fig. 5, the process trajectory is reported for a small time window corresponding to fifteen hours of continuous summer operation of the deethanizer. Following the temporal evolution, the diagrams show a process that is initially operated in *normal* condition, as for the ethane in the bottom and reference temperature. As the process moved further in time, new prototype vectors were visited and added to the trajectory

until the column leaves the normality region and crosses it towards the region of low ethane composition and starts oscillating. In a similar fashion, all the process variables changed coloring to match the visited modes. The abnormal situation was mainly due to an abrupt decrease in feed flow-rate and possibly composition. In turns, the variation triggered the action on steam to reboiler flow-rate FIC-1430, as well as the reflux to control the top temperature FIC-1436.

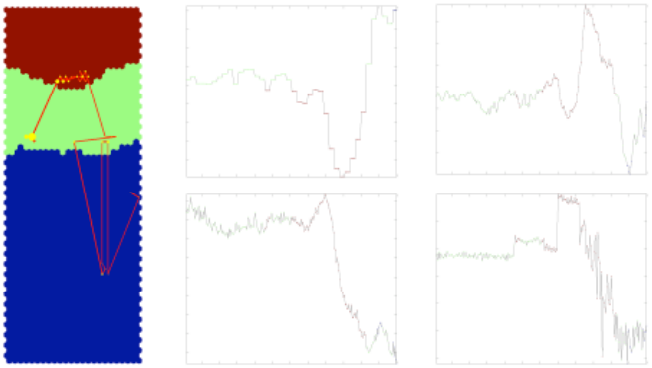


Figure 5. The trajectory and the colored time evolution for the ethane AI-1502A1, the temperature TI-1414, the steam to reboiler FIC-1430 and the reflux flow-rate FIC-1456, in clockwise order

4. Conclusions

In this work, we implemented and discussed a strategy to model, visualize and analyze the information encoded in industrial process data using the Self-Organizing-Map. In particular, the proposed strategy was applied to an industrial distillation column allowing us to illustrate on simple displays how the clustering structure of the measurements corresponds to the operational modes of the units and the possibility to define an alternative control strategy for maintaining it in the normal operating state.

References

E. Alhoniemi, 2002, Unsupervised pattern recognition methods for exploratory analysis of industrial process data, Ph.D. dissertation. Helsini University of Technology, Finland.

P. Demartines, J. Herault, 1997, Curvilinear component analysis: a self-organizing neural network for nonlinear mapping of data sets, IEEE Trans. on Neural Networks 8, 148–154.

A. Hartigan, M.A. Wong, 1979, A K-means clustering algorithm. Applied Statistics 28, 100–108.

T. Heskes, 1999, Energy functions for self-organizing maps, in Kohonen maps, 303-316.

I. T. Jolliffe, 2002, Principal Components Analysis: Second Edition, Springer.

S. Kaski, 1997, Data exploration using self-organizing maps, Ph.D. dissertation ISBN 952-5148-13-0. Helsinki University of Technology, Finland.

T. Kohonen, 2001 , Self organizing maps: Second Edition, Springer.

S. Laine, 2003, Variable selection and feature extraction to learn from industrial data. Ph.D. dissertation ISBN 951-22-6669-5, Helsinki University of Technology, Finland.

G.W. Milligan, M.C. Cooper, 1995, An examination of procedures for determining the number of clusters in a dataset, Psychometrika 50, 159–179.

M. Oja, S. Kaski, T. Kohonen, 2003, Bibliography of Self-Organizing Map (SOM) papers: 1998-2001 addendum, Neural Computing Surveys 3, 1–156.

J. Vesanto, 2002, Data exploration based on the Self-Organizing Map, Ph.D. dissertation, Helsinki University of Technology, Finland.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Delaunay tessellation and topological regression: An application to estimating product properties from spectroscopic measurements

Francesco Corona,[a] Elia Liitiäinen,[a] Amaury Lendasse,[a] Roberto Baratti,[b] Lorenzo Sassu[c]

[a]*Dept. of Information and Computer Science, Helsinki University of Technology TKK, P.O. Box 5400, Espoo FI-02015 TKK, Finland*
[b]*Dept. of Chemical Engineering and Materials, University of Cagliari, Piazza d'Armi, Cagliari I-09123, Italy*
[c]*Process Department., Sartec - Saras Ricerche e Tecnologie S.p.A. Strada 5 Traversa C, Assemini I-09032, Italy*

Abstract

The Delaunay tessellation and topological regression is a local simplex method for multivariate calibration. The method, developed within computational geometry has potential for applications in analytical chemistry and process monitoring. This study investigates the applicability of the method for estimating the aromatic composition in Light Cycle Oil (LCO) by Near Infrared (NIR) spectroscopy.

Keywords: Process monitoring, multivariate calibration, Delaunay tessellation, nearest-neighbor regression.

1. Introduction

Real-time monitoring is an essential component in modern process industry for optimizing production toward high-quality products while reducing operating and off-specification costs. The tools of process analytical chemistry like Infrared (IR) and Near Infrared (NIR) spectroscopy fulfill the necessary requirements for real-time analysis of important properties for a broad variety of materials (Workman, 1999).

The principle underlying process monitoring from spectra is the existence of a relationship between the spectrum of a given product and the property of interest. The relationship is rarely known *a priori* but it can be reconstructed from data by calibrating multivariate models. Multivariate calibration methods are often divided into local and global approaches. The latter use all known (calibration) observations to learn the parameters of a single regression model. The former use only small subsets of the calibration data to build different calibration models located in the neighborhood of the observation whose properties have to be estimated. Classical parametric approaches like Principal Components Regression (PCR) and Partial Least Squares Regression (PLSR) are available in both local and global variants (Gemperline *et al*., 2006). Among local methods, non-parametric approaches as topological regression (Jin *et al*., 2003a and Jin *et al*., 2003b) have gained recent interest, mostly driven by industrial motivations; e.g., an inherent ability to handle nonlinearities and the possibility to minimizing models' maintenance tasks while retaining the prediction accuracy.

With the goal to investigate alternative calibration methods that reduce the maintenance costs associated to continuous recalibrations, this paper discusses the Delaunay

Tessellation and Topological Regression (DTR) method (Jin *et al*., 2003b) and its application to calibrating the aromatic composition in Light Cycle Oil (LCO) by NIR spectroscopy. The application is framed within the intense research activity that has characterized the recent trends in refining aimed at optimizing the use of low-value products. LCO is a low-value stream in the diesel boiling range produced in Fluid Catalytic Cracking (FCC) units. Due to its poor characteristics (i.e., high total aromatic content and a considerable percentage of compact structure poliaromatics, among others), LCO is usually not blended directly in the finished diesel fuel pool but it can be upgraded to the higher value diesel if processed in hydrotreatment units (where aromatic compounds are hydrogenated). Within this framework and in order to assure the required process and environmental standards of the hydrotreated products, rapid and cost effective (and possibly on-line) evaluation of the aromatic content is mandatory. Among the many chemometric methods available for such purposes, the DTR method has recently been proposed as an application of computational geometry. It singles out from standard approaches because of its potentiality to achieve accuracies comparable with PCR and PLSR while being much simpler to develop and maintain/upgrade (Jin *et al*., 2005 and Jin *et al*., 2006). In order to assess the potential of the method, a feasibility study with comparison to standard regression methods was performed.

2. Delaunay tessellation and regression

The Delaunay Tessellation and Topological Regression is a local multivariate calibration method developed from arguments in computational geometry. In its basic formulation, the DTR consists of three main steps: 1) a dimensionality reduction based on the known input observations (e.g., NIR spectra), 2) the generation (in the low-dimensional space) of an unstructured mesh by Delaunay tessellation and 3) a nearest-neighbor (or topological) regression for estimating the outputs (e.g., the aromatic content in hydrocarbon mixtures) for a set of unknown observations.

2.1. Dimensionality reduction

The dimensionality reduction step aims at projecting the input observations onto a system of lower coordinates in such a way that certain properties of the data points are preserved as faithfully as possible. The mapping can be either driven only by the input data points (as in PCR) or by both the input and output data points (as in PLSR).

In general, there is a wide range of other methods to performing dimensionality reduction (e.g., see Lee and Verleysen, 2007) that can be considered in this phase. This study is confined to a projection based only on input data, because this representation can be common to all the output properties; thus, minimizing problems and costs associated to the models' recalibration and maintenance. For the sake of simplicity, a PCA was used to characterize the experiments; in that sense, the local property of the data that is preserved is in the set of pair-wise distances between them (Jolliffe, 2002).

2.2. Delaunay tessellation

Once the input observations are projected onto a low dimensional system of coordinates (e.g., the principal components), the known part of this space is partitioned, by generating a mesh using all the available data points. The elements of the mesh are simplexes delimited by known observations (i.e., input data points with known values for the output properties, the calibration set). Within each simplex, the locality conditions are assumed because similar data should be mapped close to each other.

A well-known method for generating a mesh of simplexes is the Delaunay tessellation (Gudmundsson *et al*. 2002). For a given set of point observations in two dimensions, the Delaunay tessellation constructs a mesh of triangular simplexes (hence, the common

name Delaunay *triangularization*) by using all the input data points as vertices; one triangle is a valid simplex if and only if its circumcircle does not enclose any other point in the set (the empty circle condition). In 3 dimensions, the simplexes are tetrahedrons, and so on. Generally, the Delaunay tessellation of a point set exists and is unique.
Notice that the DTR method requires a mesh generation performed on as many dimensions as those obtained in the dimensionality reduction step. For a reduction to *s* dimensions, the elements of the tessellation are simplexes defined by *k=s+1* points.

2.3. Topological regression

Once the mesh is built, it is used for estimating the properties of new observations; that is, data points for which only the input values are known. Also the new observations must be projected onto the same low dimensional system obtained in the first step.
The standard case for estimation is when a new observation **x**(i) (in the *s*-dimensional projection space) happens to fall within the convex hull that contains the know data points. In such a situation, the new point also falls within one of the simplexes.
In two dimensions, the simplex is a triangle with vertexes **x**(1), **x**(2) and **x**(3) (three known observations, also in the *s*-dimensional space), and the position of the new observation with respect to its three neighboring points (the vertexes) is found as a weighted sum of their input coordinates subjected to the convexity constraints (i.e., the weights, $\boldsymbol{\alpha}$(i), are bounded to the unit interval and sum up to one). For each such new data point, only the enclosing triangular simplex fulfills the convexity constraints.
The following equations are used for calculating $\boldsymbol{\alpha}$(i) in two dimensions:

$$\alpha_1(i) = \frac{\big(x_2(i) - x_2(2)\big)\big(x_1(3) - x_1(2)\big) + \big(x_1(i) - x_1(2)\big)\big(x_2(2) - x_2(3)\big)}{\big(x_2(1) - x_2(2)\big)\big(x_1(3) - x_1(2)\big) + \big(x_1(1) - x_1(2)\big)\big(x_2(2) - x_2(3)\big)},$$

$$\alpha_2(i) = \frac{\big(x_2(i) - x_2(1)\big)\big(x_1(3) - x_1(1)\big) + \big(x_1(i) - x_1(1)\big)\big(x_2(1) - x_2(3)\big)}{\big(x_2(2) - x_2(1)\big)\big(x_1(3) - x_1(1)\big) + \big(x_1(2) - x_1(1)\big)\big(x_2(1) - x_2(3)\big)}, \quad (1)$$

$$\alpha_3(i) = \frac{\big(x_2(i) - x_2(2)\big)\big(x_1(1) - x_1(2)\big) + \big(x_1(i) - x_1(2)\big)\big(x_2(2) - x_2(1)\big)}{\big(x_2(3) - x_2(2)\big)\big(x_1(1) - x_1(2)\big) + \big(x_1(3) - x_1(2)\big)\big(x_2(2) - x_2(1)\big)}.$$

The weights α_1(i), α_2(i) and α_3(i), with $0 \le \alpha_1(i), \alpha_2(i), \alpha_3(i) \le 1$ and $\alpha_1(i)+\alpha_2(i)+\alpha_3(i)=1$, are the contributions of the known observations (i.e., the vertexes of the triangle**)** to the new data point **x**(i) and determine its position within the enclosing simplex. Any property *y*(i) of the new observation is estimated as a weighted sum of the properties of the known points:

$$\hat{y}(i) = \alpha_1(i)\, y(\mathbf{x}_1) + \alpha_2(i)\, y(\mathbf{x}_2) + \alpha_3(i)\, y(\mathbf{x}_3), \quad (2)$$

where $y(\mathbf{x}_1)$, $y(\mathbf{x}_2)$ and $y(\mathbf{x}_3)$ are the values of the property for the vertexes. Once the weights are calculated, the estimation of any property is immediate. It is also straightforward to generalize Equation 1 and 2 to more than two dimensions.
The special case is for the estimation of a new observation that does not fall inside the convex hull defined by the known data points; hence, not even in any of the constructed simplexes. In this situation Equation 1 does not hold as the convexity constraints cannot be satisfied. Some of these observations are outliers (in a strict sense) but can also be good data located in region of the input space that was unknown when the initial calibration set was defined. It is worthwhile noticing that, in this sense, the main

limitation of the DTR method is its near-absolute lack of extrapolation ability. However, this limitation is not as dramatic as it may appear, because of the simplicity to update both the projection and the tessellation to account for the *outlying* data points
For the estimation of the properties of such observations, several approaches are reported in the literature (Fernández Pierna *et al*, 2002, Fernández Pierna *et al*, 2003, Jin *et al*., 2003a and Jin *et al*., 2003b) and can be adopted. In this applicative study, an estimate for a new observation situated outside the mesh is obtained by projecting it onto the closest simplex, as identified by its centroidal point (i.e., an artificial observation with a set of identical positive weights that sum up to one). The property is then estimated as equal to what calculated for the centroid of such a simplex.

3. Estimation of aromatics in LCO

A total of 91 LCO and Hydrotreated LCO (HDT LCO) samples were acquired for the present study. The HDT LCO samples were obtained at Sartec S.p.A. in a bench-scale pilot unit (Vinci Technologies) operating at various temperatures and pressures, by processing from different LCO feeds provided by the Saras Refinery (Sarroch, Italy). The unit mimics the most typical industrial operations and ensures the range of variation in the total aromatic content and in the distribution of the mono-, di- and tri+ aromatic classes expected in the full-scale case. The NIR spectra of the samples were recorded using a Varian Cary 500 Scan double-beam spectrophotometer in the wavelength range 1600-800 nm (*1 nm* resolution). The aromatic content (*w%*), was determined with the HPLC method EN-12916 using a modular Agilent 1100 Series system with refraction index detection. For the development of the multivariate models, the observations were divided in a calibration and validation set (respectively, 58 and 33 data points) chosen in order to contain samples that provide examples of all products' qualities and spanning the entire range of variation in the concentrations of the aromatics.

3.1. Results and discussion

Based on the data points in the calibration set, a dimensionality reduction with the Principal Components Analysis was performed, as a first step. As discussed in Section 2, the technique of choice uses only the input observations (the NIR spectra, Fig. 1a). After mean centering the observations, PCA was performed and the calibration spectra projected, Fig. 1b. The number of retained principal components is two. The selection was performed by inspection of the eigenvalues of the covariance matrix of the data; the two retained directions account for over *90%* of the total variation in the input space.
Upon projecting the input observations in the calibration set onto the first two principal directions, a Delaunay tessellation was performed, Fig. 1c. Each element of the mesh is a triangular simplex; the set of simplexes is enclosed in a bi-dimensional convex hull.
Subsequently, also the testing observations were mean-centered and projected onto the principal components space (Fig. 1d) where topological regression is performed. Again, it is worthwhile noticing that being the DTR model the same (the weights are calculated only once, from Eq. 1) for all the properties to be estimated, only a single regression model (Eq. 2) is needed; thus, minimizing the calibration and maintenance tasks.
The regression model was found by resolving for the coordinates of the testing observations and, then, calculating the respective properties from the known measurements. As for the testing observations that do not fall inside the convex hull determined by the calibration points (only 5, and all rather close to the boundaries of the convex hull, Fig 1d), the centroids method was used. The estimation results for the independent set of testing observations are depicted in the Fig. 2 and reported in Table 1 for the mono-, di- and tri+ aromatic content. In addition, we are reporting also the

results obtained when estimating the density of the samples. Density (*g cm-3*) was measured according to analytical method ASTM-4052.

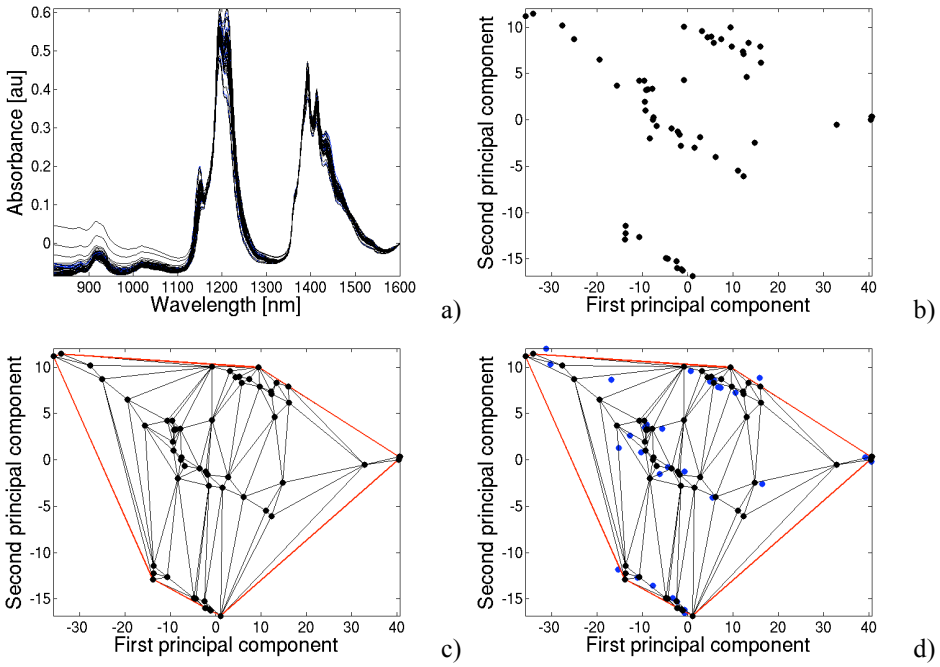


Figure 1. The spectral observations (calibration set in black and testing set in blue) in the original wavelength/absorbance space (a), projected in the principal components space (b), the mesh of simplexes generated by Delaunay triangulation (c) and, with the projected testing data points (d).

The accuracy of the estimates, expressed in terms of Root Mean Squared Error (RMSE) is within the repeatability range of the analytical measurements, and comparable with what obtained using a set of PLSR models independently cross-validated by Leave One Out (for the number of latent factors, 5). Table 1 also presents the results obtained with a set of PLSR models calibrated with 11 factors.

	Mono AH (*w%*)	**Di AH** (*w%*)	**Tri+ AH** (*w%*)	**Density** ($g\ cm^{-3}$)
DTR (2)	2.73	1.23	0.50	0.0021
PLSR (5)	2.06	0.98	0.80	0.0039
PLSR (11)	1.27	0.72	0.52	0.0021

Table 1. The estimation results (as RMSE) on the independent set of testing observations.

4. Concluding remarks

The results using the Delaunay Tessellation and Topological Regression presented here for the calibration of the aromatic content in Light Cycle Oils are at least comparable with the results of standard PLSR. The main advantages of the DTR method are in the simplicity of the calibration and ease to upgrade, but also the fewer components, thus leading to more robust and manageable models.

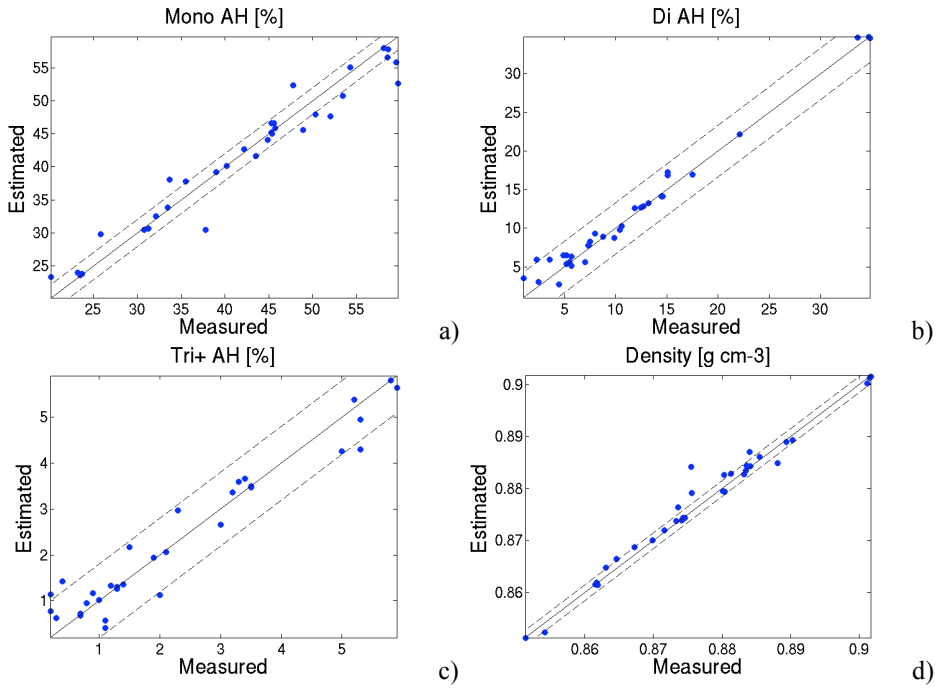


Figure 2. Estimation results (as scatter-plot) and the ASTM repeatability bands.

References

J.A. Fernández Pierna, L. Jin, M. Daszykowski, F. Wahl, D.L. Massart, 2003, A methodology to detect outliers/inliers in prediction with PLS, Chemometrics and Intelligent Laboratory Systems, 68, 17-28.

J.A. Fernández Pierna, F. Wahl, O.E. de Noord, D.L. Massart, 2002, Methods for outlier detection in prediction, Chemometrics and intelligent Laboratory Systems, 63, 27-39.

P. Gemperline et al., 2006, A Practical Guide to Chemometrics: Second Edition, CRC Press – Taylor & Francis.

J. Gudmundsson, M. Hammar, M. van Kreveld, 2002, High order Delaunay triangulations, Computational Geometry, 23, 85-98.

L. Jin, J.A. Fernández Pierna, F. Wahl, P. Dardenne, D.L. Massart, 2003, The Law of Mixtures method for multivariate calibration, Analytica Chimica Acta, 476, 73-84.

L. Jin, J.A., Fernández Pierna, Q. Xu, F. Wahl, O.E. de Noord, C.A. Saby, D.L. Massart, 2003, Delaunay triangulation method for multivariate calibration, Analytica Chimica Acta, 488, 1-14.

L. Jin, Q.S. Xu, J. Smeyers-Verbeke, D.L. Massart, 2005, Updating multivariate calibrations with the Delaunay triangulation method, Applied Spectroscopy, 59, 1125-1135.

L. Jin, Q.S. Xu, J. Smeyers-Verbeke, D.L. Massart, 2006, Updating multivariate calibration with the Delaunay triangulation method: The creation of a new local model, Chemometrics and Intelligent Laboratory Systems, 80, 87-98.

I.T. Jolliffe, 2002, Principal Components Analysis: Second Edition, Springer.

J.A. Lee, M. Verleysen, 2007, Nonlinear Dimensionality Reduction, Springer.

J.J. Jr. Workman, 1999, Review of process and non-invasive near-infrared and infrared spectroscopy: 1993-1999, Applied Spectroscopy Reviews, 34, 1, 1-89.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

An Evolutionary Approach to Derive Adaptive Optimal Control Policy for Chemical Reaction Processes

Yoshiaki Shimizu

Department of Production Systems Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, tenpaku-cho,Toyohashi 441-8580, Japan

Abstract

To obtain a near optimal control policy for real world chemical processes, in this paper, we focused our attention on a new meta-heuristic method termed differential evolution (DE). Compared with conventional approaches characterized by variational logic as well as by the inconveniences of simultaneous optimization methods for differential-algebraic equation, we have shown DE's ability to derive a near optimal solution adaptive to various requirements in practice. Since the algorithm is straightforward and flexible to manage various conditions that other conventional approaches could not cope with effectively, we can cope with high dimensionality and outstanding non-linearity peculiar to chemical process in real world applications. In numerical experiments, we provided three popular reaction processes, applied the proposed method under various meaningful conditions, and validated its adaptability in comparison with other methods.
Keywords: Process system, Optimal control, Differential Evolution, Piece-wise constant, Chemical reaction process.

1. Introduction

By virtue of the advanced progress of meta-heuristic optimization method, we are now ready to derive a near optimal control policy for multi-dimensional and complex real world applications in chemical processes. With such understanding, in this paper, we focus our attention on a new meta-heuristic method termed differential evolution (DE). In comparison with the conventional variational approaches, it can derive a near optimal solution efficiently but very simply without any differential information and additional relations like adjoint equations in Maximum Principle (MP) of Pontryagin (1962). Since it is also free from the complicated usage in simultaneous optimization methods, we can cope with high dimensionality, discontinuity, and outstanding non-linearity peculiar to real world chemical reaction processes.
The aim of this paper is to facilitate a wide application of simulation-based optimization methods like DE for optimal control problems. Finally, its manifold effects are validated through numerical experiments.

2. Problem statements

Optimal control problems , which belong to a class of the off-line dynamic optimal control problem, has been solved traditionally by a certain variational method innovated by the MP. Since singular control and state inequality constraint problems often appear in practical operations of chemical processes, it is almost impossible to solve high dimensional problems for such cases.
Roughly speaking, numerical methods to cope with differential-algebraic equation (DAE) optimization problems are classified into sequential and simultaneous strategies

(Biegler, Cervantes and Wachter, 2002). The former is a control variable parameterization method while the latter uses the difference equation forms both of control variables and state variables. Therefore, the simultaneous approach needs to solve an extremely large-scale non-linear programming problem though it can handle the algebraic constraints by the same manner, i.e., difference equations.

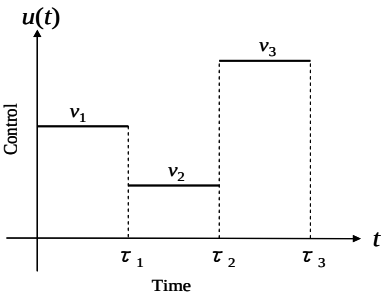


Fig.1 Piece-wise constant control

On the other hand, in the sequential or control variable parameterization approach, the search space is limited within a small region since the differential equations are not transformed into difference equations. However, it is not straightforward to handle the algebraic constraints.

Since, in many chemical processes, some controlling actions are operated discontinuously along time or space co-ordinate, this sequential approach should be highlighted by virtue of recent progress of optimization methods. Among them, piece-wise constant control (PWCC) (Fig.1) is one of the control classes that is easy to implement in reality. Hence, some meta-heuristic methods were successfully applied to obtain the optimal PWCC (Cruz et al., 2003; Kapadi & Gudi, 2004). However, these studies simply applied the methods to search only value of control variables and neglected an interval search and various practical conditions.

Since there has never existed any satisfactory answer to obtaining the optimal control from a practical point of view, it is essential to develop an approach realizing adaptive application to large-scale and complex systems in a flexible and straightforward manner.

2.1. Fixed-time Discrete Problem

Let us consider a system whose dynamics are described by

$$\dot{x} = f(x(t)), u(t), t), \quad x(t_0) = x_o \tag{1}$$

where $x(t)$ denotes an n-dimensional state vector, $u(t)$ an r-dimensional control vector, $f(\cdot)$ an n-dimensional vector function and t a time or time analog. Assuming a scalar control variable $u(t)$, for simplicity of description, the control policy of PWCC is prescribed by a set of pair (v_i, τ_i), where v_i is the value of $u(t)$ at i-th interval and τ_i is the i-th switching point. Then, the continuous problem can be transformed into the following $2N$-1 dimensional optimization problem where v and τ are the vector forms of the sequences of v_i (i=1,2,$\cdots$,N) and τ_i (i=1,2,$\cdots$,N-1), respectively.

$$\text{(p.1)} \ \underset{v,\tau}{Max} \quad J = K(x(\tau_N), \tau_N) \ \text{ subject to } \begin{cases} \dot{x} = f(x(t), v_i), (\tau_{i-1} \le t < \tau_i) \\ x(\tau_{i-1}^+) = x(\tau_{i-1}^-) \\ \underline{u}_i \le v_i \le \overline{u}_i \quad (i = 1, \ldots, N) \end{cases}$$

where $\underline{u}$ and $\overline{u}$ are lower and upper bound vectors of $u(t)$, and N denotes the number of division and controls the fineness of the control policy. In the case of equivalent interval, i.e., $\tau_j - \tau_{j-1} = \tau_{j+1} - \tau_j$, ($i$=1,2,$\cdots$,$N$-1), the dimension of search problem (v-mode search) reduces to just N, apparently. Moreover, it should be noted that $\tau_0 = t_0$ and $\tau_N = t_f$ are implied, and each τ_i ranges from 1 to N_T, satisfying the condition that $\tau_j > \tau_{j-1}, (j = 1, \ldots, N)$ where $N_T = (t_f - t_0)/H$ (H is an increment width of the numerical solution of the differential equation).

2.2. Free-time Discrete Problem

A typical example of the free-time problem is known as the time optimal control problem described in the following.

$$\text{(p.2) Min } J=t_f \text{ subject to } \begin{cases} Eq.(1) \text{ and } \underline{u} < u(t) \le \bar{u} \\ x(t_f) = x_f \end{cases}$$

Here let us note the additional boundary condition of the state variables at the final time. The present simulation-based approach is almost impossible to always satisfy this condition directly. To cope with this problem, we use the penalty function method by augmenting the boundary condition into the objective function as follows.

$$\text{(p.3)} \qquad t_f + P\left\|x(t_f) - x_f\right\| / \left\|x_f - x_0\right\|$$

where || || denote a certain norm of vector, and P a penalty coefficient whose value should be increased adaptively according to the following formula.

$$P := \begin{cases} P + \Delta p & \text{if} \quad G^k < G^{k-1} \qquad (\textit{coming up the target}) \\ P & \text{if} \quad G^k \geqq G^{k-1} \qquad (\textit{leaving the target}) \end{cases} \tag{2}$$

where G^k denotes the value of $\left\|x(t_f) - x_f\right\| / \left\|x_f - x_0\right\|$ at the k-th iteration, and Δp a certain incremental amount.

3. Evolutionary Optimization to Derive Adjustable Optimal Control

3.1. Application under various control policies

For solving (p.1) and (p.2), it is straightforward to apply DE in the following manner. The genetic code of DE is made of the sequence of $v_i \ \ (i = 1, \ldots, N)$ for the equal interval problem, and in addition to it, the set of the interval length, i.e., $z_i = (\tau_i - \tau_{i-1}) / H, \ \ (i = 1, \ldots, N)$ for the unequal interval problem. Once these values are prescribed, we can evaluate the objective function directly after solving the differential equations numerically. Then, we can utilize simply this value as the fitness of the DE. Since the interval length z_i should be integer, rounding-off is carried out to satisfy this condition. The initial population is generated randomly by scattering the value between the lower and upper bounds of the decision variables.

Generally speaking, when N is small, the control often changes rapidly between the consecutive sequences. Hence, it likely happens that actual control action cannot catch up such a large change instantly. When this is the case, it is possible to impose the following condition.

$$\left|v_{i+1} - v_i\right| / T_{setup} \le v_{speed}, \ (i = 1, \ldots, N-1) \tag{3}$$

where T_{setup} and v_{speed} denote a transient interval and an admissible switching speed of the control policy, respectively. Similarly, when the duration interval of each v_i is prescribed, we can design the strategy by imposing conditions such that $\tau_j - \tau_{j-1} \ge D$, ($i$=1,2,…,$N$) for the unequal interval case and $(t_f - t_0) / N \ge D$, for the equal interval case. Here D denotes the minimum duration interval.

On the other hand, when N becomes large, this often causes the chattering action within the bounded control region in numerical solution. To prevent such chattering, the condition $\left|v_{i+1} - v_i\right| \le \varepsilon, \ (i = 1, \ldots, N-1)$ is available to smooth the control policy, where ε is an appropriate small positive value.

3.2. State inequality constrained problem

Chemical processes are often controlled by manipulating flow rates and are restricted within a certain state space prescribed by temperature, pressure, and concentrations of chemical species, for examples. The latter condition requires us to augment the additional algebraic conditions like $g(x(t),u(t)) \leqq 0, \quad t_0 \leq t \leq t_f$ into the original model. Instead of the tedious procedures discussed previously (Takamatsu and Shimizu, 1978a), the present approach can manage this problem much more simply even for this DAE system. That is, it is enough to degrade the fitness value $Fit(v,\tau)$ of the DE according to the degree of violation of the state-inequality condition in the course of the transition as $Fit'(v,\tau) = Fit(v,\tau) + P \cdot \sum_{i=1}^{N} \delta_i(g)$, where P denotes the penalty coefficient and the psedo-heaviside function δ_i takes the value only when anyone of the state inequality constraints are violated at the *i*-th interval.

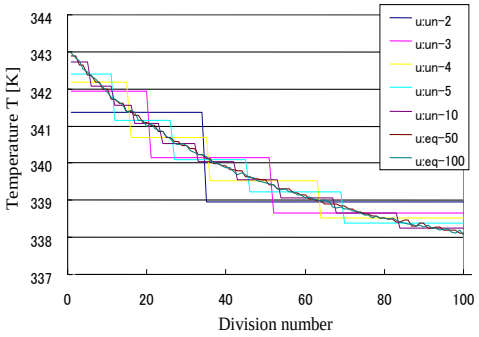


Fig.2 Control policies for various *N*
(Only *v*-mode search when *N*=50 and 100)

3.3. Insensitive control against parameter deviations

Describing explicitly a parameter vector *p* of the system model, let us note the existence of uncertainty in the model. Then we tried to obtain an insensitive control policy against the uncertain parameters. For this purpose, we derive the sensitivity equations with respect to *p*. Adjusting a certain norm of the sensitivity in an appropriate manner, e.g., $\|(\partial K/\partial x)(\partial x(t_f)/\partial p)\|$ for the objective function value, $\|(\partial x(t)/\partial p)\|$ for each state variable, and $\|(\partial g/\partial x)(\partial x(t)/\partial p)\|$ for the state-inequality constraints, for examples, we can give a suitable answer for the problem under consideration. Finally, this insensitivity goal is augmented to the original objective function J as $J' = wJ + (1-w)S$, where S represents the sensitivity term given by the appropriate norm exemplified in the above, and w is the weighting factor to adjust the trade-off between the two goals, i.e., J and S.

4. Numerical Experiments

In the following numerical experiments, DE/rand/1/bin (Storn & Price, 1997) is applied. After some preliminary numerical experiments, we set the tuning parameters as F=0.8, P_c=0.5, N_p=3(2N-1) for the unequal interval problem while 3N for the equal one and the sopping generation=$20N_p^2$, where F, P_c and N_p denote the scale factor, the cross-over rate and the population size, respectively. Under these conditions, we observed a good profile of convergence all in the following numerical experiments.

4.1. Temperature control of completely mixed batch reactor (A)

For the reaction scheme like $A \xrightarrow{k_1} B \xrightarrow{k_2} C$, the optimal temperature control problem that maximizes the production of the desired product *B* is well known and is formulated as follows.

$$\underset{T}{Max} \quad J = x_2(t_f) \quad subject\ to \begin{cases} dx_1/dt = -k_1 x_1, & x_1(0) = x_{10} \\ dx_2/dt = k_1 x_2 - k_2 x_2, & x_2(0) = x_{20} \end{cases}$$

where x_1 and x_2 are the concentrations of components *A* and *B*, respectively, t_f is the batch time, and k_1 and k_2 are the reaction rate constants of Arrhenius type.
As shown in Fig.2, the DE derived the solutions efficiently from N=2 up to the minimum grain interval problem, i.e., N=N_T=100. The only *v*-mode search has been done for N=50 and 100. When N=100, the result is almost equivalent to that of the variational method.

4.2. Catalyst blending problem (B)

The catalyst blending problem studied originally by Hofer (1976) is known to involve a singular control portion. With the prior knowledge that the optimal control is composed of three segment, i.e., upper bound, constant singular, and lower bound segments, a sensitivity method (Takamatsu and Y. Shimizu, 1978b) can derive the result very quickly by virtue of the differential information and small problem size. However, without such prior information, the DE outperformed this sensitivity method. Figure 3 depicts the results obtained by the proposed method. They are nearly identical to the profile of the sensitivity method when N=3 which is, in turn, almost same as the optimal control.

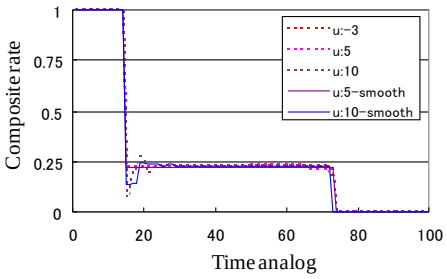


Fig.3 Result of control policy

4.3. The flow rate control under state inequality constraint (C)

4.3.1. Fixed-time Solution

For the first-order reversible exothermic chemical reaction (Siebenthal and Aris, 1964), let us suppose that the temperature in completely mixed batch reactor with cooling jacket must not exceed the upper limit $\overline{T}$. The optimal cooling control problem that will achieve the maximum production of *B* is described with dimensionless variables as follows (Takamatsu and Shimizu, 1978a).

$$Max \quad J = x_1(1) \qquad \text{subject to}$$

$$\begin{cases} \dot{x}_1 = p_1(1-x_1)\exp(p_2/x_2) - p_3 x_1 \exp(p_4/x_2), & x_1(0) = x_{10} \\ \dot{x}_2 = p_5[p_1(1-x_1)\exp(p_2/x_2) - p_3 x_1 \exp(p_4/x_2)] - p_6 u, & x_2(0) = x_{20} \\ \underline{u} \le u(\tau) \le \overline{u}, \quad x_2(\tau) \le \overline{x}_2 \quad (0 \le \tau \le 1) \end{cases}$$

where the kinetic rate constants are of the Arrhenius type. Due to the state-inequality constraint over the time horizon and the linearity of the control variable, this problem is known to be very difficult to solve.
Numerical results revealed that the good performance of the simplex method of Nelder and Mead is limited only to the small N, and the unequal interval search with moderate N (= 5) is enough to attain a satisfactory performance level (the gap to the optimal value is less than 0.3%). The chattering becomes heavily for N greater than 50 to imitate the optimal trajectory.

4.3.2. Insensitive control against parameter deviations

It is likely that the fouling degrades the performance of the cooling jacket, and the reactor temperature may exceed the limit. To prevent such defect a priori, an insensitive control policy against the parameter deviations (fouling) is more desirable to keep the safety or the long life of the reactor. To realize this aspect, the sensitivity

equations with respect to p_6 are derived. The present strategy for the insensitive control is to take an action when the state trajectory approaches the limit beyond a certain threshold value Δx_2 (= $0.01\overline{T}/T_0$). That is, $\|(\partial g/\partial x)(\partial x(t)/\partial p)\|$ in section 3.3 is defined as $\left\{\int_{t_2}^{t_1}(\partial x_2(t)/\partial p_6)^2\,dt\right\}^{1/2}$ where t_1 and t_2 prescribe the beginning and the ending points of the interval featuring the situation, $\bar{x}_2 - x_2(t) \le \Delta x_2$. By considering the insensitivity, the state trajectory tends to keep inside of the upper limit of temperature as illustrated Fig.4. Accordingly, the violation does not occur even with the 20% parameter variation (p_6' = 0.8 p_6). Moreover, it only causes a small degradation of the system performance both at the nominal (ordinary=0.9153 vs. insensitive=0.9143) and at the deviated (0.9074 (infeasible) vs. 0.9065).

4.3.3. Free Time Solution

The objective function of this problem will be changed as $Min \quad \tau_n \quad (\tau_n \le \tau_N$=100). Additionally, we need to add the terminal condition like $x(\tau_n) = x_f$. For three problems with different terminal conditions shown in Table 1, the proposed approach solved all cases within the square root error 0.0001 of the terminal condition. Figure 5 depicts the control policy for Run 1 when *N*=10.

Table 1 Numerical results for free time problem

Run	x_f	τ_n [-]
1	(0.88, 0.92)	63
2	(0.9, 0.9)	83
3	(0.9, 0.95)	86

(*N*=10; unequal interval policy)

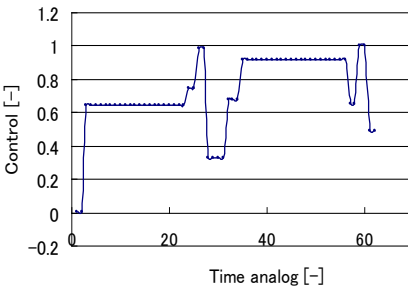


Fig.5 Profile of control for Run 1

5. Conclusion

To derive the optimize control policy for multi-dimensional and complex systems aimed at real world applications, we focus our attention on a new meta-heuristic method termed DE. Compared with the conventional approaches, we have shown DE's ability to derive optimal PWCC efficiently under various conditions popular in real world applications. In comparison with other methods through various numerical experiments, we have confirmed the great promise of DE as a practical solution method for optimal control problems.

References

L. T. Biegler, A. M. Cervantes and A. Wachter, 2002, Chem. Eng. Sci., 57, 575-593

I.L. L.Cruz, , L.G. Willigenburg and G. Straten, 2003, Applied Soft Computing, **3**, 97-122

M. D. Kapadi and R. D. Gudi, 2004, Process Biochemistry, **39**, 1709-1721

R. Storn and K. Price, 1997, Journal of Global Optimization, 11, 341-359

T. Takamatsu, Y. Shimizu and T. Murata, 1978a, J. of Chem. Eng. of Japan, 11, 59-66

T. Takamatsu and Y. Shimizu, 1978b, J. of Chem. Eng. of Japan, 11, 221-226

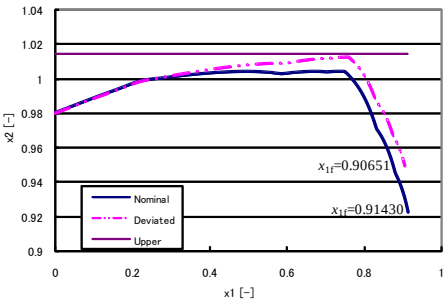


Fig.4 Nominal and deviated trajectories

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Novel Technique to Estimate Valve Stiction Based on Pattern Recognition

"Marcelo Farenzena, Jorge O. Trierweiler"

"Department of Chemical Engineering, Federal University of Rio Grande do Sul," Porto Alegre, RS, 90040-040, Brazil

Abstract

This work proposes a novel method to estimate valve stiction parameters that requires only normal operating data (process variable (PV) and controller output (OP)). The technique is based on pattern recognition, where a neural network, called Stiction Inference Model (SIM), predicts the stiction parameters. Thus, the quantification is very fast and robust, what allows the application in the industrial scenario, where a large number of loops should be analyzed. The SIM is described, as well as the procedure to its construction: the dataset construction, the variable selection, and the network training. The SIM obtained has a compact size (10 neurons) and a good performance for both parameters (R^2=0.99). The proposed method was applied in a set of valves, providing reliable results.

Keywords: Process control; Process monitoring; Valve stiction; Hysteresis.

1. Introduction

The diagnostics of the "valve health" is important to eliminate plant oscillations, allowing the process to achieve a more profitable operating point. The villain in chemical plants is high static friction (or stiction). Reports show that 20-30% of all control loops have bad performance because of stiction (Bialkowski, 1993). That is why preventive maintenance is essential to avoid off-spec products and waste of energy or raw materials. This scenario demands an automatic technique to detect and quantify valve.

In the last ten years, many works and methods have been proposed to diagnose and solve this valve disorder. A first group of works aimed at diagnosing stiction automatically, using only process variable (PV) and controller output (OP) (He et al., 2007; Horch, 1999; Rossi and Scali, 2005; Singhal and Salsbury, 2005; Yamashita, 2006). Some works have proposed specific stiction models for diaphragm type valves. A good survey about stiction models was recently written by Garcia (2008).

Furthermore, some authors have proposed to compute stiction parameters in real time. Choudhury et al. (2006) proposed two methods to quantify stiction parameters, based on ellipse fitting and *c-clustering*, using a one parameter empirical model. Subsequently, some authors have proposed to quantify stiction parameters using a more reliable model, with two parameters. Choudhury *et al.* (2008) have proposed a method based on optimization and grid search. Jelali (2008) has introduced one methodology to compute both stiction parameters, based on least-squares and global search algorithms. This method is computationally expensive what restrict its on-line application.

The main contribution of this work is to overcome this limitation, proposing a method to compute the two stiction parameters, where the stiction parameters are estimated with a very fast computational time. The technique is based on pattern recognition, where a neural network, called Stiction Inference Model (SIM), predicts the stiction parameters.

Thus, the quantification is very fast and robust. Moreover, the method needs only normal operating data, i.e. based on process variable (PV) and controller output (OP) patterns. No information about the valve stem is necessary. Based on these premises, the method can be applied industrially, where a large number of valves is analyzed. The procedure to build the SIM has used the stepwise regression to select the input variables and a neural network to capture the desired behavior input-output The SIM accuracy is slightly poorer – with less than 30% relative error when compared to the least-squares and global search algorithm. However, our objective is to propose a fast algorithm to allow its large-scale application.
The paper is segmented as follows: in section 2 the stiction phenomenon is defined and modeled. In the next section, Stiction Inference Model (SIM) is defined, as well as the procedure to build it. Next, in the section 4, the SIM is applied in a set of case studies. The paper ends with the concluding remarks.

2. Stiction: definition and model

The phenomenon of stiction can be summarized as follows (Ruel, 2000): "*Stiction is the resistance to the start of motion, usually measured as the difference between the driving values required to overcome static friction upscale and down scale. The word stiction is a combination of the words stick and friction, created to emphasize the difference between static and dynamic friction. Stiction exists when the static (starting) friction exceeds the dynamic (moving) friction inside the valve. Stiction describes the valve's stem (or shaft) sticking when small changes are attempted. Friction of a moving object is less than when it is stationary. Stiction can keep the stem from moving for small control input changes, and then the stem moves when there is enough force to free it. The result of stiction is that the force required to get the stem to move is more than is required to go to the desired stem position. In presence of stiction, the movement is jumpy.*"

2.1. Stiction model

A sticky valve has in the phase plot (Valve output (MV) versus Controller output (OP)), shown in Fig. 1, four components: deadband (*DB*), stickband (*SB*), slip jump (*J*) and moving phase (*MP*). The method assumes that the process and controller have linear behavior, while the sticky valve inserts in the loop nonlinear behavior.
When the valve changes the direction (A), the valve becomes sticky. The controller should overcome the deadband (AB) plus stickband (BC), and then the valve jumps to a new position (D). Next, the valve starts moving, until its direction changes again or the valve comes to rest, between D and E.
The stiction model consists of these two parameters: S, called staticband, (deadband+stickband) and J (slipJump). The deadband and stickband represent the behavior of the valve when it is not moving, although the input of the valve keeps changing. The magnitude of staticband and slipjump is essential to determine the limit cycle amplitude and frequency.

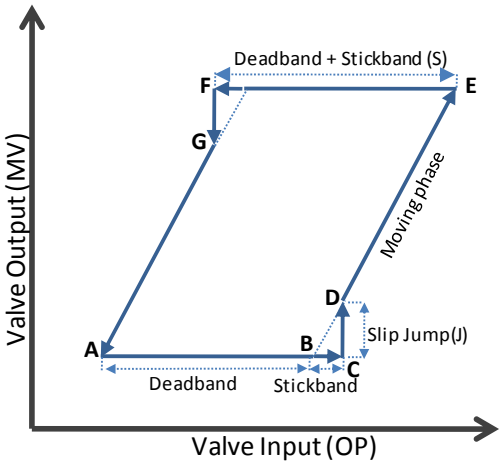


Fig. 1. Relation between controller output (OP) and valve position (MV) for a sticky valve.

The stiction model used in this work is proposed by Kano (2004), which is an extension of Choudhury's method, where stiction is modeled using two parameters. Their main advantage is that it can deal with both stochastic and deterministic signals.

2.2. Stiction evaluation

In the literature, two methods are available to compute the stiction parameters, using a two parameter model. The method proposed by Jelali (2008), a two step procedure is introduced. In the first step the stiction parameters are quantified using pattern search methods or genetic algorithms (GA). Following, the low-order linear plant model is identified, using a least-squares estimator. A second work of Choudhury *et al.* (2008) proposes a method based on a grid search. Initially, a grid using several different values of *J* and *S* is built and then based on the process output, the plant model is identified. Based on the mean square error (MSE) between predicted process output and actual output in each grid point, the stiction parameters are estimated.

3. Stiction Inferential Model (SIM)

The Stiction Inference Model (SIM), proposed in this work, can be defined as an inference model to evaluate the Staticband (*S)* and the SlipJump (*J*), whose inputs are parameters obtained only from normal operating data.

The procedure to construct SIM is analogous to the procedure to build a softsensor using neural networks. The steps are following listed and they will be explained in the next sections.

1. Dataset design: select plant models, disturbances, controller performance, and stiction parameters;
2. Generate the dataset and compute candidate inputs and outputs parameters;
3. Divide the dataset into test and training datasets;
4. Select the input variables for each stiction parameter;
5. Train the neural networks using, with variable number of neurons in the hidden layer;
6. Evaluate network performance for both datasets;

Each step will be described in details in the following subsections.

3.1. Dataset design

The dataset was designed based on a SISO loop, whose scheme is shown in Fig. 2.

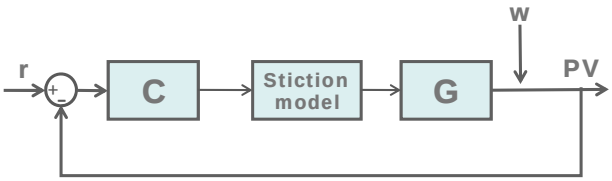


Fig. 2: SISO loop scheme

Where *C* is the PI controller, *G* is the first order plus time delay plant. The input *w* is white noise and *r* is the reference. In this work, we propose only a stiction inference model for first order plus time delay models. However it could be extended to higher order models or integrators. Tab 1 shows the variable parameters to generate the dataset and the interval for each parameter.

Tab. 1: Variable parameters used in the SIM data generation

Param.	Description	Interval
$\mathbf{K_p}$	Controller gain	[0.1: 0.2: 1.1, 2: 1: 5]
$\boldsymbol{\tau_I}$	Controller integral constant	[0.3τ: 0.3τ: 3τ]
$\boldsymbol{\tau}$	Process time constant	[5: 5: 20, 30: 10: 100]
S	Static band	[0.5: 0.5: 5]
J	Slip Jump	[0.5: 0.5: 5]
$\boldsymbol{\sigma^2(w)}$	White noise variance	[100 10 1]

The number of different scenarios in the dataset is 360,000. Subsequently, the complete dataset was randomly divided in two datasets - training dataset = 80% of the points and test dataset = 20% of the points

3.2. Input variables

The candidate variables to build the inference model and the procedure to select the best subset of them to describe the stiction are subsequently listed:

1. Δ_{PV}: Difference between the maximum and minimum value in PV;
2. Δ_{OP}: Difference between the maximum and minimum value in OP;
3. Ratio between Δ_{PV} and Δ_{OP}.
4. E_W: integral of square error of difference between PV and the best wave-shape curve interpolation (PV_W) (see (Rossi and Scali, 2005));
5. E_T: integral of square error of difference between PV and the best triangular curve interpolation (PV_T);
6. ZC: Number of zero-crossings in the zero-mean data;
7. ZC_{ACV}: Number of zero-crossings in the autocovariance function (Thornhill et al., 2003).

The variables were selected using stepwise regression (Rawlings, 1988). The inputs selected to estimate *J* are the same as the inputs to predict *S:*

- Δ_{PV}
- Ratio between Δ_{PV} and Δ_{OP}
- ZC_{ACV}

3.3. Neural-network training

To interpolate the training dataset, we use a feedforward backpropagation neural network (Haykin, 1999), from Matlab 5.3 (R11, Neural Network toolbox ver. 3.0.1). In the hidden layer, hyperbolic tangent sigmoid neurons were used, while in the output layer, linear neurons are used. In the training step, the layer performance is measured using mean square error. The training function used was Levenberg-Marquardt backpropagation (*trainlm*) and the learning function was Gradient descent with momentum weight and bias (*learngdm*).

We try networks with different number of neurons in the hidden layer to select the one that better predict the stiction parameters, with the minimum number of neurons. Initially, the SIM for Staticband (S) was designed. Tab. 2 shows the networks performance, with different number of neurons, for the training and test datasets.

Tab. 2: networks performance – output variable: Staticband (S)

Neurons	R^2 Training dataset	R^2 Test dataset
5	0.99	0.99
10	0.99	0.99
20	0.99	0.99
50	0.99	0.99

Based on Tab. 2, the best net configuration has 5 neurons in the hidden layer. Subsequently, the model for Slipjump (*J*) was designed. Tab. 3 shows the networks performance, with different number of neurons, for the training and test datasets.

Tab. 3: networks performance – output variable: Slipjump (J)

Neurons	R^2 Training dataset	R^2 Test dataset
5	0.97	0.97
10	0.98	0.98
20	0.98	0.98
50	0.98	0.98

Again, the best configuration, according to Tab 3, has 10 neurons in the hidden layer.

The final SIM structure for both parameters is:

- Neural network model with 2 layers;
- 1st layer (hidden): 10 hyperbolic tangent sigmoid neurons;
- 2nd layer (external): 1 linear neuron.

4. Case studies

This section illustrates the SIM application using simulation datasets. The comparison between the predicted and real values for stiction parameters will be made for a set of 6 valves, which are not used to train or test the SIM neural networks. Tab. 4 shows the results for original *J* and *S*, and its predicted values (J_P and S_P) by the SIM.

Tab. 4 clear shows the good performance of SIM, where the relative error for all controllers was lower than 30%. Using a Pentium D, 1 GB ram, the computation time for all cases was lower than 1 second. Using the same computer, to predict valve

stiction to 100,000 valves took 1 second, what allows the real-time and large-scale application of the proposed methodology.

Tab. 4: SIM prediction for 6 different sticky valves

Cont.	S	S_p	J	J_p
Cont1	10	13	9	9,9
Cont2	20	23	12	11
Cont3	20	20	15	15
Cont4	20	27	40	39
Cont5	30	33	30	30
Cont6	40	37	60	53

5. Conclusions

In this work, a novel algorithm to evaluate valve stiction, called Stiction Inference Model (SIM), has been proposed. It allows the computation of both stiction parameters (staticband *S* and slipjump *J*) using only routine operating data. The SIM is based on neural networks. Comparing the proposed method with to the least-squares and global search algorithm, the SIM accuracy is slightly poorer. However, its large-scale application is viable, because of its computational time.

References

W.L. Bialkowski,1993, Dreams versus reality: A view from both sides of the gap, Pulp and Paper Canada, 94, 11, 19-27.

M.A.A.S. Choudhury, M. Jain, and S.L. Shah, 2006, Detection and quantification of valve stiction, Proceedings of the American Control Conference, Vol. 2006, pp. 2097-2106.

M.A.A.S. Choudhury, M. Jain, and S.L. Shah,2008, Stiction - definition, modelling, detection and quantification, Journal of Process Control, 18, 3-4, 232-243.

C. Garcia,2008, Comparison of friction models applied to a control valve, Control Engineering Practice, 16, 10, 1231-1243.

S.S. Haykin, 1999, Neural networks : a comprehensive foundation, Prentice Hall, Upper Saddle River, N.J.

Q.P. He, J. Wang, M. Pottmann, and S.J. Qin,2007, A curve fitting method for detecting valve stiction in oscillating control loops, Industrial and Engineering Chemistry Research, 46, 13, 4549-4560.

A. Horch,1999, A simple method for detection of stiction in control valves, Control Engineering Practice, 7, 10, 1221-1231.

M. Jelali,2008, Estimation of valve stiction in control loops using separable least-squares and global search algorithms, Journal of Process Control, 18, 7-8, 632-642.

M. Kano, H. Maruta, H. Kugemoto, and K. Shimizu, 2004, Practical model and detection algorithm for valve stiction, 7th IFAC DYCOPS,IFAC, Boston, USA.

J.O. Rawlings, 1988, Applied regression analysis : a research tool, Wadsworth & Brooks/Cole Advanced Books & Software, Pacific Grove, Calif.

M. Rossi, and C. Scali,2005, A comparison of techniques for automatic detection of stiction: simulation and application to industrial data, Journal of Process Control, 15, 5, 505-514.

M. Ruel,2000, Stiction: The hidden menace, Control Magazine.

A. Singhal, and T.I. Salsbury,2005, A simple method for detecting valve stiction in oscillating control loops, Journal of Process Control, 15, 4, 371-382.

N.F. Thornhill, B. Huang, and H. Zhang,2003, Detection of multiple oscillations in control loops, Journal of Process Control, 13, 1, 91-100.

Y. Yamashita,2006, An automatic method for detection of valve stiction in process control loops, Control Engineering Practice, 14, 5, 503-510.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Two-layer planning and scheduling of batch production processes under uncertainty

Martin Hüfner, Thomas Tometzki, Sebastian Engell

Process Dynamics and Operations Group, Department of Biochemical and Chemical Engineering, TU Dortmund, 44221 Dortmund, Germany

Abstract

This paper proposes a new combined algorithm to solve planning and scheduling problems of batch production processes under uncertainty. For the planning problem under uncertainty, a two-stage stochastic integer program is formulated and solved by a stage-decomposition algorithm. To evaluate the feasibility of a production plan, a priced timed automata model is set up for the scheduling problem of the next planning period and solved using reachability analysis. If the targets from the planning layer cannot be met, a penalty term is returned to the planning layer which is used to re-evaluate the candidate production plans. The two-layer concept is applied to a multi-product batch plant demonstrator.

Keywords: evolutionary algorithms, MILP, timed automata, multi-product batch plants

1. Introduction

For the economic operation of multi-product batch plants, high-quality *production plans* that reflect the uncertainties in market and technical parameters as well as feasible or optimal or at least feasible *operation schedules* are needed. *Production planning* is understood as the decision *which* products have to be produced in the planning periods with the limited set of plant equipment. *Scheduling* concerns the decisions *how* the scarce resources are allocated to the operations over time to produce the planned products in order to meet the targets. The combination of long-term production planning and of scheduling ensures the meeting of the economic objectives and timely completion of the tasks associated with the production plans. On the other hand, it is not necessary to determine precise schedules over long horizons because the uncertainties render such schedules obsolete after short periods of time.

The decision structure in production planning with uncertainties is reflected well by a mixed-integer recourse model with a finite number of scenarios in the form of a *two-stage stochastic integer program* (2-SIP) [1]. The recourse property reflects the ability to compensate the effects of a set of *here-and-now* (first-stage) decisions which have to be made under uncertainty by a second set of *recourse* (second-stage) decisions which can be made after the uncertainty has realized. This type of models is particularly appropriate for moving horizon production planning where the movement of the horizon is accompanied by the realization of uncertainties and their compensation.

In [2], a similar temporal decomposition approach to the coupled production planning and scheduling problem is applied, where first the time period and the products that are considered in short-term scheduling are determined, and then the scheduling problem is solved for the immediate future. In [3], a simultaneous solution of production planning and scheduling problem is addressed through a hierarchical framework. The planning problem aggregates orders in the planning period and considers uncer-

tainty utilizing multi-stage stochastic programming. Using a moving horizon strategy, the production targets for the current stage are provided to the scheduling layer, which is solved using a continuous-time MILP formulation.

Optimal scheduling of multi-product batch plants (resource allocation, sequencing and timing) has been in the focus of research for many years now. A promising new approach to scheduling problems is to model them by *timed automata* (TA) and to perform optimization by *reachability analysis* [4, 5]. Modeling is completely graphical and the problem can be represented in a modular and transparent fashion with independent small models for recipes, resources, etc. that are connected by synchronization labels and shared variables and can be composed automatically.

2. Two-layer algorithm for production planning and scheduling

In this contribution, the coupled problem of production planning and plant scheduling is solved by a two-layer algorithm, where the planning (upper-layer) problem under uncertainty is represented as a *two-stage stochastic integer program* (2-SIP) and the scheduling (lower layer) problem is solved using a *timed automata* (TA) approach. The algorithm for the upper layer problem (see Section 4) interacts with the algorithm for the lower layer problem (see Section 5) in a feedback loop.

The upper layer algorithm iteratively generates a set of production plans. After every iteration, the upper layer passes the first stage solution with the best objective value to the lower layer. The lower layer algorithm computes an optimal or at least a feasible schedule for the first-stage periods and returns a penalty value to the upper layer if the due dates are not met. The penalty value is added to the planning objective value of the passed solution. Thus another production plan from the set may have a better objective. Then the scheduling is repeated for the new best solution and the resulting penalty is passed to the upper layer. The procedure iterates until the best solution from the set has been found.

3. Multi-product batch plant case study

In order to illustrate the approach, a small demonstrator batch plant [6] is considered (see Fig. 1). The end-products blue (**B)** and green (**G)** are produced from three raw materials, yellow (**Y)**, red (**R)** and white (**W)**. Two batches of materials **Y** and **W** react to produce one batch of product **B**, similarly two batches of **R** and **W** react to produce one batch of product **G**. The plant consists of 3 stages. The first stage consists of three buffer tanks which are used to store the raw materials **Y, R** and **W**. The second stage consists of three reactors that perform the reaction process to produce the end-products. Each reactor can be filled from each raw material buffer tank in the first stage; and it is possible to produce either product **B** or product **G** in each reactor. After processing the materials, a reactor contains one batch of product. The third stage consists of two buffer tanks which are used to store the end-products **B** and **G** exclusively. Once started operations must be finished without any interruption (non-preemptive scheduling). The production of one batch of a recipe (**B** or **G)** consists of 6 operations and involves timing

Fig. 1: Demonstrator batch plant

constraints between individual operations. After the materials are processed by the reactors, the end-products must be drained into the buffer tanks in the third stage immediately, imposing a zero-wait constraint between the operations.

The planning problem is based on a simplified discrete time multi-period model which represents capacity constraints and material balances. The buffer tanks in the first stage contain at most two batches of the raw materials. Each of the buffer tanks has a maximum capacity to store three batches of the respective products. Raw materials can be stored temporarily in a depot. The planning problem is subject to uncertainties in the product demands, revenues, production costs, and possible breakdowns of the reactors. The planning decisions which have to be made are the optimal choices of the purchases of raw materials and the batch numbers of products. The profit that has to be maximized is calculated from sales revenues, production costs, storage costs, and penalties for lateness. The scheduling problem is based on a continuous-time formulation containing all important plant characteristics such as non-intermediate storage policies, exclusive assignment of resources to tasks, intermediate release dates for the raw materials and intermediate due dates for the end-products. The purchased raw materials are released at time points which correspond to the beginning of the discrete periods. For each batch of a product a due date at the end of a period has to be met. The objective of the scheduling problem is to schedule the production orders with no delay.

4. Upper layer algorithm: Stage decomposition based hybrid approach

4.1. Two-stage mixed-integer program

The production planning problem under uncertainty is modeled as a 2-SIP. The uncertainty in the planning problem has Ω=32 realizations that are modeled by a discrete set of scenarios ω = 1,…, Ω. In a 2-SIP, the decisions are divided into the first-stage decisions **x** which have to be taken before the uncertainty is disclosed and second-stage decisions $\mathbf{y}_\omega$, which have to be taken after the uncertainty is realized.

$$\min_{\mathbf{x},\mathbf{y}_1,\ldots,\mathbf{y}_\Omega} \mathbf{c}^T\mathbf{x} + \sum_{\omega=1}^{\Omega} \pi_\omega \mathbf{q}_\omega^T \mathbf{y}_\omega \tag{1}$$

$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \le \mathbf{b} \tag{2}$$

$$\mathbf{W}_\omega \mathbf{y}_\omega \le \mathbf{h}_\omega - \mathbf{T}_\omega \mathbf{x} \tag{3}$$

$$\mathbf{x} \in \mathrm{X},\ \mathbf{y}_\omega \in \mathrm{Y},\ \forall \omega = 1,\ldots,\Omega.$$

The objective of the problem (1) consists of the first-stage costs and the expected value of the second stage costs. The costs are calculated as linear functions of the first-stage variables **x** and the second-stage variables $\mathbf{y}_\omega$ with vectors of cost parameters **c** and $\mathbf{q}_\omega$. The two-stage model consists of inequality constraints on the first-stage decisions (2) and on the first- and second-stage decisions (3). The finite sets X and Y may contain integrality requirements.

4.2. Stage decomposition approach

The main idea of stage decomposition [7] is to remove the ties between the second-stage scenario subproblems by fixing the first-stage decisions **x**. The scenario subproblems are of significantly smaller size than the full two-stage problem. The master problem is a function of the vector of first-stage variables **x** only:

$$\min_{\mathbf{x}} f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} + \sum_{\omega=1}^{\Omega} \pi_\omega Q_\omega(\mathbf{x}) \qquad \text{s.t.} \qquad \mathbf{A}\mathbf{x} \le \mathbf{b}, \mathbf{x} \in \mathrm{X} \tag{4}$$

The evaluation of the second-stage value function $Q_\omega(\mathbf{x})$ for a given $\mathbf{x}$ requires the solution of Ω independent MILP subproblems over the second-stage variables y_ω:

$$Q_\omega(\mathbf{x}) = \min_{\mathbf{y}_\omega} \mathbf{q}_\omega^T \mathbf{y}_\omega \quad \text{s.t.} \quad \mathbf{W}_\omega \mathbf{y}_\omega \le \mathbf{h}_\omega - \mathbf{T}_\omega \mathbf{x}, \mathbf{y}_\omega \in Y, \forall \omega = 1, \ldots, \Omega. \tag{5}$$

The constraints of the master problem (2) are scenario independent, while the parameters of the second-stage constraints in (3) may vary from scenario to scenario. The vector of the first-stage variables $\mathbf{x}$ appears as a vector of fixed parameters in the constraints of the second-stage scenario problems. The main algorithmic idea is to address the master problem given by (4) by an *evolutionary algorithm*. To evaluate f(**x**), the Ω subproblems given by (5) are solved independently by a *MILP solver*.

4.3. Evolutionary Algorithm

The master problem (4) is tackled by an adapted (integer) *(μ ,κ, λ)-evolutionary algorithm*. Each individual $\mathbf{x} = (x_1, \ldots, x_n)$ of the population represents the choice of the numbers of batches of raw materials and batches of products for all first-stage periods. A population of μ individuals is initialized randomly within the bounds of the box-constrained first-stage feasible decision space $\mathbf{x}^{min} \le \mathbf{x} \le \mathbf{x}^{max}$. The fitness values of the initial population are calculated by the objective function f(**x**) of the master problem.

After the evaluation of the initial population, λ offspring individuals are generated by λ-fold application of the discrete recombination operator. The recombination chooses two individuals out of the μ parents with equal probability. Each x_i of the offspring is the corresponding value from the first or from the second parent with equal probability. After the recombination, the mutation operator mutates the λ offspring in two steps:

In the first step, the object parameters for the amounts of products are mutated. The reason for this is that the amount of products cannot exceed the amount of available raw materials. Each object parameter for the amount of a product x_i is perturbed independently from the others by a random number drawn from a normal distribution with an expected value of zero and a variance of one. This choice leads to an expected perturbation of 1 for each object parameter. For integer parameters, the random number is rounded to the nearest integer value. To maintain the bounds for the number of products, the search space is transformed into a periodic space, such that $x_i^{max} + 1 = x_i^{min}$.

In the second step, based on the amounts of products, the amounts of raw materials are generated. The minimum amount of raw materials which is needed to produce the desired amount of products is calculated for each period. Additional raw materials are added using the absolute value of random numbers from a discretized normal distribution. Beginning with the last period of the first-stage, the raw materials are shifted to earlier periods with low probability. In order to maintain feasible production plans, the amount of raw materials can only be shifted to earlier periods. All raw materials which are not needed can be stored.

The λ offspring candidate solutions are evaluated by the objective function f(**x**) of the master problem. The best solution **x*** is passed to the *lower layer scheduling algorithm* (see Section 5). The lower layer algorithm computes an optimal schedule for the immediate future (first-stage periods) and returns the penalty value $0 < p(\mathbf{x}^*) < p_{max}$ if the due dates are not met or a large penalty value $p(\mathbf{x}^*) = p_{max}$ if no feasible schedule exists. The penalty value p(**x***) is added to the fitness value f(**x***). By adding the penalty to the objective value f(**x***): = f(**x***) + p(**x***), the fitness based order of the individuals in the current population may change. Then the new best solution is passed to the scheduling algorithm and a new optimal schedule is generated and the penalty costs are calculated. The procedure iterates until the first solution in the sorted population no longer changes.

In the selection for population replacement, a truncation selection chooses the μ best ($1 \leq \mu \leq \lambda$) individuals out of the union of μ parents and λ offspring which do not exceed the maximum age of κ for the next iteration. The age of an individual is increased by one after each generation. After the new population is created, a new iteration loop starts if the termination criterion is not fulfilled.

5. Lower layer algorithm: Timed automata based approach

5.1. Background of timed automata

The underlying scheduling problem is solved using a priced timed automata approach. Timed automata (TA) are *finite state automata* extended by the notion of *clocks* to model discrete event systems with timed behaviour. An extension of TA with the notion of costs is known as priced TA. A priced TA is a timed automata equipped with an additional function which assigns costs for staying in locations and costs for transitions. The resources, jobs and timing constraints are modelled as sets of priced TA. For the problem considered in the case study, release dates and due dates for the production orders exist. The release date for a production order is modelled by the guards on the transition that represents starting of the first operation. The due date for each production order is modelled by a separate TA (due date automaton) and synchronized with the respective job automaton. The delay cost for the production order which miss the due date are computed using the clock in the due date automaton. A detailed explanation of the approach to model release dates and due dates using TA can be found in [8].

5.2. Cost-optimal reachability analysis

Once the individual automata have been created, the *composed priced TA* is realized by the tool. The composed priced TA is automatically created by synchronizing the transition labels in the sets of the individual priced TA. TAOpt performs a symbolic reachability analysis to explore the solution space and to derive production schedules with minimal cost. The reachability tree is created from the composed automaton. It consists of nodes and transitions; a node represents a combination of a state and clock valuations of all clocks including the global clock. A cost-optimal reachability analysis is performed starting from the initial location where no jobs are started and trying to find a path to the target location where all jobs are finished within the defined due date. A depth first search algorithm is used to explore the reachability graph. The search space reduction techniques weak non-laziness, sleep-set method and passed list inclusions were employed to prune parts of the reachability tree (for detailed explanations see [5]). The objective is to minimize the sum of the delay costs. A penalty of $p(\mathbf{x}) = 0$ is passed to the *upper layer planning algorithm* (see Section 4) if all due dates are satisfied or costs $0 < p(\mathbf{x}) < p_{max}$ which are identical to the cost of delays in the objective function on the planning layer.

6. Experimental Results

The proposed two-layer approach was tested for the case study. The decisions in periods 1 to 3 are considered as first-stage decisions, those in periods 4 to 6 as second-stage decisions. The uncertainties are represented by 4 realizations of the product demands, 2 realizations of the revenues, 2 realizations of the production costs, and 2 realizations of possible breakdowns of the reactors. The realizations (B, G) of the product demands for the end-products B and G are shown in the Table below. Overall the second-stage uncertainty is represented by 32 scenarios of equal probabilities. The upper layer hybrid algorithm was implemented in MATLAB 7.3. The MILP scenario problems were for-

Realization	1	2	3	4
Period 1	(3,2)	(3,2)	(3,2)	(3,2)
Period 2	(1,3)	(1,3)	(1,3)	(1,3)
Period 3	(4,2)	(4,2)	(4,2)	(4,2)
Period 4	(2,3)	(3,3)	(3,3)	(4,1)
Period 5	(3,2)	(2,2)	(4,3)	(2,3)
Period 6	(2,3)	(1,1)	(4,4)	(1,2)

mulated using GAMS 22.5 and solved by CPLEX 10.2. In total, the scenario subproblems consist of 8065 equations, 1249 continuous variables, and 4352 integer variables. TAOpt is employed for the optimization of the TA models in the lower layer. The global TA model is composed out of 69 TA models. It consists of 200 locations, 203 transitions, and 70 clocks. It was observed that TAOpt finds the first feasible solution after 3,000 explored nodes in most instances. TAOpt was stopped either after 20,000 nodes are explored or after a schedule with no tardiness is found. 10 runs were performed on a dual Xeon machine with 3 GHz speed. A population size of $\mu = 10$, an offspring/parents-ratio of $\lambda/\mu = 7$, and $\kappa = \infty$ were chosen. The progress plot in Fig. 2 shows the evolution of the objective function $f(\mathbf{x})$. The different lines represent the 10%-, 50%-, and 90%-quantiles. The proposed algorithm generates feasible solutions after 6 minutes. The quality of the objective is improved considerably in short time, thereafter convergence slows down. In the best run, the algorithm finds the optimum solution (determined by complete enumeration of the first-stage decisions in 36 hours) in 48 minutes which has an objective function value of -21.9775. The algorithm needs about 3 sec for the evaluation of one individual without scheduling. The computation of a schedule takes up to 3.5 sec.

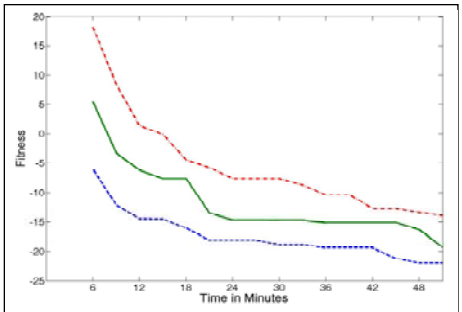


Fig. 2: Evolution of the two-layer algorithm

7. Conclusions

This work presents a successful application of a stage decomposition based hybrid algorithm for production planning and of the timed automata based approach for schedule generation to solve the coupled problem of production planning under uncertainty and scheduling. Ongoing work concerns the improvement of the interplay between planning and scheduling and the application to larger examples.

References

1. J. R. Birge, F. Louveaux: *Introduction to Stochastic Programming,* Springer, 1997.
2. S.L. Janak, C.A. Floudas, J. Kallrath and N. Vormbrock (2006). Production scheduling of a large-scale industrial batch plant. I. Short-term and medium-term scheduling. *Ind. Eng. Chem. Res.* 45, 8234-8252.
3. D. Wu and M. Ierapetritou (2007). Hierarchical approach for production planning and scheduling under unvertainty. *Chem. Eng. Proc.* 46, 1129-1140.
4. Y. Abdeddaiem, E. Asarin, O. Maler (2006). Scheduling with timed automata. *Theoretical Computer Science* 354, 272-300.
5. S. Panek, S. Engell, S. Subbiah, O. Stursberg (2008). Scheduling of multi-product batch plants based upon timed automata models. *Comp. Chem. Engg.* 32, 275-291.
6. N. Bauer, S. Kowalewski, G. Sand, T. Löhl (2000). A case study: Multi product batch plant for the demonstration of scheduling and control problems. *Proc*. ADPM 2000*,* Shaker.
7. J. Till, G. Sand, M. Urselmann, S. Engell (2007). A Hybrid Evolutionary Algorithm for Solving Two-stage Stochastic Integer Programs in Chemical Batch Scheduling. *Comp Chem Engg.* 31 , 630-647.
8. S. Subbiah,T. Tometzki, S. Engell (2008). Batch scheduling with intermediate due dates using timed automata models. *Proc.* ESCAPE 18, 151-156.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

On line monitoring and diagnosis of the operation of a hybrid CSTR by using PCA models

Luis G. Bergh and David Gómez

Chemical Engineering Department, Santa Maria University, Valparaiso, Chile

Abstract

Testing strategies for process monitoring, diagnosis and control is expensive, and usually requires either complex pilot plant facilities or to deal with the hard constraints posed by experimentation in real plants. The experience developing hybrid systems, combining pilot plants and models is presented. A model is used to simulate the reaction, given measured and virtual variables. Feed temperature and flow rate and reactor temperature are measured on line, while feed component concentrations are given as a stochastic disturbance. The temperature in the reactor can be controlled by regulating the cold water flow rate in the coil.

The potential of developing this class of hybrid systems is illustrated by the use of principal component analysis (PCA) methodology to build a model from experimental data. Hotelling T^2 and Q residual tests were running on line every time a steady state is reached. Both tests are used to identify the main causes of two kinds of problems: when the process target (product concentration) is not met and when a measuring device failed. Examples of these applications are discussed.

Keywords*: reactor, control, modelling, monitoring, diagnosis*

1. Introduction

In the last decade a number of techniques have been proposed to improve the benefits of chemical process operations. Some of them are: process modelling, data reconciliation, soft sensors and pattern recognition, process monitoring, fault detection and isolation, control loop monitoring, control algorithms and supervisory control. What is common to most of these areas is that the theoretical advantages of novel methods have to be confronted and tested in real plants. Simulation of the application of novel methods by using a plant model has some advantages. Once a model of a plant is available, the new methodology can be tested under a number of different conditions, at low cost. However, some disadvantages of this approach are how the model used represents the real nature of the process, the error in measurements and the disturbances coming into the plant. Most of the time, the simulations are conducted under some idealties, neglecting those aspects, and thus increasing the uncertainty and reducing the validity of results. Comparisons of different methods made on such bases may become useless, and cannot replace experimentation in the real world.

On the other hand, experimentation carry on real plants poses its own difficulties. For example, the scale factor forces that changes on some variables must be inside a narrow band, to avoid risky operating conditions or high loses in products. Some disturbances cannot always be managed, interrupting and degrading the experiments, leading to

confusing results. In summary, it is more difficult and costly to conduct a designed experience, which collected data can be properly analysed to make right decisions.

Pilot plants can reproduce most of the advantages of experimenting in a real plant, but at a small scale that relaxes some of those constraints, at the same time that provides the chance to a wide experimentation. However, expensive chemicals, large storing facilities, waste treatment units and expensive on stream analysers remain as drawbacks.

2. A different approach

Some processes can be analyzed separating the phenomena in different levels. A first level is the process hydrodynamics, heat and mass balances. A second level is the chemical change by reaction. Sometimes changes in solute concentration may also change transport properties, and the artificial decoupling will not be effective. When process hydrodynamics are not significant influenced by changes in solute concentration, then the experimentation approach on pilot plants can be simplified in two senses: Experimentally each fluid may be substituted by low cost water in a liquid phase reaction, and the solute concentration will not actually change, but an estimation can be obtained from detailed models relating measured operating conditions, such as flow rates, temperatures, pressures, levels, and concentrations.

Therefore, if a model is available, and the pilot plant is operated by using these low cost fluids, a hybrid system is developed. The real plant behaviour is simplified but the main hydrodynamics characteristics are still well represented. The variables representing the target of the process are predicted by using the model, under a wide operating region. Now, distributed control of local objectives can be administrated by supervisory control strategies based on estimation of the crucial variables. Also process monitoring, diagnosis and fault detection, isolation and remediation studies can be developed under low cost and reasonable approximation to the behaviour of a real process.

Models of different kinds can be built and solved on-line (Bergh, 2007). Here, a steady state model is used to generate the data to build a PCA model, and then this model is used to test on-line the operation of the reactor.

3. Pilot CSTR unit

A P&ID of the pilot SCTR is shown in Figure 1. Both, feed and cooling water are measured and controlled by variable speed peristaltic pumps. The electric power supplied is under control, and it is used to simulate the heat of reaction. Inlet and outlet temperatures are measured. Virtual measurement of reactant A is generated in the computer, while the outlet concentration is calculated from the model by using the measured operating variables.

A first order kinetics exothermic and irreversible reaction is considered. The input variables measured are: feed flow rate (F) and temperature (Ti), cooling water flow rate (Fw) and temperature (Twi). The output variables measured are reactor temperature (T), cooling water temperature (Two) and ambient temperature (Ta). Given an input concentration (Cao) the output concentration (Ca) is calculated from:

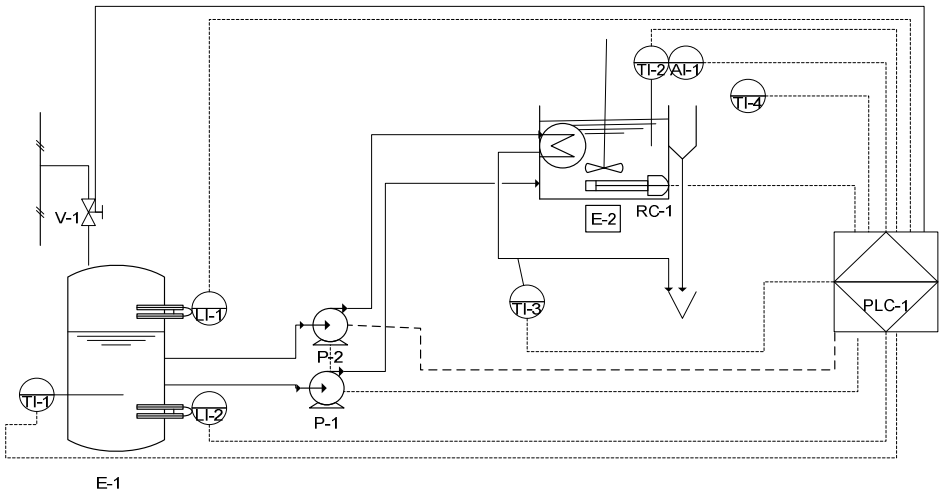


Figure 1. P&ID of pilot CSTR.

$$Ca = \frac{Cao}{(1+k\tau)} \tag{1}$$

with $k = k_0 e^{-\frac{E}{RT}}$ (2)

and $\tau = \frac{V}{F}$ (3)

The heat of reaction is calculated from: $Q = Ca \cdot k(-\Delta H)_r \cdot V$ (4)

The reactor temperature is measured at each sampling time (every 10 s), and k, Ca and Q are recalculated by using the model. The new value of heat of reaction is implemented by the electrical heater RC-1. Convergence is obtained after three or four iterations (less than one minute) and then the reactor reaches a steady state, where all material and heat balances are met. Similar iteration will occur if any of the input variable changes. This hybrid system will provide complete operating data that will be strictly valid only at steady state. In the transient period between steady states the estimation of reactor concentration will not be equal to the real one, but close enough to show smoothly the tendency. More details can be found in Gomez (2008).

4. PCA models

The key feature of PCA method is their ability to mathematically project high dimensional process and quality data into smaller dimensional, summary data sets via the development of linear models. The practical value of PCA modeling method is that this technique allows for the systematic examination and interpretation of the model outputs. Examination of the model outputs can provide insight into the operation of an industrial process during monitoring and quality assurance activities. With PCA, the systematic interpretation of dominant patterns in the data and the isolation of the most important contributors to these patterns are possible. This allows the classification of

data relationships according to normal and abnormal operation. Some of these numerous advantages of PCA method has over traditional monitoring and prediction technologies are: provision for data dimension reduction and robustness to highly correlated, noise and missing data (Kourti and MacGregor, 1995).

The concept of a latent variable model is that the true dimension of a process is not defined by the number of measured variables, but by the underlying phenomena that drive the process. The latent variables themselves are modeled as mathematical combinations of the measured variables and describe directions of variation in the original data. A latent variable model can contain many fewer dimensions than the original data, it can provide a useful simplification of large data sets, and it can allow better interpretation of the measured data during analysis (MacGregor *et al*., 2007).

5. PCA model building

Nine variables were considered, seven were measured on-line, the initial concentration was given as a date and the output concentration was calculated using heat and mass balances and kinetic model. The data were collected for different feed flow rates and feed concentration. An example of a set of data collected every 5 seconds is shown in Figure 2. When all the variables have reached a steady state, this point is a candidate to form the data matrix, after filtering. A normal operation is defined when a conversion over 95 % is reached in the reactor. More than 200 steady states were collected and filtered. This data was processed using PLS_Toolbox from Eigenvalue Research.

A PCA model was built. A model with 4 latent variables was found to explain at least 95 % of the variance in the centered and scaled pretreated data. The T^2 Hotelling, at 95 % confidence interval (using the F distribution), was 10.99 and the residuals Q limit (following Jackson, 2003), at 95 % confidence interval, was 0.29. These statistics will be used to test when the reactor operation reached an abnormal situation, and which variables are contributing the most. Two kinds of problems were studied. First, when the conversion in under 95% and when a gross error in the measurement of one operating variable occurred.

When the reactor operated with a feed flow rate of 13 mL/s, a cooling water flow rate of 20 mL/s, and a initial concentration of 1.0 mol/L, the temperature was 32.5 °C and the heat of reation was 1420 W. The conversion was 93.7% and the Q test was satisfied while the T2 Hotelling test was failed. The contribution of each variable to the failure is shown in Figure 3. The output concentration was low mainly due to a combination of high feed flow rate and high cooling water flow rate (low k and τ).

If the feed flow rate measurement is deviated by a fractional error, then the T^2 test is satisfied and the Q test is failed for errors less than -0.2 and great than 0.4, as can be observed from Figure 4. Similarly, if the input concentration measurement is deviated by a fractional error, then the T^2 test is satisfied and the Q test is failed for errors greater than 0.2.

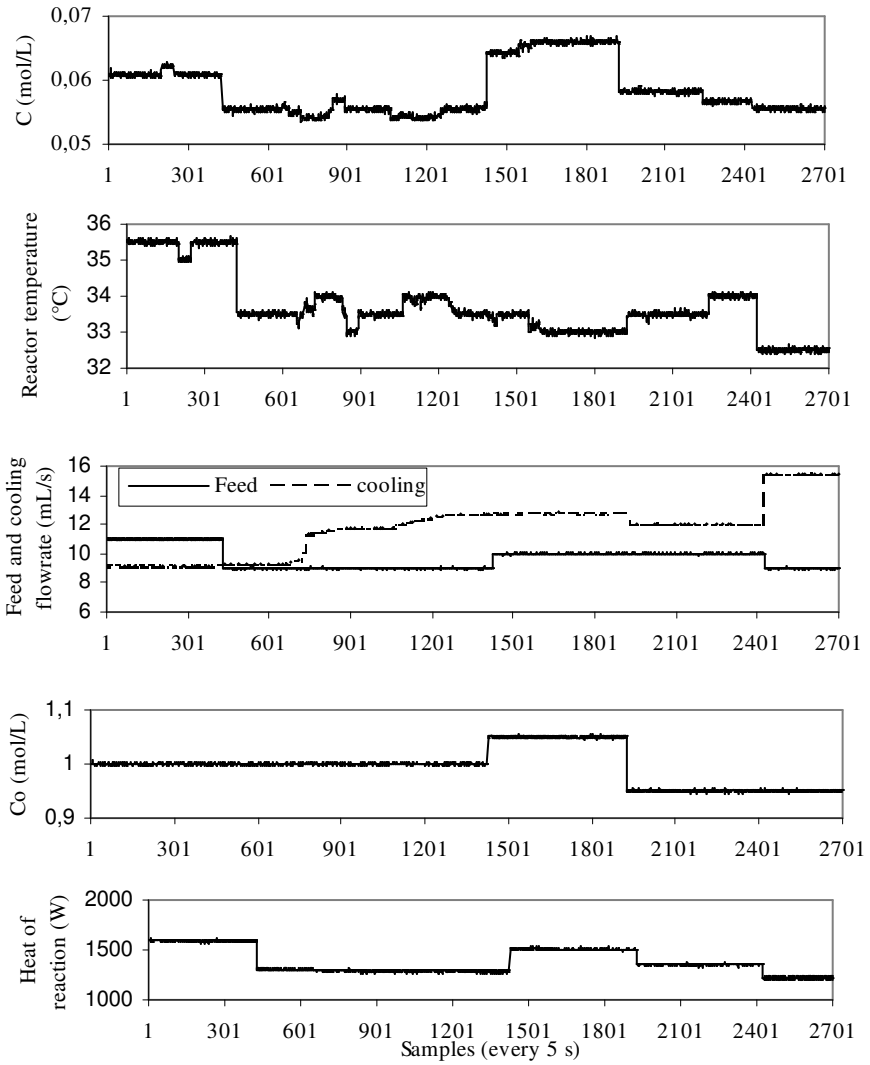


Figure 2. Operating data collected from pilot CSTR.

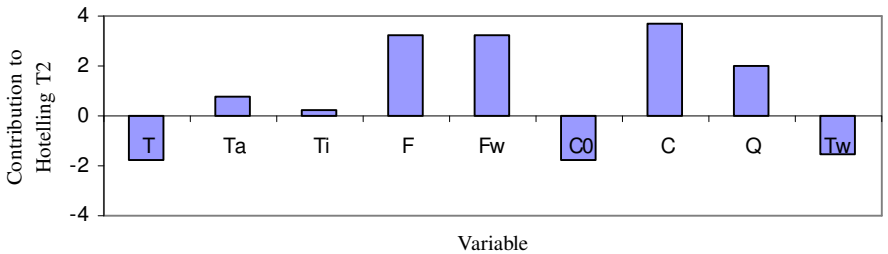


Figure 3. Contribution of variables to Hotelling's T^2.

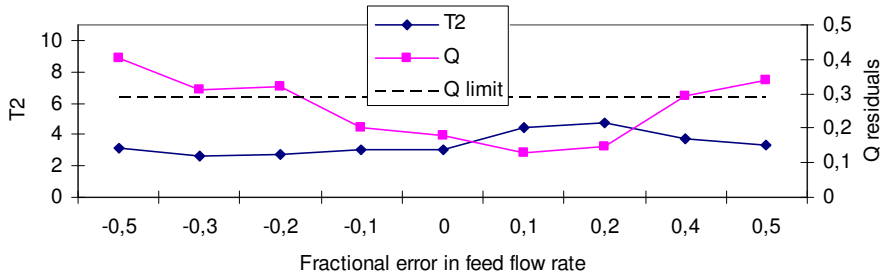


Figure 4. Sensitivity to detect fractional error in feed flow rate.

5. Conclusions

A heated and cooled tank has been operated combined with a phenomenological model to predict on-line the output concentrations of a CSTR reactor. This configuration permited to operate the reactor in a wide range, at low cost, and to test novel methodologies to detect, isolate and remediate operating and measurement failures.

The application of multivariate statistical methods, and particularly PCA, is a powerful tool to build linear models containing the essencial of the process phenomena with the minimum number of latent variables. The application of PCA models to monitoring a pilot CSTR has been demonstrated. These PCA models can be effectively used as part of a supervisory control strategy, specially when control decisions are infrequently made.

Acknowledgement

The author would like to thanks Conicyt (Project Fondecyt 1070106) and Santa Maria University (Project 270968) for their financial support.

References

Bergh L.G. (2007). Combining on-line process measurements and models to empirically test strategies for process monitoring, diagnosis and control, Proceedings 12th IFAC Symposium MMM 07, Quebec City, August 21-23, 193-198.

Gómez D. (2008). Monitoreo en linea de reactor agitado calefaccionado usando métodos de proyección, Tesis magister Ing. Quím., Universidad Santa María, Valparaíso, Chile.

Jackson J.E. (2003). A user´s guide to principal components, John Wiley, New Jersey.

Kourti T. and J.F. MacGregor, (1995). Process Analysis Monitoring and Diagnosis using Multivariate Projection Methods – A Tutorial, Chemometrics and Intelligent Laboratory Systems, 28, 3-21.

MacGregor J.F., T. Kourti, J. Liu, J. Bradley, K. Dunn, H. Yu (2007). Multivariate methods for the analisys of data.bases. process monitoring, and control in the material processing industries, Proceedings 12th IFAC Symposium MMM 07, Quebec City, August 21-23, 193-198.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Using Wavelet Texture Analysis in Image-Based Classification and Statistical Process Control of Paper Surface Quality

Marco S. Reis,[a] Armin Bauer [b]

[a]*CIEPQPF, Department of Chemical Engineering, University of Coimbra, Rua Sílvio Lima, 3030-790, Coimbra, Portugal.*
[b]*Voith Paper Automation GmbH & Co KG, Linzer Strasse 55, 3100, St, Pölten, Austria*

Abstract

Paper formation (the level of homogeneity in the distribution of fibers on the surface of paper) is a key quality parameter for paper products, being currently monitored off-line, at low sampling rates relatively to the high production speeds achieved with modern paper machines. In this paper, we address the problem of conducting such monitoring activity on-line, in situ, using wavelet texture analysis (WTA) on raw images acquired with a especially design sensor. Our analysis shows that either a reduced set of features derived from WTA (suggested by an ANOVA analysis), or a low dimensional sub-space (with two dimensions, obtained using PCA or PLS-DA), enable an adequate separation of the several paper formation quality grades, meaning that we can indeed follow, on-line, the quality status of paper formation. Furthermore, we will also show how statistical process control (SPC) can be properly conducted using WTA features, in a simple and robust way.

Keywords: Multivariate Image Analysis, Wavelet Texture Analysis, Paper Formation, Multivariate Statistical Process Control, Principal Components Analysis.

1. Introduction

Paper formation (the level of homogeneity in the distribution of fibres on the surface of paper) is a key quality parameter for paper products, as it impacts not only the aesthetic evaluation and visual perception of quality by the end users, but also affects other relevant properties such as the average strength of paper (a sheet with poor formation is weaker than a comparable sheet with a better formation), the quality of printing operations (which is inferior for papers with poor formation) and, furthermore, leads to process problems, such as, for instance, the increase of coating consumption, as paper formation gets worse. For these reasons, paper formation constitutes a matter of great concern to paper producers, being therefore routinely measured in the quality control laboratories in modern paper production facilities, either through visual inspection and comparison with a series of standards representing different quality levels of formation or, more frequently, with resource to instrumentation that scan the paper sheet, measuring the variation of light transmission through it. This monitoring activity occurs only a few times per day, according to the routine testing plans established in each paper mill. For instance, a possible routine may consist on taken a sample for each paper reel produced, which is then sent to the control laboratory, where, after some time, the measurements are performed and results introduced in the product quality information system. This whole process introduces a very significant delay in the supervision and

control tasks, given the high production rates achieved by modern paper machines, where paper is being currently produced at linear speeds that can be over 100Km/h. Therefore, we can easily conclude that such off-line process monitoring has strong limitations, being rather ineffective in promptly detecting problems in paper formation and leading to significant losses in quality when process upsets do arise. In this context, we address in this paper the development of technology for performing paper formation monitoring, on-line and in situ, using a state-of-the-arte texture analysis methodology: Wavelet Texture Analysis (WTA). WTA enables the extraction of informative image features regarding phenomena contained in the textures (Bharati *et al.*, 2004), which are computed from the 2D wavelet transform coefficients of raw images. Wavelet analysis has been already applied to the analysis of paper formation, but only the wavelet details coefficients for the second decomposition level were used (Bouydain *et al.*, 1999), something that, according to our results, may seriously hinder the discrimination ability that can be achieved among different level of quality in paper formation.
This article is organized as follows. In the next section we briefly describe the technology underlying the image acquisition system and address the main methodological steps of our analysis approach. Then, we present some results obtained, regarding both the discrimination ability that can be achieved using WTA features, and address how can they be used for conducting statistical process control (SPC) in a proper way. Finally, we briefly conclude with a summary of the main results achieved.

2. Materials and Methods

2.1. Image acquisition sensor

The image acquisition sensor is a critical element for the on-line monitoring and analysis of paper formation. Its function is to capture good quality images in the harsh environmental conditions prevailing in the forming section of the paper machine, which is essentially steamy and wet, with drops erratically flying everywhere around the camera. The technological solution developed for the formation sensor, consists of a digital camera within a housing that is able to rotate at high speed, in order to prevent dirt accumulation on the housing surface. Such design protects the camera from the environmental conditions, ensuring that the sensor will function properly with little maintenance. The camera used was a Jai A10 CL, with a Navitar DO-2595 lens. The strobe light consists of a led array emitting red light (CCS LDL-TP).

2.2. Methods

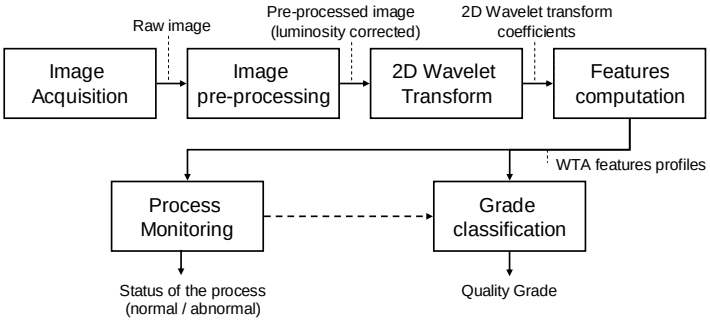


Figure 1. Scheme of the WTA methodology for monitoring and classification of paper formation.

The basic methodological steps considered in our analysis are as follows (Figure 1). Images are first acquired (i) and corrected for non-uniform illumination effects (ii), even

though this is not a critical step in the analysis, as our procedure is rather robust to this kind of problems (i.e., from our experience, final results do not change much upon removing such pre-processing task from the analysis). Then, we (iii) computed the 2D wavelet transform for each raw image (Mallat, 1999), which essentially consists of transforming a matrix of numbers (pixel intensities, as we are analyzing single-channel or grey-level images) into another, with the same size (same overall number of wavelet coefficients), containing blocks of coefficients regarding details for different scales (from the finest to the coarsest scale, which is known as the decomposition depth, J_{dec}) and along three different directions (horizontal, vertical and diagonal). In our case, the decomposition depth considered was $J_{dec} = 5$, and therefore each image will give rise to 16 blocks of data or sub-matrices ($=J_{dec}$x3+1; the "1" stands for the block of final approximation coefficients, containing the lowest resolution version of the image, after removing all the details corresponding to the finer scales). The next stage in WTA is (iv) to summarize the information contained in the wavelet coefficients in each of these blocks into a single number, from which a vector of WTA features will result. From the several possibilities available (energy, entropy, averaged l_1-norm, standard deviation) we used the standard deviation (the wavelet transform of uncorrelated homogeneous white noise is also white noise with the same standard deviation; therefore by using the standard deviation to summarize the information in each block, we are tactilely using a stochastic uncorrelated frame of reference against which we compare the texture in the images). As the end of this fourth stage, we have compressed the original analysis space composed by 512x512 pixels or intensity numbers, to just 16 numbers that are expected to contain the essential information necessary to discriminate different formation patterns. Finally, the extracted WTA features can be used for (v) classifying the acquired image by proposing a quality grade and/or for conducting statistical process control of this property. The first activity (classification) requires the development of a suitable classifier, which usually involves a feature selection/compression stage and a subsequent classification stage (Pal and Mitra, 2004). For the first task we have considered here variable selection (using ANOVA) as a feature selection methodology and partial least squares for discriminant analysis (PLS-DA), as well principal components analysis (PCA), for features compression. As to the second task (process monitoring), we applied multivariate statistical process control (MSPC) based on PCA, in order to deal with the high collinearity present in the WTA features (Jackson and Mudholkar, 1979; Kresta *et al.*, 1991).

3. Results

A set of 24 images, representing different paper formation patterns, was collected for analysis. These raw images were classified, by visual inspection, into three quality levels regarding paper formation: 1-"good", 2-"poor", 3-"bad" (Figure 2). For two images were given special annotations: a dubious sample was assigned the class label "1,5", and an outlier, possibly resulting from an image acquisition problem, was labeled with score "9". Fallowing the sequence of steps represented in Figure 1, these images were pre-processed for non-uniform illumination correction and transformed to 2D wavelet coefficients. Finally, the wavelet feature profiles were computed, containing 16 features for each image, based on which the subsequent analysis was performed. This analysis had two main components: assessing the discrimination ability of wavelet features to discern among different formation quality grades (classification) and to study how statistical process control could be conducted based on such features (monitoring).

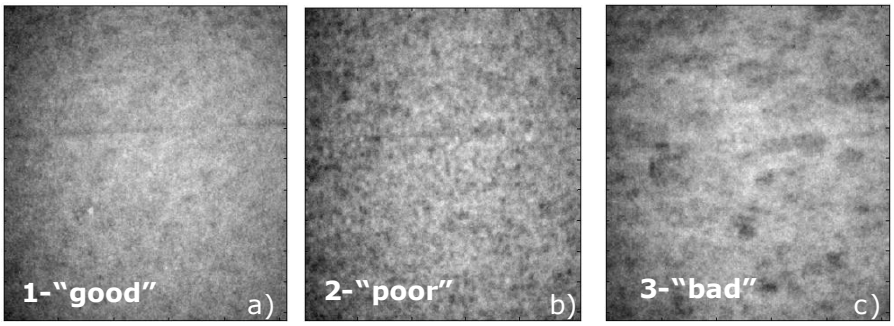


Figure 2. Images representative of the three main quality grades of paper formation.

3.1. Classification of paper formation

Even though the compression magnitude of the entire image information to 16 features is indeed quite large, such features are still highly redundant. For instance, the first principal component (PC) in a PCA analysis over these 16 features, represents 97.96% of the overall data variability. However, in order to achieve a proper discrimination, a minimum of two PC must be retained (we get to this number by computing the error rate estimated by "leave-one-out" cross-validation using an increasing number of retained PC, and applying the simple linear classifier to perform the classification task, from which we verified that, after retaining the second PC, the error rate dropped from ~ 23% to 0%). Figure 3 illustrates the separation ability obtained from different compression strategies (to reduce the number of features from 16 to 2), followed by the estimation of a linear classifier in the resulting reduced map (as the separation is so good, there is no need to use more complex classifiers): PCA, PLS-DA (Barker and Rayens, 2003) and variable selection (using ANOVA to select those features that most descriminate the various formation quality levels), followed by Fisher Discriminant Analysis (FDA), as this method alone is not appropriate to handle situations were the classifying features are heavily correlated, due to matrix inversion problems, as happens in this case. As we can verify, all methods successfully discriminate among the three levels of quality present in the sample images, meaning that WTA does provide an adequate basis for analyzing paper formation.

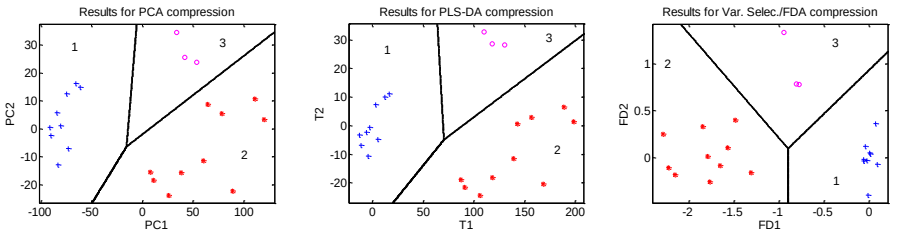


Figure 3. Results for the classification of samples after compressing the number of WTA features from 16 to 2: a) using PCA; b) using PLS-DA; c) using variable selection (based on ANOVA), where 3 variables were selected, and FDA.

3.2. Monitoring paper formation

In order to use the WTA feature profiles for monitoring purposes, a suitable statistical process control scheme was adopted, with well established results and adequate for

addressing the highly correlated nature of the data matrix of wavelet features: multivariate statistical process monitoring (MSPC) based on PCA (PCA-MSPC) (Jackson, 1959; Kresta *et al.*, 1991). This methodology consists of modelling systems operation under normal conditions using a PCA model, in order to capture the normal correlation structure and variability relative to the so called "common" causes. For each (new) multivariate observation (x_i), two statistics are computed, one relative to the (Mahalanobis) distance from its projection onto the PCA subspace to the centroid of the normal operation conditions (NOC) cluster, also in the PCA subspace, usually referred as T_i^2, and another relative to the (squared) Euclidean distance to such a subspace, Q_i.

In order to assess the application of this approach to the on-line monitoring of paper formation, a sub-set of four images from class 1-"good" were selected, representing normal operation conditions (NOC; images 1, 2, 7 and 13), from which the method parameters, including the statistical control limits, were estimated. All the remaining images, with representatives from all grades of quality, were used for testing the monitoring procedure. Figure 4 depicts the projection of all samples into the PCA subspace estimated from the wavelet features for the four reference images, where it is clear that all reference and testing samples from normal operation conditions fall inside the NOC region (95% control ellipse, also computed only using the reference set), whereas abnormal samples are also correctly signalled as such, falling outside the NOC ellipse.

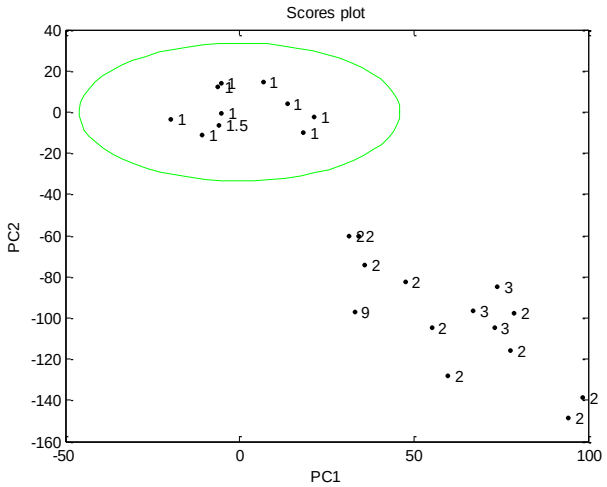


Figure 4. Projection of all samples (train and test) on the PCA subspace estimated from the four training images of quality grade 1-"good", taken as representative of NOC, with the associated 95% NOC control ellipse. All abnormal samples in the test set (class 2 or 3) fall outside the control ellipse.

From Figure 4, it is possible to see that the dubious sample labelled "1.5" also falls inside the NOC region, which is not unexpected given its similarity to other images from this class. However, an analysis of the method statistics (T^2 and Q) clarify its status and rather abnormal texture (see plot for Q), indicating that it is indeed an abnormal image (Figure 5; note the logarithmic scale used in the YY' axis, given the large magnitude difference between the statistics values for the normal and abnormal situations).

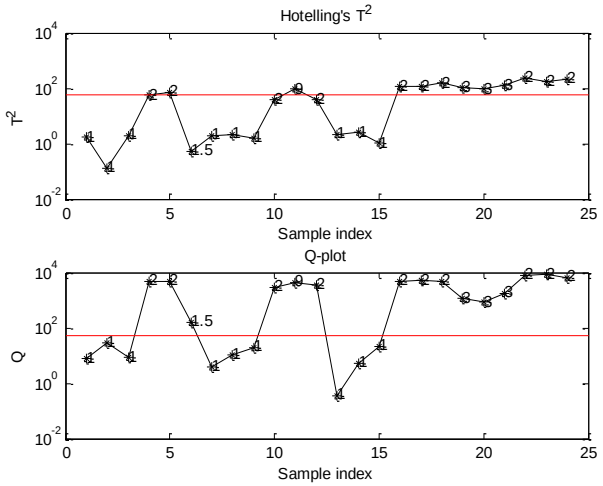


Figure 5. Control charts for the T^2 and Q statistics, regarding the implementation of PCA-MSPC over all images (train and test). Using the combination of the two statistics, all samples are correctly classified as normal or abnormal. (Note the logarithmic scale used in the plots).

4. Conclusions

In this article, we have presented an approach for classifying the quality grade of paper formation based on the WTA features profiles computed from images taken on-line from a pilot scale paper machine. Furthermore, their application to process monitoring was also addressed. Our analysis showed that the classification task is successfully conducted in a reduced dimensional space (2 dimensions), using either a reduced set of features (suggested by an ANOVA analysis), or a low dimensional sub-space (obtained through PCA or PLS-DA). Regarding the monitoring task, it was also possible to verify that statistical process control (SPC) can be properly conducted using WTA, in a simple and robust way. Therefore, we can conclude from the results presented, that it is possible to simultaneous monitor and assess formation quality, on-line and in-situ, using WTA-analysis of images acquired with the special designed sensor developed.

References

Barker, M. & Rayens, W., 2003, Partial Least Squares for Descrimination, *Journal of Chemometrics, 17*, 166-173.

Bharati, M. H., Liu, J. J. & MacGregor, J. F., 2004, Image Texture Analysis: Methods and Comparisons, *Chemometrics and Intelligent Laboratory Systems, 72*, 57-71.

Bouydain, M., Colom, J. F. & Pladellorens, J., 1999, Using Wavelets to Determine Paper Formation by Light Transmission Image Analysis, *Tappi Journal, 82*, 7, 153-158.

Jackson, J. E., 1959, Quality Control Methods for Several Related Variables, *Technometrics, 1*, 4, 359-377.

Jackson, J. E. & Mudholkar, G. S., 1979, Control Procedures for Residuals Associated with Principal Component Analysis, *Technometrics, 21*, 3, 341-349.

Kresta, J. V., MacGregor, J. F. & Marlin, T. E., 1991, Multivariate Statistical Monitoring of Process Operating Performance, *The Canadian Journal of Chemical Engineering, 69*, 35-47.

Mallat, S., 1999, *A Wavelet Tour of Signal Processing* (2nd ed.). San Diego: Academic Press.

Pal, S. K. & Mitra, M., 2004, *Pattern Recognition Algorithms for Data Mining*. Boca Raton: Chapman & Hall/CRC.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Numerical Pitfalls by State Covariance Computation

Nina P. G. Salau,[a] Jorge O. Trierweiler,[a] Argimiro R. Secchi,[b]

[a]*GIMSCOP/DEQUI/UFRGS, Campus Central, Porto Alegre/RS, CEP 90040-040, Brazil*
[b]*PEQ/COPPE/UFRJ, CP 68502, Rio de Janeiro/RJ, 21945-970, Brazil*

Abstract

With the error past information, conveyed by state covariance matrix, a state estimator can predict the state vector one step ahead, thus providing system operators with more information to make control decisions such as security assessment and other related functions. A well-computed state covariance matrix avoids error propagation due to numerical pitfalls and, thereby, is crucial for a successful state estimator design. In this paper we investigate the numerical robustness of four EKF formulations and derive some results giving insights into their numerical performance. As benchmark case study, we have chosen a batch chemical reactor with reversible reactions whose system model and measurement are such that multiple states satisfy the equilibrium condition. Finally, we have shown that CEKF is the best alternative to EKF for such case study.

Keywords: nonlinear state estimation, state covariance, multiple solutions.

1. Introduction

It is well known that a suitable design of state estimators requires a representative model for capturing the plant behavior and knowledge about the noise statistics, which are generally not known in practical applications (Vallapil & Georgakis, 2000). In practical applications, however, some pitfalls such as numerical round-off, plant-model mismatch and state unobservability deserve also special attention because they can lead to divergence problems (Brown & Hwang, 1996). Any state covariance matrix equation is composed by states and measurements linear models and noise covariance statistics and, hence, all of the mentioned pitfalls may increase the error propagation conveyed by this matrix. In the literature, several modified implementations of the EKF are presented in an effort to avoid numerical pitfalls (Simon, 2006). The basic difference between these formulations is concerned with the state covariance matrix computation. In this work we investigate the numerical robustness of four EKF formulations applied to a system tending to an equilibrium condition. Firstly, we outline a condition that leads to a classical EKF formulation to converge to physically unrealizable state and demonstrate the potentiality of two other EKF formulations to handle with this numerical pitfall. Next, we demonstrate that the constrained extended Kalman filter (CEKF) is the best alternative to EKF for such system due to the possibility of incorporating constraints into an optimization problem preventing, hence, the estimator from converging to physically unrealizable states.

3. Formulation and Solution of the Estimation Problem

Consider the following nonlinear dynamic system to be used in the state estimators of this work:

$$\dot{x} = f\left(x,u,t\right) + \omega\left(t\right) \tag{1}$$

$$x_k = f_{k-1}\left(x_{k-1}, u_{k-1}, t_k\right) + \omega_{k-1} \tag{2}$$

$$y_k = h_k\left(x_k, t_k\right) + \nu_k \tag{3}$$

where u denotes the deterministic inputs, x denotes the states, and y denotes the measurements. The process-noise vector ω (ω_k) and the measurement-noise vector v_k are assumed to be white Gaussian random processes with zero mean and covariance Q (Q_k) and R_k, respectively.

1.1. Unconstrained EKF formulations

We have selected three different unconstrained EKF formulations to be compared in our work: a classical EKF formulation known as discrete EKF (DEKF), the EKF with continuous Riccati equation (EKF-CRE) and the reduced-rank extended Kalman filter (RREKF). The basic difference of these formulations is related to state covariance matrix computation, as given in Table 1.

Table 1. Three unconstrained formulations of EKF

Filter	State Covariance Matrix		
DEKF	Transition	$P_{k\|k-1} = F_{k-1}P_{k-1\|k-1}F_{k-1}^T + Q_{k-1}$	(4)
	Update	$P_{k\|k} = \left[I_n - K_k H_k\right]P_{k\|k-1}$	(5)
EKF-CRE	Transition & Update	$P_{k\|k} = P_{k-1\|k-1} + \int_{k-1}^{k}\begin{bmatrix} F(\tau)P(\tau) + P(\tau)F(\tau)^T + Q(\tau) \\ -P(\tau)H(\tau)^T R(\tau)^{-1} H(\tau)P(\tau)\end{bmatrix} d\tau$	(6)
RREKF	Transition	$P_{k\|k-1} = F_{k-1}P_{k-1\|k-1}F_{k-1}^T + Q_{k-1}$	(7)
		$P_{q,k\|k-1}^{rr} = \sum_{i=1}^{q}\lambda_{P,i}\upsilon_{P,i}\upsilon_{P,i}^T$	(8)
	Update	$P_{k\|k} = \left[I_n - K_k H_k\right]P_{k\|k-1}$	(9)

1.2. Moving Horizon Estimator

MHE is described by the following optimization problem (Muske & Rawlings, 1994; Rao & Rawlings, 2003).

$$\min_{\substack{\omega_{k-N-1\|k},\cdots,\omega_{k-1\|k} \\ \nu_{k-N\|k},\cdots,\nu_{k\|k}}} \Psi_k^N = \begin{bmatrix} \hat{\omega}_{k-N-1\|k}^T \left(P_{k-N-1\|k-1}\right)^{-1} \hat{\omega}_{k-N-1\|k} \\ + \sum_{j=k-N}^{k-1} \hat{\omega}_{j\|k}^T \left(Q_{k-1}\right)^{-1} \hat{\omega}_{j\|k} + \sum_{j=k-N}^{k} \hat{\nu}_{j\|k}^T \left(R_k\right)^{-1} \hat{\nu}_{j\|k} \end{bmatrix} \tag{10}$$

The state covariance matrix is computed by the Riccati discrete equation (Rao & Rawlings, 2003).

1.3. Constrained EKF

CEKF follows from MHE for a zero-length horizon (N=0). If the measurement equation is linear, the resulting problem is a quadratic program which can be solved with small computational effort (Gesthuisen et al., 2001). CEKF is similar to EKF because also takes into account a zero-length horizon in the updating stage.

2. Example of Numerical Pitfalls by State Covariance Computation

2.1. Case Study

As benchmark case study, we have chosen a batch reactor process, as introduced by Haseltine & Rawlings (2005). All reactions are reversible and follow the ideal gas law:

$$\begin{aligned} A &\underset{k_2}{\overset{k_1}{\rightleftharpoons}} B + C \\ 2B &\underset{k_4}{\overset{k_3}{\rightleftharpoons}} C \end{aligned} \tag{11}$$

Multiple states satisfy the equilibrium condition for a given measurement, which in this case is the system pressure at the equilibrium, evaluated by the following equation:

$$y = p = \left(c_A + c_B + c_C\right) RT \tag{12}$$

Table 2 presents the possible theoretical solutions, without measurement or state noise, at the equilibrium pressure given by the initial state: $c_0 = \begin{bmatrix} 0.5 & 0.05 & 0 \end{bmatrix}$. Note that only the solution EQ1 has physically realizable states (nonnegative concentrations).

Table 2. Equilibrium points

Component	EQ1	EQ2	EQ3
c_A	0.0122	-0.0267	-1967.4
c_B	0.1826	-0.2372	-9.9454
c_C	0.6669	1.1257	1978.2

The state estimation parameters and the poor initial guess $\bar{x}_0$ used for this example were obtained from Haseltine & Rawlings (2005):

$$\Delta t = t_k - t_{k-1} = 0.25\ min \tag{13}$$

$$P_0 = diag\left(0.5^2 \quad 0.5^2 \quad 0.5^2\right) \tag{14}$$

$$Q = diag\left(0.001^2 \quad 0.001^2 \quad 0.001^2\right) \tag{15}$$

$$R = 0.25^2 \tag{16}$$

$$\bar{x}_0 = \begin{bmatrix} 0 & 0 & 4 \end{bmatrix} \tag{17}$$

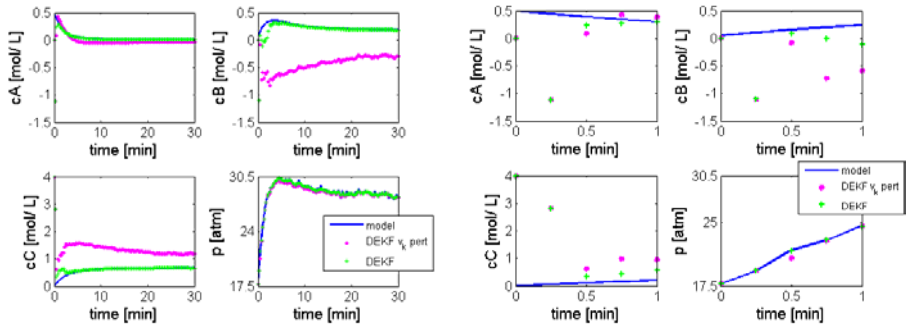
According to the authors, EKF fails in this example because the system model and measurement are such that multiple states satisfy the equilibrium condition and is given a poor initial guess of the state for the estimator. Nonetheless, we cannot assert that EKF fails because two equilibrium points (EQ1 and EQ2 in Table 2) are possible due to the poor and incoherent initial state and its respective state covariance initial guess used in this example. Further, the theory of EKF formulations assumes system observability, i.e., the state dynamic system is uniquely determinable from its inputs and outputs, given a model for the dynamic system. Since at least two states exist that realize the same equilibrium pressure, the batch reactor system is considered unobservable according to the nonlinear observabilty. Because this system satisfies the observability rank condition (linear observability), it is also considered locally weakly observable (Hermann and Krener, 1977). As the nonlinear observability assumption is violated, it is not surprising that the theory of EKF formulations may not work for this example. Details on observability definition can be found in Sontag (1998).

2.2. Results and Analyses

2.2.1. Measurement Noise Perturbation

Firstly, uniform and normally distributed random measurement noises were simulated. Either solution EQ1 or solution EQ2 is obtained in accordance with the set of random measurement noise employed into DEKF. Thus, we have chosen a set of random measurement noise which leads DEKF to converge to the solution EQ1 and added a noise measurement perturbation in order to lead it to converge to solution EQ2. As it

can be seen in Fig. 1, a noise measurement perturbation of 0.754 *atm* at t = 0.5 *min* changes the estimated states trajectory from solution EQ1 to solution EQ2.



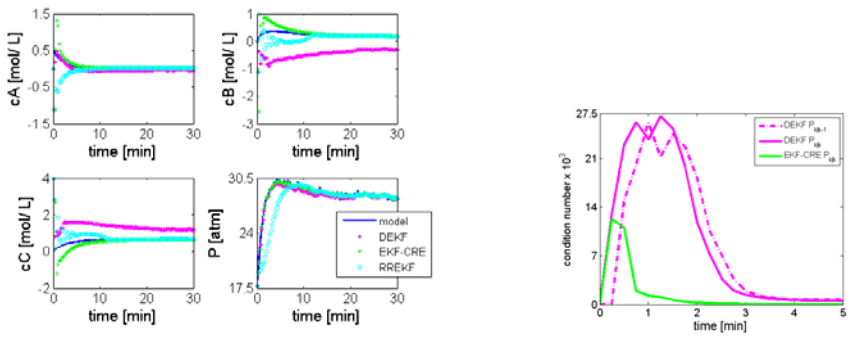
(a) (b)

Fig. 1. Comparison between the performances of DEKF with and without a noise measurement perturbation of 0.754 *atm* at t=0.5 *min*: (a) until final batch time and (b) until t=1 min.

We have also experienced that other numerical pitfalls can lead DEKF to converge to solution EQ2, such as low-precision arithmetic, gradient (linear model) calculation by finite differences, poor guess of P_0 and incorrect values of tuning parameters Q and R.

2.2.2. Comparison between Unconstrained EKF formulations

To prevent the state estimator from converging to EQ2, two other unconstrained EKF formulations were applied to this example. In spite of a measurement noise perturbation, RREKF and EKF-CRE have converged to solution EQ1, as shown in Fig. 2(a).



(a) (b)

Fig. 2. Comparison between (a) DEKF, EKF-CRE and RREKF performances; (b) the condition number of DEKF and EKF-CRE state covariance matrices.

RREKF disregards non-dominant eigenvectors which implies zero variance in the respective directions and no effect of measurement updates. For this example, the non-dominant eigenvector is orthogonal to the tangent of solution EQ1 (the scalar product approaches zero). RREKF converges to EQ1 because it applies no correction in the non-dominant eigenvector direction.

EKF-CRE is not subjected to errors due to model discretization and the term $PH^T R^{-1} HP$ inserts nonlinearity into the respective state covariance matrix. Besides, the state covariance matrix computed by EKF-CRE in a single stage (Eq. 6) presents a smaller condition number, i.e., it is less sensitive to perturbations than the states

covariance matrices computed by DEKF (Eqs. 4 and 5), as shown in Fig. 2(b). The mentioned advantages of EKF-CRE over DEKF justify the convergence of this formulation to EQ1, even with a measurement noise perturbation.
Although EKF-CRE and RREKF have prevented from physically unrealizable states at the final batch time, physically unrealizable states (negative concentrations) were unavoidable during the batch. This fact is justified by the poor state and respective state covariance initial guesses.

2.2.3. Comparison between clipped DEKF and CEKF

To prevent physically unrealizable states, we have constrained DEKF with an ad hoc clipping strategy in which negative update values of the state are set to zero (i.e., if $\hat{x}_{k|k} < 0$, set $\hat{x}_{k|k} = 0$). The comparison between clipped DEKF and CEKF performances are shown in Fig. 3(a). Before eventually converging to the actual states, the pressure filtered by the clipped DEKF is quite larger than the measured one. Due to a lower initial guess of the state covariance matrix, $P_0 = diag\left(0.022^2 \quad 0.022^2 \quad 0.022^2\right)$, the clipped DEKF in our work has presented a better performance in comparison to performance of the clipped DEKF in Haseltine & Rawlings (2005). In their paper, the clipped DEKF drives the predicted pressure 3 orders of magnitude larger than the measured pressure before eventually converging to the actual states at a longer time scale (1 order larger than the converge time obtained in our work). A clipped DEKF, however, disregards the assumption that v_k is a Gaussian random noise and does not let the Kalman gain to distribute properly the measurement residual throughout the estimated state and, thereby, corrects them. On the other side, CEKF swiftly converges to the actual states and minimizes ω_k and v_k in a least square sense, incorporating constraints into an optimization problem, which prevents from bad noises distribution. Fig. 3(b) illustrates the distribution of v_k for the clipped DEKF and the CEKF.

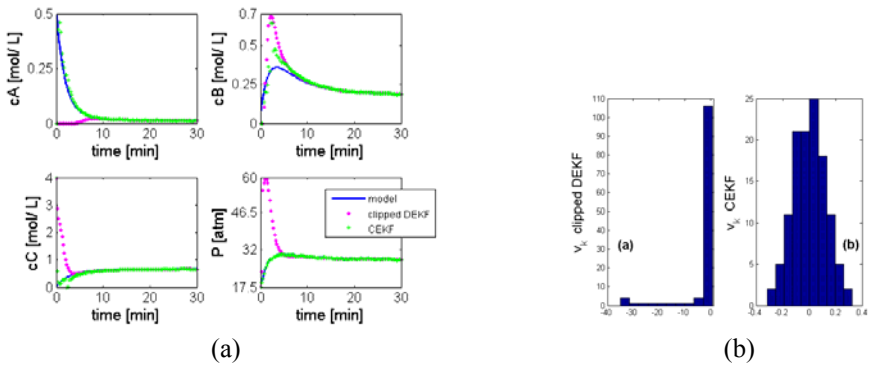


Fig. 3. Comparison between (a) clipped DEKF and CEKF performances and (b) the distribution of v_k for clipped DEKF (non-Gaussian) and CEKF (Gaussian).

Finally, we have compared the performances between CEKF and MHE (N=2). The high computational effort of MHE (2 orders larger) is not justified for this process as the improvement of estimation results is marginal, as shown in Fig. 4.
Due to the small computational effort of CEKF, possibility of incorporating constraints into an optimization problem, and the comparable estimation results to MHE, this approach seems to be the best choice for our case study.

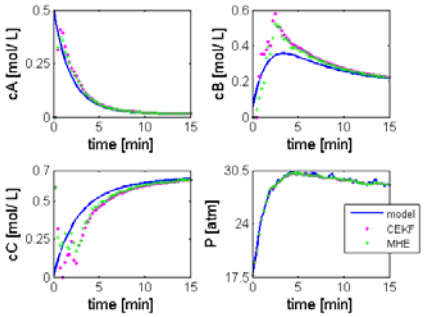


Fig.4. Comparison between CEKF and MHE (N=2) performances.

3. Conclusions

This paper outlines the numerical robustness of four EKF formulations for a chemical engineering example whose system model and measurement are such that multiple states satisfy the equilibrium condition. With a measurement noise perturbation, we outline a condition that has led a classical EKF formulation (DEKF) to converge to physically unrealizable state. According to our results, EKF-CRE and RREKF are more numerically robust in computing the state covariance than the DEKF. As both formulations have avoided an increase in error propagation due to a measurement noise perturbation, they were able to converge to the desired final state solution. Thus, a suited choice of the EKF formulation based on the state covariance equation is essential to prevent from physically unrealizable states. However, due to the poor initial states and its respective state covariance initial guess for the batch reactor, EKF-CRE and RREKF have not prevented from physically unrealizable states during the batch. Hence, CEKF can be seen as the best alternative to systems with such behavior due to the possibility of incorporating constraints into an optimization problem to prevent from physically unrealizable states. Besides, it is simpler, computationally less demanding than the MHE, and has comparable performance.

References

Brown, R.G.; Hwang, P.Y.C (1996). Introduction to Random Signals and Applied Kalman Filtering: With MATLAB Exercises and Solutions. John Wiley & Sons, U.S.A.

Gesthuisen, R.; Klatt, K. -U.; Engell, S. (2001). Optimization-based State Estimation: a Comparative study for the Batch Polycondensation. In Proceedings of ECC-2001.

Haseltine, E.L.; Rawlings, J.B. (2005). Critical Evaluation of Extended Kalman Filtering and Moving-Horizon Estimation. Ind. Eng. Chem. Res., 44, 2451-2460.

Hermann, R. and A. Krener (1977). Nonlinear controllability and observability. Automatic Control, IEEE Transactions, 22(5), 728–740.

Muske, K.R.; Rawlings, J.B. (1994). Nonlinear Moving Horizon State Estimation. NATO ASI Series, Kluwer Academic, 293, 349–365.

Rao, C.V.; Rawling, J.B.; Mayne, D.Q. (2003). Constrained State Estimation for Nonlinear Discrete-Time Systems: Stability and Moving Horizon Approximations. IEEE Trans. Autom. Control, 48(2), 246-258.

Simon, D. (2006). Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches. John Wiley & Sons, U.S.A.

Sontag, E.D. (1998). Mathematical Control Theory: Deterministic Finite Dimensional Systems. 2nd Edition, Springer, New York.

Valappil, J.; Georgakis, C. (2000). Systematic Estimation of State Noise Statistics for Extended Kalman Filters. AIChE J., 46(2), 292-398.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

IPL2&3 Performance Improvement Method for Process Safety Using the Event Correlation Analysis

Junya Nishiguchi, Tsutomu Takai

Yamatake Corporation, 1-12-2, Kawana, Fujisawa-shi, Kanagawa, Japan

Abstract

Alarm management efforts have recently intensified in the world, and there are many tools based on a guideline EEMUA191 to evaluate alarm system performance. However, the improvement and rationalization of alarm systems have not made satisfactory progress, because these tools usually just provide with alarm system performance metrics but not useful information to improve and rationalize alarm systems. In this paper, a novel method for performance improvement of IPL2&3 including alarm system using the event correlation analysis was proposed.

Keywords: EEMUA, IPL, alarm management, event correlation, clustering

1. Introduction

Safe operation is the top priority for process plants. As a concept of safe design to provide protection from hazardous incidents, Independent Protection Layers (IPLs) which consist of eight layers as shown in Table 1 has been extensively applied to the various plants (AIChE/CCPS 1993). The second layer (IPL2) and third layer (IPL3) are related to alarm system. The primary purpose of IPL2 is the supervision of the plant under the normal operation with Basic Process Control System. When the process variables deviate from the set points, the system activates alarms and requires the corrective actions of the operator. On the other hand, IPL3 represents the critical alarms and corresponding operator interventions. Failure in the functions of IPL2&3 is resulting to production loss. Therefore, it is important that IPL2&3 should be able to function effectively from both safety and production standpoint.

Table 1. Independent Protection Layers for Process Safety

Layers	Definitions	Functions	
IPL8	Community Emergency Response	Minimize damage from an incident	
IPL7	Plant Emergency Response		
IPL6	Physical Protection (Dikes)		
IPL5	Physical Protection (Relief Devices)		
IPL4	Automatic Action SIS or ESD		
IPL3	Critical Alarms, Operator Supervision, and Manual Intervention	Ensure safe and productive operations	Prevent an incident from happening
IPL2	Basic Controls, Process Alarms, and Operator Supervision		
IPL1	Process Design		

For rationalization of alarm system, EEMUA191 (EEMUA 1999) is widely accepted as a de facto standard guideline. In EEMUA191, alarm system is defined as *a very important way of automatically monitoring the plant condition and attracting the attention of the process plant operator to significant change that require assessment or*

action. It specifies several performance metrics that can be used to assess the alarm system performance:

- *Operator questionnaires*
- *Alarm usefulness surveys*
- *Assessment of number of alarms in a system*
- *Measurement of average alarm rate*
- *Measurement of number of alarms following a major plant upset*
- *Measurement of operator response time*
- *Measurement of number of standing alarms*
- *Analysis of the priority distribution of alarms configured and occurring*
- *Correlation techniques*

Unfortunately the effective correlation techniques from the standpoint of rationalizing IPL2&3 had not been defined well. Analyzing correlation between the alarms and operator actions in a process upset is expected to estimate the propagation path, source origin, and nuisance alarms, which can be useful information to improve the alarm system and understand the IPL2&3 performance properly.

2. New Method for IPL2&3 Performance Improvement

2.1. Concept

As referred to in EEMUA191, every alarm should be defined uniquely to notify a specific upset situation, and every operator action should be defined to solve the corresponding alarm. Once a certain upset causes alarms in the plant with the poor IPL2&3, it will propagate via the process fluid and consequently leads to the alarm flood; meanwhile, alarms without the corrective actions will increase operators' workload as nuisance alarms. Therefore, in order to rationalize IPL2&3 performance, every alarm and operator action should be properly related each other.

As a conventional method for fault propagation analysis, Signed Directed Graphs (Shiozaki et al. 1989) has been actively researched. However, it is difficult for this method to put into practical use because it requires much effort to adjust process model whenever the process devices are replaced. Also, to our knowledge there has not been any systematic method for rationalizing operator actions.

The concept of our method is to provide a fast, easy, and effective way to improve IPL2&3 performance. The relationship and occurrence order between events, which contain alarm and operator action events, are extracted from event log data with Event Correlation Analysis (Nishiguchi et al. 2005). In this method, the event pairs with constant occurrence time lags are considered to be similar events, since the time lags depend on the time delay of the process dynamics and human reaction time. From the similarity and the occurrence time lag calculated by the Event Correlation Analysis, we can estimate the group with the consequential events as well as their occurrence order. Our method enables us to extract the relationships between process variables change and operators' behavior only with accumulated log data. The solutions to improvement for the IPL2&3 design and management are found rapidly and easily. The examples of extracted relationships and corresponding solutions are as follows.

I. Consequential Alarms

When the several alarms are strongly related, these alarms are likely to be consequential alarms, which can be reduced by alarm filtering techniques.

II. Complex Operator Actions

When the several operator actions are strongly related, the operator actions are likely to be complex sequential operations, which can be reduced by automating operations.

III. Redundant Alarms

When there is a high frequent alarm without the related operator actions, it may be redundant alarm, which can be reduced by changing set point or replacing into message.

IV. Causes of upset

When constant occurrence order of alarms and operator actions is found, the first occurrence event is likely to be the source of plant upset.

2.2. Event Correlation Analysis

Event Correlation Analysis quantifies the relationship and occurrence order between alarms and operator actions. Although correlation coefficient is usually used to measure the relationship between two continuous values, it is well known fact that the coefficient cannot be applied for event data (Li 1990), such as alarms and operator actions log data. In this method, similarities between all event pairs are calculated from the log data with probability distribution of correlation regarding independent event pairs (Figure 2).

The alarms and operator actions log data obtained from Distributed Control System (DCS) contains occurrence time and event kind. The log data is converted into event time series $s_i(t)$ defined as binary series for each event i, which is 1 if the event occurs within the time window Δt and 0 if does not.

$$s_i(t) = \begin{cases} 1, & \text{if some points in } (n\Delta t, (n+1)\Delta t] \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where Δt is window size, and n is time unit.

The cross correlation function (2) between event i and j indicates the occurrence counts that event j follows event i in the time range of ($m\Delta t$, (m+1)Δt].

$$c_{ij}(m) = \begin{cases} \sum_{n=1}^{T/\Delta t - m} s_i(n) s_j(n+m) & m \geq 0, \ -K \leq m \leq K \\ c_{ji}(-m) & m < 0 \end{cases} \quad (2)$$

where T is observation time, and K is maximum lags of time unit.

From (2) maximum correlation value c^*_{ij} and time lag at maximum correlation m^*_{ij} are defined as follows.

$$c^*_{ij} = \max c_{ij}(m), \ m^*_{ij} = \arg\max_m c_{ij}(m) \quad (3)$$

As a result, the similarity between two events R_{ij} is defined with the probability that the correlation between two independent events is lower than c^*_{ij} within time lag K.

$$R_{ij} = P\left(c_{ij}(m) < c^*_{ij} \mid -K \leq m \leq K\right) \cong \left\{ \sum_{l=0}^{c^*_{ij}-1} \frac{\nu^l e^{-\nu}}{l!} \right\}^{2K+1} \quad (4)$$

where ν is the expected value of Poisson distribution regarding the occurrence of independent events, which is approximated by the average co-occurrence number between independent two events as shown in equation (5).

$$\nu = \frac{T}{\Delta t} \cdot p_i \cdot p_j \cong \frac{\Delta t}{T} \sum_{n=0}^{T/\Delta t} s_i(n) \cdot \sum_{n=0}^{T/\Delta t} s_j(n) \quad (5)$$

In other word, this method conducts the statistical test of hypothesis that two events i and j are generated independently. The actual maximum correlation c^*_{ij} is compared

with the distribution of correlation value between independent two events. The similarity R_{ij} is defined with the value, which is subtracted the reject rate from one.

By applying some clustering algorithm (e.g. hierarchical clustering) with the similarities of every event pair, the event groups with similar events will be extracted automatically. In addition, the occurrence order of each event pair is estimated with the time lag m^*_{ij}, which represents the most probable time delay.

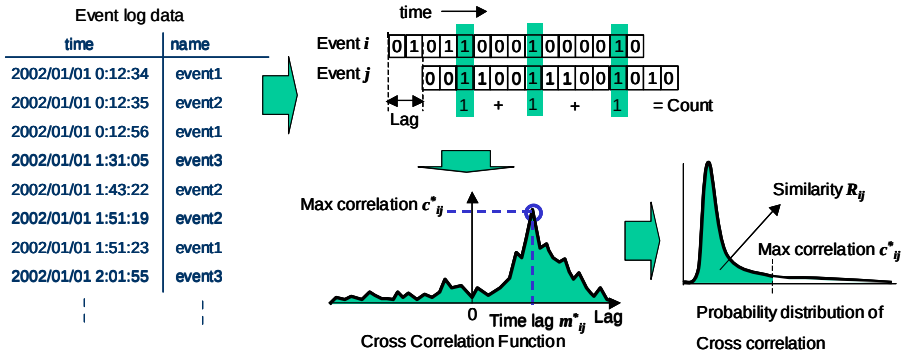


Figure 2. Event Correlation Analysis

2.3. Numerical Experiment

We evaluated the Event Correlation Analysis with synthetic data consists with six event type generated with the following rules for 30 days in one-second units.

- Event 1 and 4 are generated by Poisson distribution with mean of 0.0001 and 0.0004, respectively.
- Event 2 and 3 are generated from event 1, added with noise of Poisson distribution with mean of 0.000005 and removed with the probability 50%. In addition, occurrence time of each event is shifted with Normal distribution $N(600,100)$ and $N(1800,100)$, respectively.
- Event 5 and 6 are generated from event 2, added with noise of Poisson distribution with mean of 0.0002 and removed with the probability 50%. In addition, occurrence time of each event is shifted with Normal distribution $N(600,100)$ and $N(1800,100)$, respectively.

Table 3. Results for the synthetic data

Order i –> j	Similarity c^*_{ij}	Delay m^*_{ij}	Order i –> j	Similarity c^*_{ij}	Delay m^*_{ij}
1 –> 2	0.99	600	4 –> 5	0.98	600
1 –> 3	0.98	1800	4 –> 6	0.97	1900
2 –> 3	0.96	1200	5 –> 6	0.95	1100

Table 3 shows the similarity and delay time of the related event pairs with higher similarity than 0.9. From the table, we can see the proposed method properly extracted the related event pairs and their occurrence time lags even though the data contained noise.

3. Validation with Actual Plant Data

3.1. Target Plant

The proposed method was applied for actual chemical plant data for validation. The target plant has the typical multiple-stage gas purification unit. According to the event log data obtained from this unit, a total of 1,267 types of alarms and operator actions occurred a total of 56,350 times over a period of two months.

3.2. Results

The event log data was analyzed with Alarm Analyst R20, the software product that implements the Event Correlation Analysis. The software generated the groups with the consequential events for a few seconds. From the discussions with the plant engineers and operators, we were able to plan the solutions for IPL2&3 performance improvement regarding the top 35 groups accounted for 60% of occurrences within only six hours. We will next give some examples of the results, which show actual process property and operators' behavior correctly.

I. Consequential alarms

The five events listed in Figure 4 were occurred synchronously and the total count was 702. From the result of the Event Correlation Analysis, we found Alarm 1, 2, 3 and 4 followed after Operation 1 at around midnight everyday. From the interview with the operators, when they switched the blower to the manual operation, Alarm 4 notified them of pipe blowing. As a result, the air and steam flow rate became unstable and consequently Alarm 1, 2 and 3 occurred. Since these alarms did not have corresponding operator actions in this situation, it was decided that Alarm 1, 2 and 3 were inactive with alarm filtering technique during manual pipe blowing procedures. In addition, Alarm 4, which gave only guidance to the operators, was changed to a message.

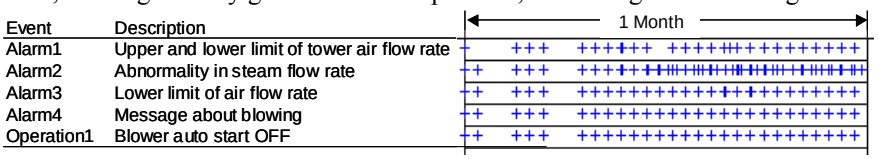


Figure 4. Example of consequential alarms

II. Complex operator actions

The five events listed in Figure 5 were also occurred synchronously and the total count was 495. From the result of the Event Correlation Analysis, we found Operation 2, 3 and Alarm 6 occurred simultaneously at around 4:00am, and Operation 2, 4 and Alarm 5 occurred simultaneously at around 4:00pm. The investigation of the daily report found that these events are related to the pipe flushing. In the system A, operators firstly invalidated alarms with Operation 2, followed by flushing procedure with Operation 3 and the notify of pipe flushing with Alarm 6. Likewise, in the system B, operators firstly invalidated alarms with Operation 2, followed by flushing procedure with Operation 4 and the notify of pipe flushing with Alarm 5. In general, operators should not intervene with such regular procedures. Therefore, Operation 2, 3 and 4 were automated with the timer, and Alarm 5 and 6 were changed to message.

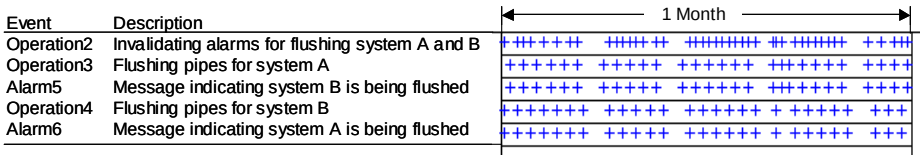


Figure 5. Example of complex operator actions

III. Redundant alarms

Alarm 7 is a flow rate alarm with lower limit type and occurred 657 times. According to the result of the Event Correlation Analysis, there were not any alarms or operator actions related to Alarm 7. Thus, this alarm is considered to occur independently with other alarms and operator actions. In fact, the operators told they usually did not take any actions in response to this alarm, but waited until it returned to stable by itself. We removed the alarms like this alarm that did not function as alarm, because they could deteriorate the sensitivity of the operators to alarms.

IV. Cause of upset

Figure 6 shows the example of the cause of plant upset. As shown in Figure 6, these six events occurred at the devices located near each other. From the event log data, Operation 5 was a manipulated variable change action, Operation 6 was an upper threshold change action, and Alarms 8, 9, 10 and 11 were all upper limit alarms. The occurrence order estimation found that Alarm 8 occurred first and Operation 6 occurred last. According to these results, it must be a problem with the pipes near the heat exchanger subjected to cooling with seawater. In fact, a field investigation revealed that the pipe easily became clogged, resulting in poor seawater cooling. Therefore, it was decided that this pipe should be cleaned regularly to solve this problem.

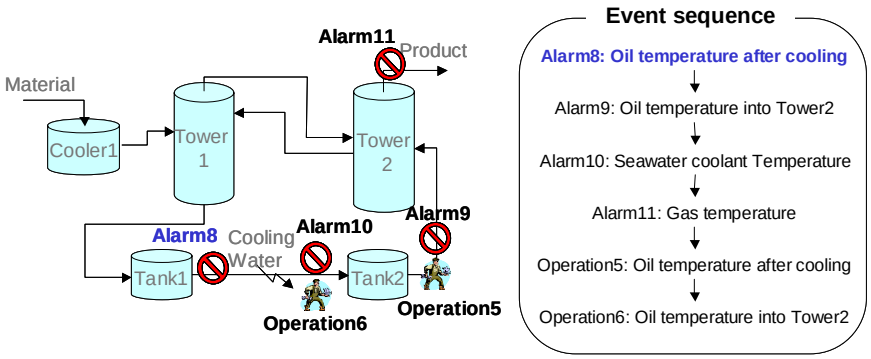


Figure 6. Example of cause of upset

4. Conclusion

In this paper, a new method for IPL2&3 performance improvement using Event Correlation Analysis was proposed. The validation result with the actual plant data shows our method is considered to improve process safety as well as productivity rapidly, easily, and effectively. In addition, since this method extracts the operators' behavior from the log data, it leads to standardization of operator actions.

References

AIChE/CCPS, Guidelines for Engineering Design for Process Safety, AIChE, 1993

Engineering Equipment & Materials Users' Association, Alarm Systems. A Guide to Design, Management and Procurement, EEMUA Publication No. 191, 1999

J. Shiozaki, B. Shibata, H. Matsuyama, and E. O'Shima, Fault Diagnosis of Chemical Processes Utilizing Signed Directed Graphs—Improvement by Using temporal Information, IEEE Transactions on Industrial Electronics, vol. 36, no. 4, pp. 469–474, 1989.

J. Nishiguchi and H. Tsutsui, 2005, A New Approach to Process Alarm Reduction Using Statistical Point Processes, SICE Annual conference 2005, pp.443-448

W. Li, Mutual Information Versus Correlation Functions, Journal of Statistical Physics, vol. 60, pp. 823-831, 1990

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Collaborative Multi - Agent based Process Monitoring System for Offshore Oil and Gas Production

Sathish S Natarajan, Kaushik Ghosh, Rajagopalan Srinivasan*

Department of Chemical and Biomolecular Engg, National University of Singapore
10 Kent Ridge Crescent, Singapore, 119260.

Abstract

Offshore oil and gas production platforms are uniquely hazardous in that operating personnel have to work in a perilous environment surrounded by extremely flammable hydrocarbons. A failure in an equipment could quickly propagate to others resulting in leaks, fires and explosions, causing loss of life, capital invested and production downtime. A method for preventing such accidents is to deploy intelligent monitoring tools which continuously supervise the process and the health of equipments to provide context–specific decision support to operators during safety-critical situations. A dynamic model of an offshore oil and gas production platform was developed using gPROMS and data to reflect operating conditions under normal, fault conditions and maintenance activities were simulated. These data are used to train three monitoring algorithms based on multivariate statistics (Principal Component Analysis), two of which are specialized in monitoring certain sections of the platform. These multivariate monitoring algorithms are considered as individual agents and the results produced by each are then integrated using a multi-agent collaborative framework. A consolidator agent, which uses voting based, Bayesian probability and Dempster Shafer fusion strategies for conflict resolution and decision fusion is developed. The ability of this agent based monitoring scheme to detect and diagnose faults in a more precise manner than any single FDI agent in offshore oil and gas production platforms is demonstrated.

Keywords: offshore production, process monitoring, agent based.

1. Introduction

Offshore Oil and Gas production has accounted for most of the increase in worldwide production of hydrocarbons over the last few decades. Given that most of the significant hydrocarbon discoveries in recent times have happened offshore, it is expected that the contribution of offshore production to our overall energy mix would only increase further. Hence it is imperative that these offshore production platforms are operated in a safe and efficient manner.

Operations personnel in offshore platforms face many unique challenges: The workforce lives and works in a restricted location for significant periods of time without a break. The environment is characterized by confined space, constant noise and activity while the production process has the potential for the rapid escalation of hydrocarbon related incidents. This inherently hazardous nature combined with the extremely harsh

*Corresponding Author
Ph: +65 65168041, Email: chergs@nus.edu.sg

and stressful environment is the prime reason for the unusually large number of accidents. A study by the UK Health and Safety Executive shows that process related failures account for nearly 80% of the risk to offshore personnel. Currently, onboard operating personnel are expected to detect failures as and when they occur and take corrective action to prevent them from escalating into severe accidents. However given the cognitive limitations of operators working in a remote facility some failures are not detected.

A method for preventing accidents on offshore platforms is to deploy intelligent monitoring tools which continuously supervise the process and the health of the equipments to provide context – specific decision support to operators during safety critical situations. Historically, the offshore processing industry has sought to improve its safety performance using the proactive strategy in which the basic process design is made inherently safe and robust. The main shortcoming of this strategy is that even if all the hazards have been identified, it may not be economically or technically feasible to eliminate them. The reactive strategy seeks to overcome this through improved process control using advanced process monitoring and supervision. There is plenty of literature on process monitoring, fault detection and identification (Venkatasubramaniam et al, 2003) for petrochemical industries but these monolithic monitoring strategies are not suited for the offshore production process due to the lack of large scale computational facilities on board. Hence, a remote agent based distributed monitoring scheme would be more suited. Also, due to the complexity involved a single FDI agent would not be capable of handling all the fault scenarios. A method in which the results of several FDI agents, each specialized in monitoring certain sections of the overall process, are meaningfully fused by a consolidator agent, is proposed.

Since relevant process data for offshore platforms is not readily available in literature a dynamic model in gPROMS is developed, as detailed in Section 2. Process faults common in offshore production platforms (Khan et al, 2002) is introduced into this model. The operations data collected from this model was used to train three different multi-variate statistical models, each specialized in identifying a fault in a particular section of the platform, as discussed in Section 3. The results from these individual methods are then combined using different fusion strategies.

2. Dynamic Model of Offshore Natural Gas Production

The first step towards achieving a representative dynamic model is always a good steady state simulation. Hence a steady state model of a typical natural gas production platform was developed in Hysys. The steady state operations data from the Hysys model served as the initial conditions to the dynamic model in gPROMS. Three flow lines involving the usual combination of safety, master, wing and choke valves were modeled. The flow lines were combined into a production header and a separator header. As shown in Figure 1, a test separator is included which is brought online occasionally when a stream has to be analyzed. An important maintenance activity in the operation of these platforms is the pig-launching operation, during which the platforms are required to be shut down; hence shutdown and startup procedures were incorporated. Other maintenance activities include corrosion inhibitor and kinetic hydrate inhibitor injection. Apart from capturing the dynamics during these normal operations, three fault cases were simulated as indicated in Figure 1.

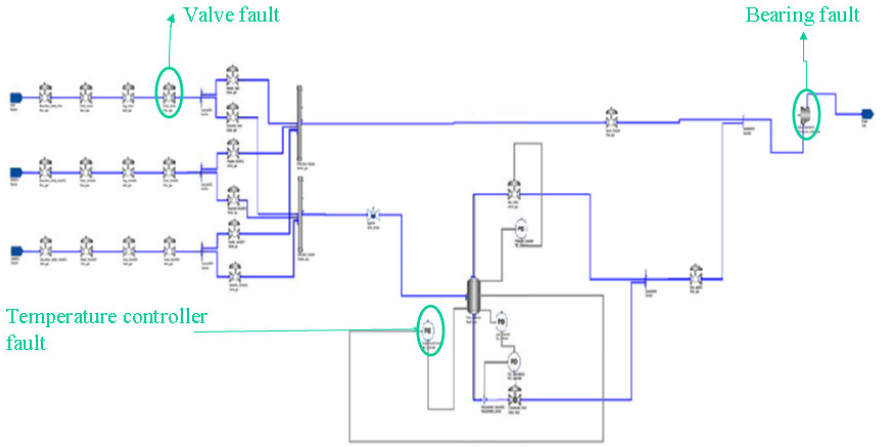


Figure 1. Process Flowsheet of Offshore Natural Gas Production from industrial case study

(1) A valve leak fault in a flow line in which the flow through the line was reduced from its normal range by 10% as a step change.
(2) Test separator temperature controller fault in which during the operation of the separator the heat input was reduced resulting in a temperature drop by 2%.
(3) A bearing fault on a compressor in which the eccentric vibrations and currents were simulated. This fault was not simulated as a complete failure of the bearing, but as an incipient fault like a dent in the raceways producing eccentricities.

Figure 2 shows the simulated data for these three faults along with their corresponding expected normal behavior.

3. Multivariate statistical models for monitoring

Three different models were developed as individual agents. The three agents were a process monitoring agent, which monitors the overall process, a specialized separator monitoring agent and a specialized compressor monitoring agent. The first two were based on principal component analysis (Jackson, 1991) while the third is based on spectral principal component analysis.

3.1. Process Monitoring Agent

The normal operating data from 34 variables measured at various points in the process were used to train a PCA model. The data from the three fault conditions were then projected onto this PCA model. The valve leak fault is detected and diagnosed with a delay of 14 samples. The temperature controller fault is detected with a delay of 4 samples and diagnosed after 16 samples. This process monitoring agent does not detect the bearing fault on the compressor.

3.2. Separator Monitoring Agent

Of the 34 variables measured, 5 were chosen around the separator. A separate PCA model is developed using normal data for these 5 variables and the fault data projected onto this. Being specially suited to detect abnormalities around the separator this agent detects and diagnoses the temperature controller fault with a delay of only 4 samples and is unable to detect the bearing fault around the compressor.

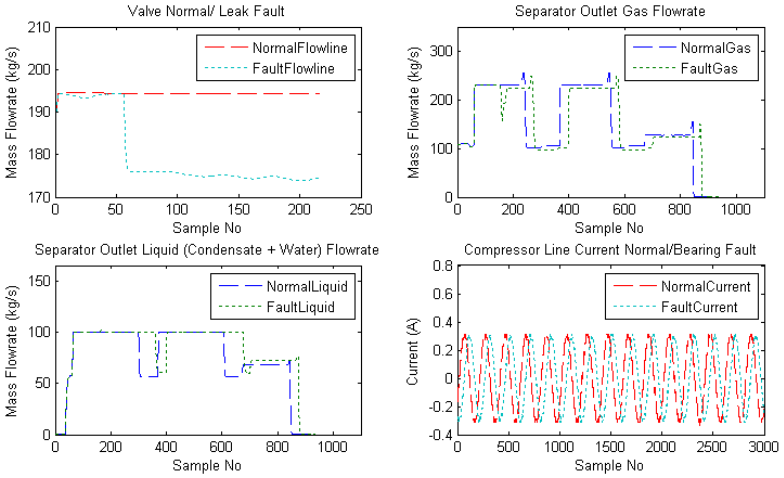


Figure 2. Simulated data for normal operating condition as well as Valve Leak Fault, Separator Temperature Controller Fault and Compressor Bearing Fault

3.3. Compressor Monitoring Agent

The eccentricities produced by the bearing fault are oscillatory in nature. Hence a Spectral PCA method is used for monitoring the compressor. The novel feature of spectral PCA is that the rows of the data matrix *X*, are the single sided power spectra P(*f*) of the signals over a range of frequencies (Thornhill et al, 2002). The data matrix *X* is then decomposed in the usual PCA procedure and the T^2 and Q statistic measured. 5 variables were chosen around the compressor including the simulated vibration and current signals. This agent is capable of detecting and diagnosing the bearing fault alone after 11 samples.

4. Collaborative Multi-agent based fusion

As is evident from Table 1, no single FDI method is capable of detecting all the faults. Instead of developing a suitable monolithic method, the results from these individual methods could be meaningfully combined using a consolidator agent to produce good results. This consolidator agent utilizes voting, Bayesian, or Dempster Shafer fusion strategies. Using such methods, the strengths of the individuals agents are accrued while their drawbacks are overcome resulting in improved performance.

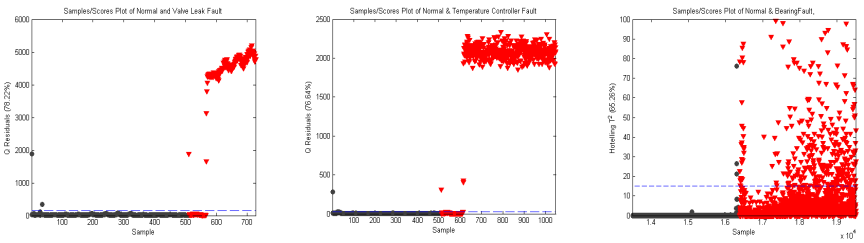


Figure 3. Sample/Score plot of Normal vs. Valve Leak Fault, Separator Temperature Controller Fault and Compressor Bearing Fault

Table 1. Performance of the individual agents, F1 - Valve Leak Fault, F2 - Separator Temperature Controller Fault, F3 – Compressor Bearing Fault

Fault	Process Monitoring PCA		Separator Monitoring PCA		Compressor Monitoring Spectral PCA	
	Detection Delay (x30s)	Diagnosis Delay (x30s)	Detection Delay (x30s)	Diagnosis Delay (x30s)	Detection Delay (x30s)	Diagnosis Delay (x30s)
F1	14	14	22	-	-	-
F2	4	16	4	4	-	-
F3	-	-	-	-	11	11
Average Delay	9	15	13	4	11	11
Recognition rate (%)	62.23		50.39		47.42	

Voting based fusion is the simplest and most commonly used decision fusion strategy. The class assigned to each sample by an FDI agent is considered as a vote for that class. The class with the maximum number of votes is the combined prediction. The major drawback of this method is that all agents are treated equally although in practice some agents outperform others and should hence be given more weightage.

Both Bayesian and Dempster-Shafer strategies utilize the prior performance of an agent for combination. The prior performance of a FDI agent is recorded in the form of a confusion matrix in which the rows are the actual classes while the columns are the assigned classes. For each FDI method, P(Ci/Cj) is the probability of assigning a sample to class Ci given class Cj. For k FDI methods, Bayesian Belief Value for class Ci is calculated as

$$Bel(C_i) = \frac{\prod_k (P_k(C_i / C_j))}{\sum_i \prod_k (P_k(C_i / C_j))}$$

The class with maximum Bayesian belief value, *Bel(Ci),* is the assigned class upon combination using Bayesian fusion strategy.

Table 2. Performance of different fusion strategies, F1 – Valve Leak Fault, F2 – Separator Temperature Controller Fault, F3 – Compressor Bearing Fault

Fault	Voting Based Fusion		Bayesian Probability based Fusion		Dempster Shafer based fusion	
	Detection Delay (x30s)	Diagnosis Delay (x30s)	Detection Delay (x30s)	Diagnosis Delay (x30s)	Detection Delay (x30s)	Diagnosis Delay (x30s)
F1	22	-	14	14	14	22
F2	4	16	4	4	4	4
F3	-	-	11	11	11	11
Average Delay	13	16	9.67	9.67	9.67	12.33
Recognition rate (%)	49.53		97.85		97.17	

The Dempster-Shafer (D-S) method is popular in uncertainty reasoning which is used to combine separate pieces of evidence by assigning a belief value. The difference between Bayesian and D-S is the fact that the Bayesian method precisely assigns a probability value to each class, whereas D-S method assigns a BPA (Basic Probability Assignment) value to the union of all possible classes. The combined BPA values are obtained by using D-S combination rule.

$$Com\_BPA(C_i) = \frac{\sum_{A \cap B = C_i} BPA_I(A)\, BPA_{II}(B)}{1 - \sum_{A \cap B = \phi} BPA_I(A)\, BPA_{II}(B)}$$

Where Ci is the subset formed by the intersection of A and B and ϕ is null or empty set. The subset with highest combined BPA value is the resulting class upon combination.

The individual agents have recognition rates, which is defined as the percentage of sample correctly classified in the range of 47-62% shown in Table 1. Upon combination using Bayesian and Dempster Shafer based fusion strategies the recognition rates are improved to around 97% shown in Table 2. Voting based scheme does not perform as well as it requires at least 2 of the agents to concur for collaboration. We are able to achieve complete fault coverage and also improvement in fault detection and diagnosis time by using Bayesian and Dempster Shafer combination strategies.

5. Conclusion

An intelligent collaborative agent based monitoring system was proposed in this paper. Using data simulated from a dynamic gPROMS model three monitoring agents were developed, two of which were specialized. The results from these individual agents were then combined using different strategies. Upon combination we achieve significant (from about 60% to 97%) improvement in fault recognition rates and reduction in fault detection and diagnosis time. Thus, we have demonstrated that by using fusion we are able to combine the advantages of the individual methods while their respective shortcomings are overcome resulting in improved monitoring performance for offshore oil and gas production. Our future work would be to develop a remote agent based monitoring strategy, building upon these results for the offshore oil and gas industry, the first step of which would be the development of an ontological framework for efficient agent interaction.

Acknowledgement

The authors are grateful for the financial support from Maritime and Port Authority of Singapore.

References

1. Health and Safety Executive (HSE), 1996. Offshore accident/incident statistics reports. OTO96.954, London, UK.
2. V. Venkatasubramanian., R. Rengaswamy., S.N. Kavuri. 2003, A review of process fault detection and diagnosis , Computers and Chemical Engineering, Vol 27,3.
3. N.F.Thornill, S.L, Shah, B.Huang, A.Vishnubotla, 2002. Spectral principal component analysis of dynamic process data. Control engineering Practice 10, 833-846.
4. F.I. Khan., R. Sadiq., T. Husain., 2002, Risk-Based process safety assessment and control measures design for offshore process facilities., Journal of Hazardous Materials A94, 1-36.
5. J.E. Jackson., 1991, A User's Guide to Principal Components, Wiley, New York.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Real-time Moving Horizon State and Parameter Estimation for SMB Processes

Achim Küpper[a], Moritz Diehl[c], Johannes P. Schlöderl[b], Hans G. Bock[b], Sebastian Engell[a]

[a]*Process Dynamics and Operations Group, Technische Universität Dortmund, Emil-Figge-Str. 70, 44221 Dortmund, Germany*
(e-mail: achim.kuepper@bci.tu-dortmund.de)
[b]*IWR - Interdisciplinary Center for Scientific Computing, Universität Heidelberg, Germany*
[c]*Electrical Engineering Department (ESAT-SCD), K.U. Leuven, Belgium*

Abstract

Advances in numerical algorithms have rendered the application of advanced process control schemes feasible for complex chemical processes that are described by high-order first-principles models. Applying real-time iteration schemes reduces the CPU requirement such that rigorous models can be applied that enable a precise forecast of the system behaviour. In this paper, a moving horizon state and parameter estimation scheme for chromatographic simulated moving bed SMB processes is presented. The simultaneous state and parameter estimation is based on a high-order nonlinear SMB model which incorporates rigorous models of the chromatographic columns and the discrete shifting of the inlet and outlet ports. The estimation is performed using sparse measurement information: the concentrations of the components are only measured at the two outlet ports (which are periodically switched) and at one fixed location between two columns. The goal is to reconstruct the full state of the system, i.e. the concentration profiles along all columns, and to identify model parameters reliably. The state estimation scheme assumes a deterministic model within the prediction horizon, state noise is only present in the state and in the parameters prior to and at the beginning of the horizon. The scheme can be applied online. The advantage of this estimation scheme is that it is applicable to all process scenarios encountered during the real operation of an SMB plant, e.g. start up, transition periods, varying flows and switching times, since no model simplification nor a state reduction scheme are applied. Numerical simulations (start up of the SMP process) of a validated model for a separation problem with nonlinear isotherms of the Langmuir type demonstrate the efficiency of the algorithm.

Keywords: Simulated Moving Bed chromatography, Moving horizon estimation, State estimation, Model identification, Real-time application, Real-time iteration.

1. Introduction

The Simulated Moving Bed (SMB) process is an efficient chromatographic separation technology that is increasingly applied in the food, fine chemicals, and pharmaceutical industries. A SMB process is realized by connecting several chromatographic columns in a closed loop as illustrated by Fig. 1. Since SMB processes and its variants are characterized by mixed discrete and continuous dynamics, spatially distributed state vari-

ables with steep slopes, and slow and strongly nonlinear responses of the concentrations profiles to changes of the operating parameters, they are difficult to control. In the literature, relatively few contributions that deal with state estimation of SMB processes can be found. The published work is either based upon simplifications or deals with the engineering of tailored estimation schemes. Recently, a rigorous moving horizon estimation approach for SMB processes was proposed in [2]. In this formulation of the MHE, a deterministic behaviour of the process on the estimation horizon and Gaussian independent identically distributed measurement noise are assumed. The initial state at the beginning of the horizon and its covariance are computed by an Extended Kalman Filter (EKF). The state noise covariance and the initial error covariance of this EKF are the only tuning parameters of this scheme. A fast online solution of the underlying constrained least-squares optimization problem is obtained by using the direct multiple shooting method [3]. A full rigorous model is applied and therefore no assumption that the plant is close to the periodic steady state is needed. Along with the states, key adsorption parameters are estimated online. Simulations demonstrate that the states and critical model parameters can be reconstructed successfully. The computation times are such that the estimator can be applied online. Since a rigorous full scale SMB model is used, the MHE approach of [2] can also handle transition periods. In this contribution, the moving horizon state and parameter estimation scheme is applied to reconstruct the state variables in the start-up phase.

2. The Simulated Moving Bed Process

Chromatographic separation is based on the different adsorption affinities of the molecules in the liquid to an adsorbent which is packed in a chromatographic solid bed. A counter-current movement between the liquid and the adsorbent is implemented by switching the ports in the direction of the liquid flow periodically, illustrated by Fig. 1.

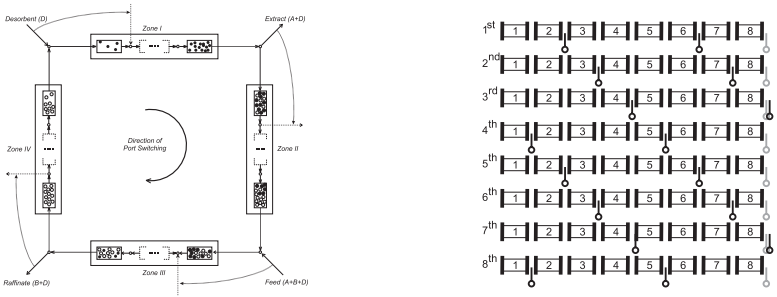


Fig. 1: Principle of the Simulated Moving Bed process (left) and positions of the extract, raffinate, and recycle port measurements over a cycle of operation (right)

In this paper, the counter-current flow of the solid and the liquid phases is modeled as it is achieved in the real plant by periodically switching the inlet and outlet ports in the direction of the liquid flow after period m with period length τ has passed. The state variables thus represent the concentrations in the physical columns and do not exhibit jumps. Only the input flow rates and the inflow concentrations change discontinuously. The dynamic simulation of the SMB process is achieved by integrating the differential equation over period m

$$\dot{x} = f(x, Q_m, p) \quad t \in [(m-1)\tau, \tau]; \qquad x(t_0) = x_{m,0} \tag{1}$$

followed by the switching of the flows $Q_{m,j}$:

$$Q_{m+1,j} = M_Q Q_{m,j} \qquad j = De, Ex, Fe, Ra, \mathrm{Re} \tag{2}$$

with differential states x and parameters p. The vector $Q_{m,j}$ defines the inlet/outlet flow of port j (desorbent, extract, feed, raffinate) and the recycle stream at period m. The respective component of $Q_{m,j}$ represents the flow of the respective port j to the respective column. M_Q is a permutation matrix that shifts the flow ports after period m has passed. The recycle flow that defines the total flow rate in the zone in front of the desorbent port is also switched in the same manner. All zone flow rates result from the port flows and the recycle flow. For this paper, three positions where the concentrations of the two substances of the mixture are measured are assumed. The measurements are positioned behind the extract port, behind the raffinate port, and behind one column in the process where physically the closing of the loop is realized. The extract and raffinate concentration measurements move together with the ports while the third measurement remains at its physical location throughout. Fig. 1 illustrates the measurement positions over a full cycle of operation that is defined by the time that passes until the ports are back at the starting positions after having moved around the closed loop of columns. The positions of the inlet and outlet ports are then repeated periodically. From mass balances of the components around the inlet and the outlet ports, the internal flow rates and the inlet concentrations can be calculated according to:

$$\begin{aligned}
&\textit{Desorbent node}: && Q_{IV} + Q_{De} = Q_I;\; c_{i,IV}^{out} Q_{IV} = c_{i,I}^{in} Q_I \qquad i = A,B \\
&\textit{Extract node}: && Q_I - Q_{Ex} = Q_{II} \\
&\textit{Feed node}: && Q_{II} + Q_{Fe} = Q_{III};\; c_{i,II}^{out} Q_{II} + c_{i,Fe} Q_{Fe} = c_{i,III}^{in} Q_{III} \qquad i = A,B \\
&\textit{Raffinate node}: && Q_{Ra} + Q_{IV} = Q_{III}
\end{aligned} \tag{3}$$

where $Q_{I\text{-}IV}$ are the flow rates in the corresponding zones, Q_{De}, Q_{Ex}, Q_{Fe}, and Q_{Ra} denote the external flow rates and $c_{i,in}$ and $c_{i,out}$ denote the concentrations of the component i in the streams leaving and entering the respective zone. The column distribution over the four separation zones considered here is 2/2/2/2, i.e. each zone consists of two columns. The chromatographic columns are modelled by the General Rate Model where it is assumed that there are no radial gradients in the column and that the particles of the solid phase are uniform, spherical, porous (with constant particle porosity ε_p), and that the mass transfer between the particle and the surrounding layer of the bulk is in a local equilibrium. The concentration of component i is denoted by c_i in the liquid phase and by q_i in the solid phase. D_{ax} is the axial dispersion coefficient, u the interstitial velocity, ε_P the void fraction of the bulk phase, $k_{l,I}$ the film mass transfer resistance, and D_p the diffusion coefficient within the particle pores. The concentration within the pores is denoted by $c_{p,i}$. The following partial differential equations of a column can be derived from a mass balance around an infinitely small cross section area of the column with a constant radial distribution of the interstitial velocity u and the concentration c_i.

$$\frac{(1-\varepsilon_b)3k_{l,i}}{\varepsilon_b}(c_{b,i} - c_{b,i|r=R_p}) = D_{ax}\frac{\partial^2 c_{b,i}}{\partial^2 z} - u\frac{\partial c_{b,i}}{\partial z};\quad (1-\varepsilon_p)\frac{\partial q_i}{\partial t} + \varepsilon_p \frac{\partial c_{p,i}}{\partial t} = \left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \frac{\partial c_{p,i}}{\partial r}\right)\right], \tag{4}$$

with appropriate initial and boundary conditions. It is assumed that the concentration q_i is in thermodynamic equilibrium with the liquid concentrations in the particle and their relationship can be described by an adsorption isotherm. In the application considered here, the adsorption isotherm is of extended Langmuir type

$$q_i = H_i^1 c_{p,i} + \frac{H_i^2 c_{p,i}}{1+\sum_k k_k c_{p,k}}, \quad i = A,B, \tag{5}$$

with H_i^j and k_i as isotherm constants.

The resulting system of coupled differential equations can be efficiently solved by the numerical approach proposed in [4] where a Galerkin finite element discretization of the bulk phase is combined with an orthogonal collocation of the solid phase. As a result,

the initial values, boundary values, and partial differential equations (PDE) are transformed into a set of initial values and a system of ordinary differential equations (ODE) $\dot{x} = f(x,u,p)$ of order 352 (for n_{fe}=5, n_c=1, number of components n_{sp}=2, and number of columns n_{col}=8) where the flows Q are summarized in the input vector u. The system output is defined as $y=h(x,u,p)$.

3. Moving Horizon Estimation in SMB Processes

For the simultaneous estimation of the states and the parameters of SMB processes, we employ the Moving Horizon Estimation scheme introduced by Diehl, Schlöder, and Bock [2], which is modified in order to handle the shift of the inputs and of the measurements of the SMB process. The Moving Horizon Estimator estimates the states and the parameters based on the past measurements at specific time points that are located in the horizon $T_N=t_K-t_L$. t_K represents the current time and t_L is the time at the beginning of the horizon. A least-squares minimization is performed that minimizes the deviations of the real measurements η_k from the simulated measurements $h(x;t_k,u,p)$ at times t_k. The optimization problem of the MHE results as:

$$\min_{x(t_L),p} \left\{ \left\| \begin{matrix} x(t_L)-\bar{x}_L \\ p-\bar{p}_L \end{matrix} \right\|^2_{P_L^{-\frac{1}{2}}} + \sum_{k=L}^{K} \left\| \eta_k - h(x(t_k),u(t_k),p) \right\|^2_{V_k^{-\frac{1}{2}}} \right\} \quad \begin{matrix} s.t. \quad \dot{x}(t) = f(x(t),u(t),p) \quad (6) \\ x_{\min} \le x \le x_{\max}; \quad p_{\min} \le p \le p_{\max}; \quad t \in [t_L, t_K] \end{matrix}$$

The second term represents the prediction errors within the horizon and the first term represents the arrival cost (the penalization of a change of the estimates of the initial values of the states and of the parameters). Note that only the initial values of the states and the parameters are free parameters of the optimization problem because no state noise is assumed within the horizon, which is equivalent to the assumption of a dynamic model that is perfect in its structure. The absence of state noise on the horizon is compensated by the simultaneous estimation of key model parameters which is an appropriate assumption since uncertainties are mostly due to model errors and not to disturbances. From the solution $x(t_L)$ and p of the optimization problem, the deterministic model is simulated forward to obtain the current estimated state x_K. The MHE takes upper and lower bounds on the states and on the parameters into account. The means $\bar{x}(t_L)$ and $\bar{p}$ and the covariance P_L in the arrival cost are determined by a smoothed Extended Kalman Filter that is implemented using QR-decompositions in order to guarantee positive definite matrices despite possible numerical errors [5]. In order to enable a smooth gradient calculation with respect to the simulated measurements which are subject to jumps due to the periodic movement of the ports, virtual measurements at constant positions at the outlet of each chromatographic column are included in the mapping of h. In order to account for the actual existence of real measurements at the considered time point k, the corresponding components on the diagonal of the measurement weight V_k^{-1} are set to σ_v^{-1} while nonexisting measurements have zero entries on the diagonal of V_k^{-1}. A zero weight can be interpreted as infinite measurement noise. The last two entries of V_k are the variances of the measurements at the internal measurement position (recycle) which are not shifted.

3.1. Multiple-Shooting Real-Time Iteration Scheme for MHE

The moving horizon optimization problem is solved by the multiple shooting methods that subdivides the time horizon into subintervals and formulates autonomous initial value problems on each individual subinterval. To retain the original problem, the subproblems are coupled by continuity conditions to guarantee the continuity of the state trajectories. The computational requirements are largely reduced by applying the *real-*

time iteration scheme for the multiple shooting method introduced by [1] that updates all relevant sensitivity matrices that are necessary to solve the optimization problem before the most recent measurement η_K is available. Another important feature is that the next optimization problem is initialized well at the current solution such that the number of iterations can be reduced to one.

3.2. Results (start-up scenario)

The separation of the enantiomer mixture EMD-53986 into the pure components is considered. The parameters of the SMB model were taken from [6].The measurements are corrupted by noise with a standard deviation of *0.025 g/l*. No cross-correlations between the state noises and between the state noises and the parameters were assumed. The same noise variances were assumed for each state. The tuning of the moving horizon estimator was performed by varying the covariances of the state variables and of the free parameters. The weighting matrix W incorporates a state standard deviation of 0.00433 *g/l* and a parameter standard deviation of 0.0316 for H^2_A and 0.0265 for H^2_B: $W^{0.5}=diag(0.0043,\ldots,0.0043,0.0361,0.0265)$. The initial weight P_0 is set to 0.005 W. The chosen state and parameter noises represent a compromise between the smooth estimation of the states and a quick adaptation of the parameters. The sampling time of the estimator is 1/10 of the period length. The moving horizon length is five sampling intervals (half a period). In the simulation scenario, the reference model and the model of the estimator are initialized with empty columns.The output, parameters, and states are reconstructed correctly, as can be seen by figures 2 to 4. The estimation of the parameters does not interfere with the reconstruction of the states. The MHE estimator can be applied online, as can be seen from Fig. 3. The CPU times are below the sampling rate at all sampling points. The CPU times of the MHE are around 19 s on a standard PC (Intel Xenon CPU 2.8 GHz, 4.0 GB RAM), the maximum and minimum values being 25.0 s

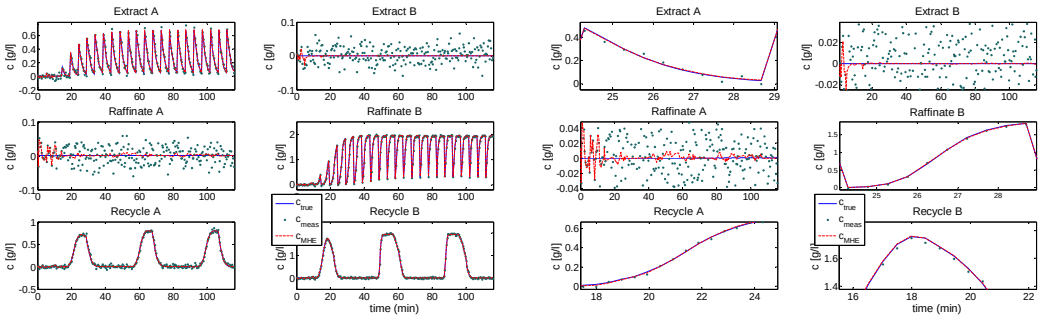


Fig. 2: Measurements (extract, raffinate, recycle); enlarged (right)

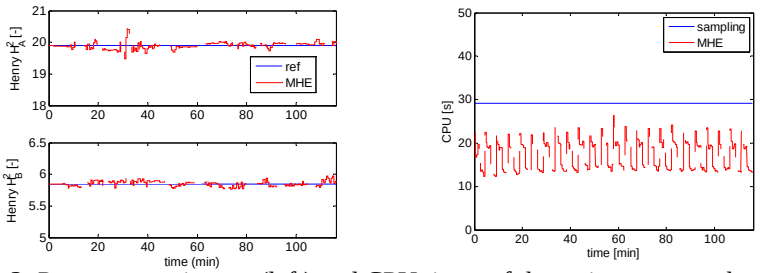
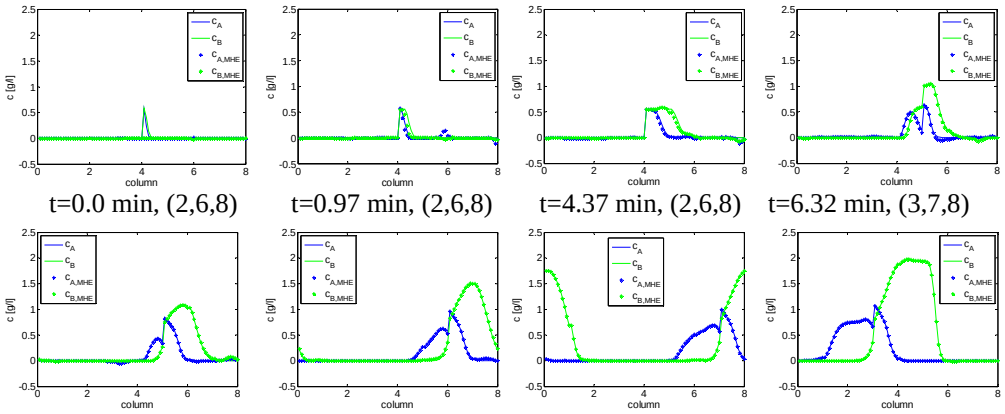


Fig. 3: Parameter estimates (left) and CPU times of the estimator at each sampling point (right)



t=0.0 min, (2,6,8) t=0.97 min, (2,6,8) t=4.37 min, (2,6,8) t=6.32 min, (3,7,8)

t=9.23 min, (3,7,8) t=14.09 min, (4,8) t=18.95 min, (1,5,8) t=38.39 min, (1,5,8)

Fig. 4: Start-up scenario: real and estimated concentrations; measurements at the outlet of the columns indicated in brackets

4. Conclusion

In this paper, a moving horizon state estimation scheme that successfully reconstructs the states and key parameters was described and applied to a high-order nonlinear process with discrete switching of the inputs, the Simulated Moving Bed process. The integration of the presented moving horizon scheme in an advanced control scheme is subject of ongoing research.

5. Acknowledgements

The financial support of the Deutsche Forschungsgemeinschaft (DFG, German Research Council) in the context of the research cluster "Optimization-based control of chemical processes" (RWTH Aachen, IWR Heidelberg, Universität Stuttgart, TU Dortmund) is very gratefully acknowledged.

References

[1] M. Diehl, H. Bock, J. Schlöder, R. Findeisen,Z. Nagy, F. Allgöwer (2002), Real-time optimization and nonlinear model predictive control of processes governed by differential-algebraic equations, *Journal of Process Control*,12(4), 577-585.

[2] A. Küpper and M. Diehl and J. Schlöder and H.G. Bock and S. Engell (2009), Efficient moving horizon state and parameter estimation for SMB processes, *Journal of Process Control*, doi:10.1016/j.jprocont.2008.10.004.

[3] H.G. Bock (1981), Numerical treatment of inverse problems in chemical reaction kinetics, in Modelling of Chemical Reaction Systems,K.H. Ebert and P. Deuflhard and W. Jäger (editors), 18 of *Springer Series in Chemical Reaction Systems*, Springer-Verlag.

[4] T. Gu (1995), Mathematical modelling and scale-up of liquid chromatography, Springer Verlag, New York.

[5] M. Diehl(2002), Real-time optimization of large scale nonlinear processes, 920 of *Fortschritt-Berichte VDI Reihe 8, Mess-, Steuerungs- und Regelungstechnik*, VDI Verlag, Düsseldorf.

[6] A. Jupke (2003), Experimentelle Modellvalidierung und modellbasierte Auslegung von Simulated Moving Bed (SMB) Chromatographieverfahren, Dr.-Ing. Dissertation, Fachbereich Bio- und Chemieingenieurwesen, Universität Dortmund, Shaker-Verlag, Aachen.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Channel Blockage Detection of Microreactors Using Pressure Sensors

Yasuyuki Kaburagi, Masaru Noda and Hirokazu Nishitani

Graduate School of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma 630-0192, Japan

Abstract

Blockage in microreactors is a serious problem that limits their practical usage. A proper diagnosis of blockage is indispensable to ensure these machines operate effectively and stably. We describe two systems, A and B, that can locate channel blockage using pressure sensors. System A uses a computational fluid dynamic (CFD) simulation, and system B uses a simple pressure balance model to construct databases for channel blockage diagnosis. The results of case studies showed that system A can identify the blockage location accurately using fewer pressure sensors than microchannels. System B can also identify the blockage location in many cases, but several misdiagnoses occur. In addition, we describe a method that estimates the degree of blockage in a microchannel. Simulation results showed that our method can estimate the degree of blockage precisely using the pressure sensor that is closest to the blocked location identified by systems A or B.

Keywords: Microreactor, Channel blockage diagnosis, Support vector machine, Computational fluid dynamics, Pressure balance model

1. Introduction

Microreactors have received both industrial and academic attention as a new type of production technology for specialty chemicals that are difficult to produce with conventional reactors. Blockage in the microchannels of microreactors is a serious problem that limits their practical usage. It changes the flow pattern in the microreactor and degrades the product quality. Detecting and identifying the blockage location is indispensable for ensuring more effective and stable operation in these microreactors.

Kano *et al.* (2007) proposed a blockage diagnosis system for a stacked microreactor that uses temperature sensors. The system uses the temperature differences at sensor locations between normal and abnormal operating conditions to identify the blockage location. When the channel blockage does not affect the temperature in the microchannels due to a high surface/volume ratio, which is one of the characteristics of microreactors, the system may not work. Also, they did not address how to estimate the degree of blockage after identifying the location.

This paper describes two blockage diagnosis systems that use pressure sensors instead of temperature sensors. One uses a computational fluid dynamic (CFD) simulation, and the other uses a simple pressure balance model for constructing a database for a channel blockage diagnosis. We also describe a method for estimating the degree of blockage after the location is identified. A CFD simulation was also used in this study to generate the validation data for our systems in case studies.

2. Microreactor

A microreactor generally consists of three parts, namely, an inlet manifold for flow distribution, parallelized microchannels for the reaction, and an outlet manifold for mixing, as shown in Fig. 1. The industrial scale production using a microreactor requires a high number of parallelized microchannels because each one provides only a small amount of product. The blockage in the microchannel causes poor uniformity in the residence time distribution between them, which may make product quality worse.

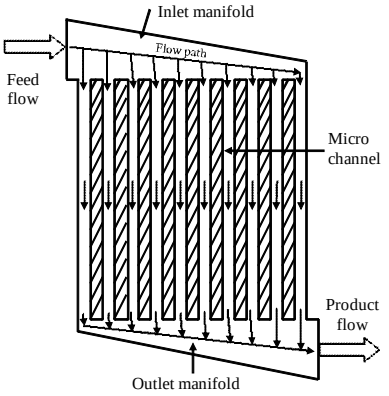


Fig. 1 Structure of microreactor

This study introduces the following assumptions for developing a blockage diagnosis system of a microreactor:

(1) Pressure sensors are set up at the inlets of microchannels. The total number of sensors, N_C, and their locations, $s \in S$, are given in advance.

(2) Channel blockage occurs in only one microchannel at a time.

3. Blockage diagnosis system A based on CFD simulation data

When channel blockage occurs in a microreactor, the pressure distribution changes. If the pressure distribution has certain patterns due to the blockage location, comparing measured pressure-distribution data with the prepared pressure distribution data when channel blockage occured can identify the blockage location.

In this study, we used a CFD simulation to execute a number of simulations for preparing pressure distribution data at sensor locations under various degrees of blockage in each microchannel. The pressure distribution data calculated by the CFD simulation are used as a training data set for a support vector machine (SVM) to construct a discriminate function for identifying the blockage location.

The detailed procedure for building System A is as follows. We assume that pressure distribution data vector $\tilde{P}_N$ consisting of measured pressure $\tilde{p}_{N,s}(s \in S)$ in a normal operating condition is known. Here, $\tilde{p}_{N,s}$ denotes the pressure data measured at sensor location s without any blockage, and S is the set of sensor locations.

(1) Calculate pressure data vector $\hat{P}_{B,i,x}$ consisting of $\hat{p}_{B,i,x,s}(s \in S)$ for some degrees of blockage $x \in X$ in microchannel i using the CFD simulation. Here, $\hat{p}_{B,i,x,s}$ denotes the calculated pressure data at sensor location s under x% blockage in microchannel i (i = 1,…,N_C). X is the set of degrees of blockage for which pressure data is calculated using the CFD simulation.

(2) Calculate the pressure difference vector $\Delta\hat{P}_{i,x}$ by Eq.(1).

$$\Delta\hat{P}_{i,x} = \hat{P}_{B,i,x} - \tilde{P}_N \quad (i = 1,..., N_C, x \in X) \tag{1}$$

(3) Construct a discriminate function using the SVM. The database of blocked channel i and pressure difference data $\Delta\hat{P}_{i,x}$ $(i = 1,..., N_C, x \in X)$ are used as a training data set for the SVM.

The blockage location is identified through the following procedure:

(1) Obtain actual pressure data vector $\tilde{P}_B$ consisting of $\tilde{p}_{B,s}(s \in S)$.

(2) Calculate pressure difference data vector $\Delta\tilde{P}$ using Eq.(2)

$$\Delta\tilde{P} = \tilde{P}_B - \tilde{P}_N \tag{2}$$

(3) Identify the blockage location using the discriminate function of the SVM.

4. Blockage diagnosis system B based on pressure balance model

A CFD simulation is a powerful tool for estimating the states in a microreactor when channel blockage occurs, but it is a time-consuming and costly process. Commenge *et al*. (2002) proposed a pressure balance model (PBM). Figure 2 shows an example of compartments defined in the PBM. The PBM consists of a simplified description of pressure and material balances in compartments. They applied this model to estimate states such as the flow rate and pressure drop in each microchannel instead of using a CFD simulation.

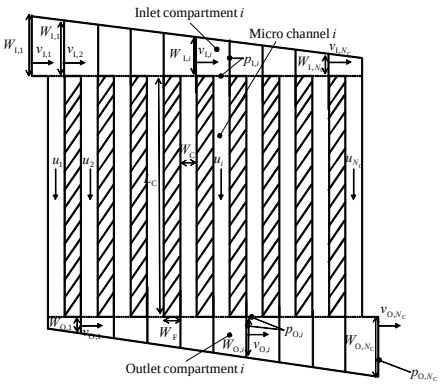


Fig. 2 Pressure balance model

This section describes a simple blockage diagnosis system, B, that is based on the pressure balance model. System B uses correlations between the pressure data vector calculated using the PBM under 50% blockage in each microchannel and the measured pressure data vector at the sensor locations. The procedure for building the database for system B is as follows:

(1) Calculate pressure data vector $\overline{P}_{B,i}$ consisting of $\overline{p}_{B,i,50,s}(s \in S)$ under 50% blockage in microchannel i (i = 1,…,N_C) using the PBM.

(2) Calculate pressure difference data vector $\Delta\overline{P}_i$ using Eq.(3).

$$\Delta\overline{P}_i = \overline{P}_{B,i} - \tilde{P}_N \quad (i = 1,..., N_C) \tag{3}$$

The blockage location is identified through the following procedure:

(1) Obtain actual data vector $\tilde{P}_B$ consisting of $\tilde{p}_{B,s}(s \in S)$.

(2) Calculate pressure difference data vector $\Delta\tilde{P}$ using Eq.(2)

(3) Identify microchannel i^*, which has the highest correlation between $\Delta\tilde{P}$ and $\Delta\overline{P}_i$ as the blocked microchannel. The correlation R_i is defined using Eq.(4).

$$R_i = \frac{(\Delta\tilde{P} - \Delta\bar{\tilde{P}})^T (\Delta\bar{P}_i - \Delta\bar{\bar{P}}_i)}{\left\|\Delta\tilde{P} - \Delta\bar{\tilde{P}}\right\| \left\|\Delta\bar{P}_i - \Delta\bar{\bar{P}}\right\|} \tag{4}$$

Here, all elements of $\Delta\bar{\tilde{P}}$ and $\Delta\bar{\bar{P}}_i$ are the means of all elements of $\Delta\tilde{P}$ and $\Delta\bar{P}_i$, respectively. The superscript *T* denotes the transpose, and $\|\cdot\|$ denotes 2-norm.

5. Estimation method of degree of blockage

After the location of the blockage is identified, the degree is estimated through the following procedure.

(1) Obtain pressure data $\tilde{p}_{B,i^{**}}$ at sensor location i^{**}, which is the closest location to microchannel i^* identified by systems A or B as the blocked location.

(2) Determine $\bar{W}_C^*$, which minimizes the objective function *J* given using Eq.(5).

$$J = \left| \tilde{p}_{B,i^{**}} - \bar{p}_{B,i^{**}}(W_C^*) \right| \tag{5}$$

Here, W_C^* is the width of microchannel i^*. $\bar{p}_{B,i^{**}}$ is the pressure data at sensor location i^{**} calculated using the PBM.

(4) Determine the degree of blockage (*BD*) of microchannel i^* using Eq.(6).

$$BD = \left(1 - \frac{Z_C \cdot \bar{W}_C^* \cdot \bar{u}_{B,i^*}}{Z_C \cdot W_C \cdot \bar{u}_{N,i^*}}\right) \times 100 \ [\%] \tag{6}$$

Here, Z_C and W_C are the original depth and width of each microchannel. $\bar{u}_{B,i^*}$ and $\bar{u}_{N,i^*}$ denote the flow rate in microchannel i^* with and without blockage calculated using the PBM, respectively.

6. Case study

Systems A and B were applied to the blockage diagnosis problems of a microreactor with 10 microchannels. The geometric parameters in Fig. 2 and the operating conditions of the microreactor are summarized in Table 1. The following four types of pressure sensor configurations were implemented as different case studies. In this study, Fluent version 6.1™ (ANSYS Japan) was used to generate validation data for our systems.

Configuration 1: N_S=10, S = {1, 2, 3, 4, 5, 6, 7, 8, 9, 10}
Configuration 2: N_S=5, S = {1, 3, 5, 7, 9}
Configuration 3: N_S=4, S = {1, 4, 7, 10}
Configuration 4: N_S=3, S = {1, 5, 9}

Eighty cases of CFD simulations were executed to obtain pressure distribution data under eight degrees of blockage: 4.8%, 9.1%, 18.5%, 36.4%, 44.8%, 52.9%, 70.8%, and 79.1% in each microchannel. The results of the CFD simulation were used as a training data set to construct a discriminate function of the SVM.

The diagnosis results of system A are summarized in Table 2. The identified number of the blocked microchannel is shown in Table 2 for ten blockage locations, for three

degrees of blockage (8.6%, 27.3%, and 61.4%), and for four sensor configurations. System A perfectly identified the blockage location using only four pressure sensors (Configuration C). Four misdiagnoses, including failure to diagnose, occurred when the degree of blockage was 8.6% and when sensor configuration D was used. These simulation results proved the usefulness of system A.

System B was also applied to the aforementioned blockage diagnosis problem. Table 2 shows that system B identified the blockage location in many cases, but several misdiagnoses occurred. A misdiagnosis tends to occur when the degree of blockage is low and when the pressure sensor interval is high. This is due to a mismatch of the PBM and the model used in the CFD simulation.

Table 3 shows the estimation results of the degree of blockage. The results show that our method can estimate the degree of blockage with less than 2.1% error.

Table 1 Geometric parameters and operation conditions of microreactor

Name	Parameter	Value	Unit	Name	Parameter	Value	Unit
Number of channel	N_C	10	-	Inlet manifold	$W_{I,1}$	5.0	mm
Channel width	W_C	100	μm		$W_{I,Nc}$	1.4	mm
Channel depth	Z_C	100	μm	Outlet manifold	$W_{O,1}$	1.4	mm
Channel length	L_C	20	mm		$W_{O,Nc}$	5.0	mm
Width of fin	W_F	284	μm	Inlet velocity	$v_{I,1}$	0.01	m/s
Viscosity	M	0.1	Pa·s	Outlet pressure	$p_{O,Nc}$	101.3	kPa

7. Conclusion

Two blockage diagnosis systems, A and B, that use pressure sensors were described. Simulation results showed that system A can diagnose the blockage location successfully using only four pressure sensors against ten microchannels, even when the degree of blockage is low. Though system B can also identify the location of blockage in many cases, several misdiagnoses occurred. The diagnosis performance of system B is worse than that of system A, but it will be useful when CFD simulation is not available.

When some experimental data under abnormal operating conditions are obtained, those data can be used as a part of a training data set for constructing a discriminate function of system A. An optimal integration of experimental data and CFD simulation data is a next problem to be discussed.

Our blockage diagnosis system and the estimation method for the degree of blockage are very useful for detecting early blockage and for identifying the blocked microchannels. In this study, we focused on a microreactor with many microchannels, but our methods can very easily be applied to various types of micro-devices when pressure distribution data can be calculated under abnormal operating conditions.

References

Commenge, J. M., L. Falk, J. P. Corriou and M. Matlosz, 2002, Optimal Design for Flow Uniformity in Microchannel Reactors, *AIChE J.*, 48, 2, 345-358

Kano, M., T. Fujioka, O. Tonomura, S. Hasebe and M. Noda, 2007, Data-Based and Model-Based Blockage Diagnosis for Stacked Microchemical Processes, *Chem. Eng. Sci.*, 62, 1073-1080

Table 2 Diagnosis results of systems A and B

Blockage location	Blockage degree [%]	System A				System B			
		A	B	C	D	A	B	C	D
	61.4	1	1	1	1	1	1	1	1
1	27.3	1	1	1	1	1	1	1	1
	8.6	1	1	1	1	1	1	1	1
	61.4	2	2	2	2	2	2	1	1
2	27.3	2	2	2	2	2	2	1	1
	8.6	2	2	2	1	2	2	1	1
	61.4	3	3	3	3	3	3	4	3
3	27.3	3	3	3	3	3	3	4	3
	8.6	3	3	3	2	3	3	3	4
	61.4	4	4	4	4	4	4	5	5
4	27.3	4	4	4	4	4	4	5	5
	8.6	4	4	4	4	4	4	5	5
	61.4	5	5	5	5	5	5	5	6
5	27.3	5	5	5	5	5	5	5	6
	8.6	5	5	5	5	5	5	5	6
	61.4	6	6	6	6	6	6	6	6
6	27.3	6	6	6	6	6	6	6	6
	8.6	6	6	6	6	6	6	7	6
	61.4	7	7	7	7	7	7	7	7
7	27.3	7	7	7	7	7	7	7	7
	8.6	7	7	7	7	7	7	7	7
	61.4	8	8	8	8	8	8	8	8
8	27.3	8	8	8	8	8	8	8	8
	8.6	8	8	8	-*	8	8	7	10
	61.4	9	9	9	9	9	9	9	9
9	27.3	9	9	9	9	9	9	9	9
	8.6	9	9	9	9	9	9	9	9
	61.4	10	10	10	10	10	9	10	9
10	27.3	10	10	10	10	10	9	10	9
	8.6	10	10	10	8	10	10	10	8

* Fail to identify

Table 3 Estimation results of degree of blockage (*BD*) of microchannel

Blockage location	Actual *BD* [%]	Estimated *BD* [%]	Blockage location	Actual *BD* [%]	Estimated *BD* [%]
	61.4	61.4		61.4	59.7
1	27.3	27.1	6	27.3	26.6
	8.6	8.4		8.6	8.3
	61.4	59.7		61.4	59.5
2	27.3	26.5	7	27.3	26.4
	8.6	8.3		8.6	8.3
	61.4	59.5		61.4	60.0
3	27.3	26.4	8	27.3	26.7
	8.6	8.3		8.6	8.4
	61.4	59.3		61.4	61.4
4	27.3	26.4	9	27.3	27.3
	8.6	8.3		8.6	8.4
	61.4	60.9		61.4	60.6
5	27.3	26.9	10	27.3	26.9
	8.6	8.6		8.6	8.4

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

PETROBRAS Experience Implementing Real Time Optimization

Fábio S. Liporace,[a] Marcos V. C. Gomes,[a] Antônio C. Katata,[b] Antônio C. Zanin,[c] Lincoln F. L. Moro,[c] Carlos R. Porfírio,[d]

[a]*PETROBRAS/CENPES/PDEDS/Gás Natural/Célula de Otimização, Av. Horácio Macedo, 950 - Cidade Universitária, Rio de Janeiro, 21949-915, Brazil.*
[b]*PETROBRAS / REVAP / Otimização, Refinaria Henrique Lage, Rodovia Presidente Dutra km 143, São José dos Campos, 12223-680, Brazil*
[c]*PETROBRAS / Abastecimento - Refino / Tecnologia de Refino / Otimização, Av. República do Chile, 65 - Sala 2102 - Centro, Rio de Janeiro, 20031-912, Brazil*
[d]*PETROBRAS / RPBC / Otimização, Refinaria Presidente Bernardes Cubatão, Praça Marechal Stênio Caio A, Lima, 1, Raiz da Serra, Cubatão, 11555-900, Brazil*

Abstract

PETROBRAS has defined Real Time Optimization (RTO) as a "High Sustainability" technology for downstream operations, due to its high economic return. Since 2001, RTO tools are being tested within the Company, either using in-house process simulators or, sometimes, using available commercial ones. This paper presents an overview of the PETROBRAS experiences on RTO, showing applications on Distillation and Fluidized Catalytic Cracking (FCC) units. Alternatives based on Sequential-modular simulators, along with reduced models (Kriging models and neural nets), as well as Equation-oriented based simulators / optimizers have been explored. The project scopes vary from covering only the Reactor / Regenerator section of a FCC unit up to a whole Crude distillation unit, including the preheat train, all distillation towers and the heat and material integration. Some of these RTO applications have been running close loop for almost 6 months, with proved expressive economical benefits. Based on the knowledge acquired during all these years, some of the future development needs for the improvement of RTO technology will be presented and discussed, as a guide for future research projects.

Keywords: real time optimization, industrial application, sequential modular simulator, equation oriented simulator.

1. Introduction

RTO technology is a powerful tool for the continuous search of the most profitable way to run petroleum and petrochemical process units. Some of the major Oil Companies, as ExxonMobil, Shell and BP to name a few, have specific groups to study it and use it to improve their business [1] [2]. Cutler and Perry [3] state that despite being a hard and complex task, its potential benefits are relevant and might provide profit increases around 6 to 10% when allied to Advanced Process Control (APC).

The task of an RTO application is to make the best of an existing process unit, adjusting its process variables for every new change of external conditions, like operational variables, feed compositions and process constraints. The RTO benefits are usually

associated with the maximization of products and minimization of the specific energy consumption and other resources, depending on the following factors [4] :

- Market availability;
- Safety and environmental constraints
- Products prices and feed costs
- Product specifications

This paper focuses on the use of RTO technologies in the optimization of petroleum downstream process units. The next section provides an overview of the structure and mechanism of RTO systems. Section 3 describes PETROBRAS experience on the development and use of in-house and commercial RTO solutions. On section 4 this paper is concluded with the discussion of the key aspects to the achievement of successful implementations.

2. RTO structure, mechanism and characteristics

Though many implementation structures are possible, RTO has been usually used as a connection between the tasks of production planning & scheduling and process control [4] [7] [8] for petroleum and petrochemical applications. The central figure of an optimization application is the mathematical model. It is expected to represent the process behavior on a wide range of operating conditions with good accuracy. It is required not only to guarantee that the predicted potential profitability matches that of the real process, but also because when the optimal solution is implemented the process constraints must not be violated. Most RTO systems used nowadays are based on rigorous, steady-state, first principles mathematical models.

The good performance of an RTO system depends on a reliable mathematical model and on reliable input data. In order to obtain that, many procedures must be executed before the economic optimization problem can be solved [4] [6] [8]:

a) Gross Error Detection
b) Steady-state Detection
c) Data Reconciliation
d) Parameter estimation.

Once that a reconciled data set and a fitted model have been obtained, the process optimization can be performed. The optimization problem usually consists of maximizing the operational profit (or minimizing costs) subject to a set of constraints. On most situations the optimization problem is posed as a non-linear programming (NLP). Most commercial applications are based on variations of the SQP (Successive Quadratic Programming) algorithm. This algorithm is also used to solve the previous Data Reconciliation and Parameter Estimation problems.

3. Real Time Optimization within PETROBRAS

Since 2004, RTO has been classified by PETROBRAS and its Strategic Downstream Committee as a “High Sustainable” technology. It means that RTO is seen as a key technology to improve PETROBRAS performance and profit, and therefore significant effort and resources will be spent on this subject. PETROBRAS implementations on RTO covered a wide range of alternatives, focusing both on profitability and on the search of the best way to deliver the technology:

- Fluid Catalytic Cracking (FCC) and Crude Distillation Units (CDU)
- proprietary and commercial process models and RTO systems
- Sequential Modular (SM) and Equation Oriented (EO) approaches [6]

3.1. First experiences

The first RTO initiatives were taken using PETROBRAS' in-house process simulator for FCC, with a small scope covering only the reactor/regenerator section. The

proprietary process model used is based on a Sequential Modular (SM) approach, many difficulties were found (see next section). Nevertheless, we were able to test the technology as well as our engineers to take a step further.

3.2. Distillation Unit / SM approach (2004)

This implementation took place at the Crude Distillation Unit (CDU) and the two Solvents Units of RECAP refinery [6].

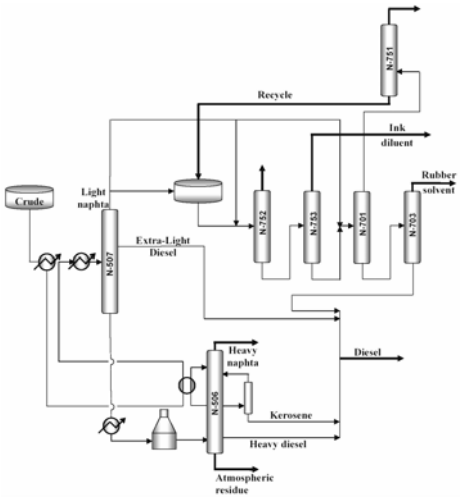


Figure 1 - Scheme of the CDU and the Solvents Units of RECAP/PETROBRAS.

The process model was built using PETROX, a proprietary sequential-modular process simulator from PETROBRAS. The simulation comprises 53 components and pseudo-components and 64 unit operation modules, including the 7 distillation columns and a recycle stream. All modules are built with rigorous, first-principles models.

For optimization applications, PETROX was linked to NPSOL, an SQP optimization algorithm. Procedures for Steady-state and Gross error detection, Data Reconciliation, Parameter Estimation and Economic Optimization were implemented. The economic optimization problem consisted of the maximization of the operational profit, constrained by limits related to product specifications, safety constraints, feed rate and performance parameters. The whole optimization problem involves 19 decision variables and 21 constraints.

Most of the reported problems of optimization based on sequential-modular models were observed in this application:

- Low computational efficiency, due to slow recycle loops and the numerical derivatives, that imply running the SM model several times. These derivatives are also inaccurate, which slows down the optimization process even more.
- Lack of reliability. The SM model is computed many times, and must converge always. If a single failure happens during the optimization, all the effort is lost.

In order to minimize these problems, a lot of effort must be spent on the conception, customization and tuning of the SM model. However, that is no guarantee of success. When the Data Reconciliation and Parameter Estimation problems were implemented, the same problems were observed.

3.2.1. Metamodel approach

In order to overcome some of these shortcomings, a metamodel approach has been studied. Metamodels or surrogate models [5] [6] are reduced models whose parameters are obtained with data that is generated with rigorous, first principles models. In this work, an optimization procedure was developed, combining metamodels and rigorous models with a Sequential approximate optimization (SAO) algorithm. The optimization problem is solved based on the metamodel, that is updated with data obtained from the rigorous model throughout the optimization procedure. The RECAP optimization problem was addressed with this approach, with kriging models and neural nets used as metamodels. Accurate results have been obtained with considerable reduction of the computational effort on most of the studied cases.

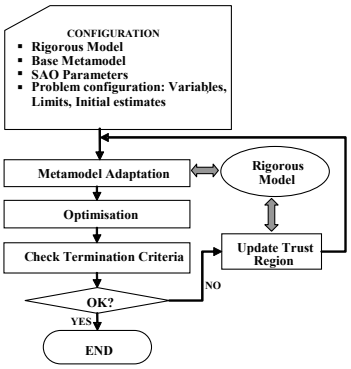


Figure 2 - SAO strategy applied to the metamodel-based optimisation.

3.3. Distillation Unit / EO (2005 to 2006)

This was the first EO RTO project PETROBRAS implemented. After an International Bid, where 3 well-known companies were invited to submit their proposals, we've decided to contract AspenTech and its AspenPlus Optimizer. The project scope included all 3 preheat trains as well as Pre-flash, Naphtha Stabilizer, Atmospheric, Vacuum and Pre-vacuum distillations towers. The unit was fully energy and mass integrated modeled on the RTO software, which allowed us to study the implications that changes on the preheat train, like feed distribution, has on the Atmospheric tower, for instance. Or to study the best pumparound heat removal distribution along this tower and its effects on the preheat train. In order to do that, all pumparounds were modeled as external streams from the tower and not as an internal model within its model (see Figure 3), as it is common on SM simulators.

The system is running on open loop since 2007. PETROBRAS intends to close loop this year after making model tuning adjustments in order to incorporate the new atmospheric trays and other unit improvements. Nevertheless, by keeping the system running open loop (around 9 runs / day), we were able to improve our knowledge of the system itself, how to overcome non convergence problems and gain expertise on how to maintain such a real time, strong data and instrumentation dependent system as well as evaluate potential benefits (around 13 000,00 dollars / day).

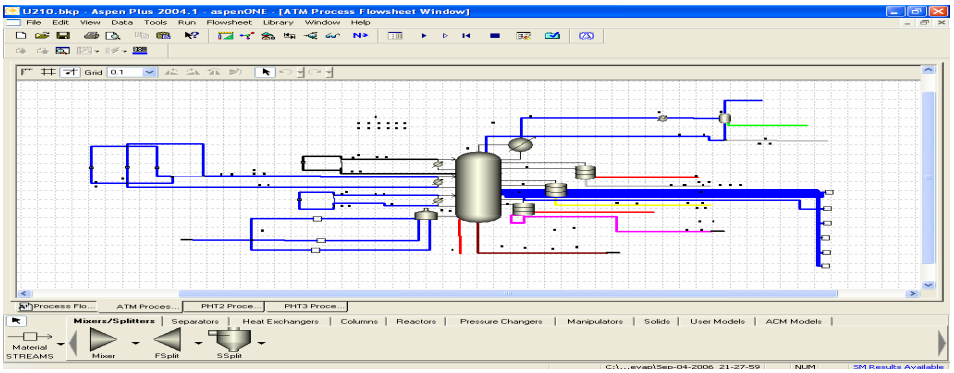


Figure 3 - Aspen Plus Optimizer Screenshot - Atmospheric tower

3.4. FCC Unit / EO (2007 to 2008)

Following the success on the distillation unit, PETROBRAS moved forward to implement an RTO on another very important unit. Again, after an international bid where the same 3 well known companies were invited to be on, we've decided to use Invensys and its ROMeo, as part of our intend to test all the available technology. The project scope included the Reactor / Regenerator section, Main Fractionator and Gas Recovery Plant. Again the unit was fully energy and mass integrated modeled.

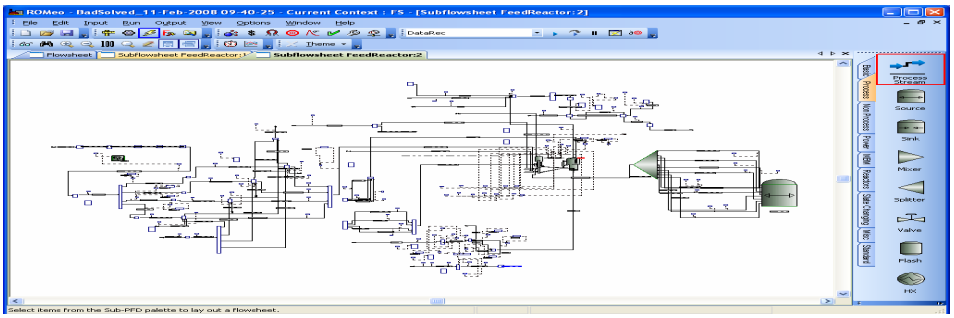


Figure 4 - ROMeo screenshot - Reactor/Regenerator Section.

The system is running on closed loop (around 8 runs / day) since June/08 with most of the independent variables active. On average, around 60% of the successful runs are being accepted by Operations and targets are being sent to Advanced Control. PETROBRAS has evaluated an average gain of US$ 0.12 / bbl of FCC feed for this application, by comparing the unit performance with and without RTO.

A few comments on both projects:

- **Lack of instrumentation on preheat train (FCC)** – implied on simplifications, which has impacts on Main Fractionator heat balance and, thus, must be evaluated from time to time
- **Low feed lab analysis frequency** – There is a need for a better way to estimate feed characterization
- **Non-convergence problems** - Mainly, due to instrumentation faulty and/or out of service heat exchanger or other piece of equipment. Although there is a kind

of standard procedure to deal with them, it is not possible to automate it. So each problem must be solved on a case to case, hands-on basis.

These facts enforce the need for a fully dedicated RTO engineer for each application, not only to assess its results and make sure they are being implemented, but to keep the system running despite of the many daily issues the application faces.

4. Conclusions

4.1. Successful implementations

A key aspect to a successful RTO application, is the good integration among different corporate teams, like Process analysis, Production and Planning Program, Optimization and Automation, IT, Instrumentation, Laboratory, Production. A long-term plan must also be done to assure that the acquired benefits are kept and expanded, with APC and RTO experts dedicated to the on-line applications after their implementations. The process model must be updated throughout the process unit life cycle. The problem constraints and the optimization results must be continuously evaluated. The advanced and regulatory control systems are the ways by which the RTO optimal solutions are implemented. Therefore, they are essential to a successful RTO as well.

4.2. Modeling approach

PETROBRAS experiences showed that the Open-equation approach is more suitable for RTO, when compared to the Sequential-modular process models, specially when process unities of higher complexity are addressed.

4.3. Future Developments

PETROBRAS automation staff is now working to apply RTO technologies to other process technologies. Another ongoing initiative is the simultaneous application of RTO to more than one process unit.

References

[1] Terry, J.D.; Righi, J. M.; Moore, J.W.; Shah, M.K. Baytown Refinery Pipestill 3 RTO. IPS North American Client Conference, September, 2008.

[2] Greuel, J. Refinery-Wide ARPM at Sarnia Refinery. IPS North American Client Conference, September, 2008.

[3] Cutler, C. R.; Perry, R. T. Real-time optimization with multivariable control is required to maximizes profits. Computers and Chemical Engineering, 7, n.5, p.663-667, 1983.

[4] White, D. C. Online optimization: what, where and estimating ROI. Hydrocarbon Processing, p. 43-51, June, 1997.

[5] Gomes, M.V.C.; Bogle, I.D.L.; Biscaia Jr., E.C.; Odloak, D. Using kriging models for real-time process optimisation. Proceedings of the 18th European Symposium on Computer Aided Process Engineering, p. 361-366, 2008

[6] Gomes, M.V.C. Otimização Seqüencial por Aproximações – Uma aplicação em tempo real para o refino do petróleo, D.Sc. Thesis, (*in portuguese*), PEQ/COPPE/UFRJ, Rio de Janeiro, Brazil, 2007.

[7] Shobrys, D.E., White, D.C. Planning, scheduling and control systems: why cannot they work together. Computers and Chemical Engineering, v. 26, p. 149-160, 2002.

[8] Romagnoli, J.A., Sanchez, M.C., 2000, Data Processing and Reconciliation for Chemical Process Operations., Academic Press.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Embedded Control for Optimizing Flexible Dynamic Process Performance

Jeonghwa Moon, Seonbyeong Kim, Gerardo Ruiz and Andreas A. Linninger

Laboratory for Product and Process Design, Department of Chemical Engineering, University of Illinois at Chicago. Chicago, IL 60607, USA.

Abstract

High performance processes require design that operates close to design boundaries and specifications while still guaranteeing robust performance without design constraint violations. In order to safely approach tighter boundaries of process performance, much attention has been devoted to integrating design and control in which dynamic controllability as well the design decisions are taken simultaneously. However rigorous methods solving design and control simultaneously lead to challenging mathematical formulations which easily become intractable numerically and computationally. This paper introduces a new mathematical formulation to reduce this combinatorial complexity of integrating design and control. We will show substantial reduction in the problem size can be achieved by embedded control decisions within specific designs. This embedded control decisions avoid combinatorial explosion of control configuration by using a full state space model that does not require pairing of control variables and loops. The current capabilities of the methodology will be demonstrated using a realistic reactor-column flowsheet.

Keywords: Design and control integration, process and operational uncertainty, design and analysis of dynamic flexibility.

1. Introduction

The aim of integration of process design and control is finding best design decision considering economics and dynamic controllability simultaneously. During the last 30 years, several methodologies have been developed for the integration of design and control (Vassilis et al., 2004; Ramirez & Gani, 2007; Seferlis P., & Georgiadis M.C., 2004). In spite of their progress in integration of design and control field, the available methodologies still offer insufficient insight to the problems. Recently, we proposed a new method, *embedded control optimization approach* (Malcolm et al, 2007). This integrated design method can operate satisfactorily under adverse input conditions, while delivering products within desired quality specifications. In this paper, this embedded control optimization methodology is reviewed and enhanced to incorporate full state space estimator for better the dynamic process performance. A case study of designing reactor-column flowsheet is considered as a more realistic example for this methodology. In this case study, we will illustrate the ability of performing design and control integration with rigorous mathematical techniques.

2. Methodology

2.1. Problem decomposition

The conceptual problem of the integration of process design and control is a stochastic infinite dimensional mixed integer dynamic optimization problem which is extremely challenging for existing mathematical programming techniques. To overcome intractability of original problem, Pistikopoulos and co-workers proposed problem decomposition algorithms (Mohideen et al., 1996). In their problem decomposition, the main design optimization problem is solved in discretized sampling space; rigorous flexibility tests in a second stage ensure the dynamic feasibility over the entire uncertain space. If current candidate designs are not feasible in critical scenarios, these critical scenarios incorporate into the sampling space and main problem is solved again with additional augmented scenario. The final optimal designs can be achieved by repeating this iterative process. This framework is also used in our algorithm.

2.2. Embedded control optimization

Even though, the problem decomposition substantially reduces the problem size, it still remains a challenge due to combinatorial complexity of the non polynomial hard search space. Specifically, introduction of control decisions such as the insertion of feedback loops, or pairing of manipulated and control variables causes combinatorial explosion in the possible integrated design and control realizations. We therefore propose to separate the design decisions from the control decisions as shown Figure 1. At the master level, we fix design decisions such as reactor sizes, and residence time that govern dynamic process performance. No control decisions are made at this level. Once the main design decisions are specified, we assess the dynamic performance of process by using a simplified, yet reasonably competitive control schemes based on full state space identification and least square regulation. Optimal control action is calculated with relative ease based on linear state space models which are obtained dynamically in each time step. In order to keep the model consistent with the nonlinear process dynamics, the process identification is repeated in each time step. The detailed description is found elsewhere (Malcolm et al., 2007).

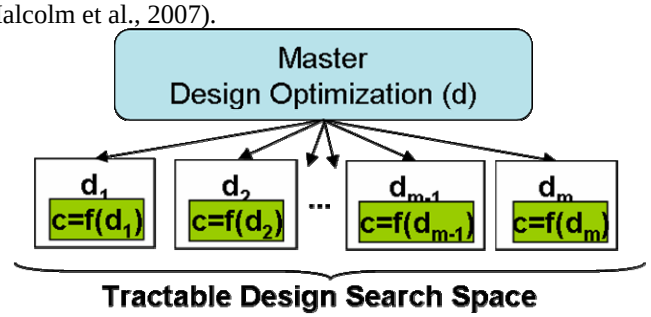


Figure 1 Proposed embedded control optimization structure. It optimizes control choice with given design.

3. Flexible design of reactor column

This section demonstrates the effectiveness of embedded control optimization by designing a reactor-column flowsheet process with uncertain reaction coefficients. This is similar to processes introduced and studied by Luyben (2007). The aim of this case study is to determine optimal design specifications with reasonable control for

dynamically flexible operations. This task of design and control integration should be done simultaneously with reasonable computational effort.

3.1.1. Process description.

This flowsheet has one reaction and one distillative separation column, as shown in ***Figure 2***. The precursor *A* enters the reactor to be converted into the product *B* with an exothermic irreversible first order reaction.

$$A \xrightarrow{k_0} B \qquad \textbf{(1)}$$

The reaction cannot be driven to full conversion, because this operation would require too high temperature causing the product *B* to be destroyed. In addition, this high temperature might violate safety constraints and lead to reactor explosion. To control temperature in reactor, it is equipped with a cooling jacket. Its effluent *F* is directed to the continuous distillation column in order to separate the product *B* from unreacted raw material *A*.

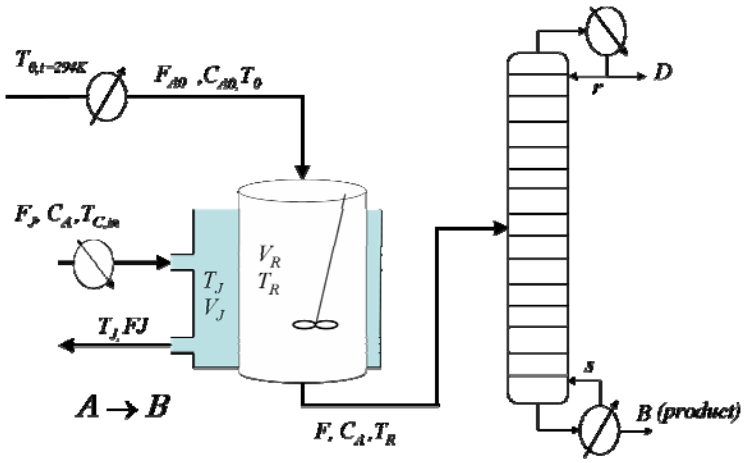


Figure 2 Reactor-column flowsheet. This is composed of a reactor and a column. The reactor is surrounded by a cool jacket.

3.1.2. Design & manipulated variables.

Reactor diameter (D_R) and length (L_R), and heat transfer area (A_J) are treated as design variables. Also the number of stages (*Ns*), feed stage, and column diameter (N_f) are design variables of the column,. For manipulated variables, we select input feed temperature (T_0), coolant temperature (T_J), reflux ratio (*r*), and reboil ratio (*s*). The values of the design variables and manipulated variables need to be determined to insure safe operation within desired quality standards, under any operating conditions and in the presence of uncertain reaction conditions.

3.1.3. Operational constraints.

For safe operation, this process needs to satisfy three constraints at all times.

For the safety, reactor temperature (T_R) should never exceed 385K,

$$T_R < 385 \qquad \textbf{(2)}$$

For the product quality, we enforce the mole fraction of component *A* (x_A) of final product B should be less than 0.05.

$$x_A < 0.05 \quad \textbf{(3)}$$

For minimum process productivity, the conversion ratio (χ) should be greater than 0.7.

$$\chi > 0.7, \quad \chi = 1 - \frac{C_A}{C_{A0}} \quad \textbf{(4)}$$

Thus, control variables are T_R, x_A, and χ. Among these constraints, the safety constraint needs to be enforced dynamically for all time periods of possible scenarios, as opposed to merely the steady state which is the sole concern in classical flexibility analysis. A certain design specification does not satisfy safety constraint in dynamic state, even though it does not have any violations in steady state. Therefore, we need to consider the dynamic flexibility for the optimal dynamic process performance. Productivity and quality constraints are soft constraints, which lead to performance losses that cost money, but need not to be enforced rigorously for all time intervals.

3.1.4. Uncertainty scenarios.

We wish to investigate the impact of two main uncertain parameters, associated with chemical reactions. The first parameter is preexponental factor k_0, and the second is heat of reaction λ. Their nominal values and variance are illustrated in **Table 1**.

Table 1 Nominal values and expected deviation of uncertain parameters

	θ^N	$\Delta\theta^+$	$\Delta\theta^-$	var
k_0	20.75e6	2.07e6	2.07e6	10%
λ	-17.43e3	-1.74e3	-1.74e3	10%

As a first attempt to perform the stochastic optimization, we chose 10 samples in the uncertain space of reaction conditions using Latin hypercube method, and evaluated the probabilities of each parameter set to calculate expected cost.

We now wish to rigorously determine the design and manipulated variables in the manner that, with every realization of the reaction conditions and all variability in the dynamic performance by uncertainty, the process does not violate the constraints and produces product *B* in desired purity limit.

3.2. Identification test

In order to employ the embedded control optimization, our methodology dynamically performs of adaptive state space identification. This repeated identification is to convert highly nonlinear process model into a simple linear state representation in open loop. With this simplified linear model, the linear regulator can find optimal control moves. The performance of process in response to step change in feed rate as well as the impact of sinusoidal disturbance is show in **Figure 3**. We have now identification phase which is occurring until time step of 40,000. In each step, there is a little change because the process is operated with approximately constant conditions, only in sinusoidal disturbances slightly impacts processes. However, a big impact occurs at t=40,000, in which the flow rate is suddenly increased to double. Now dynamics of whole process is changed and we need to adjust to new state space. This is done very satisfactorily, as shown **Figure 3**. The blue lines represent nonlinear data from mathematical model; red lines represent predicted data by identifier. Even though model is linear, because it is updated in every time step, it can maintain reasonable predictability even in a transition phase.

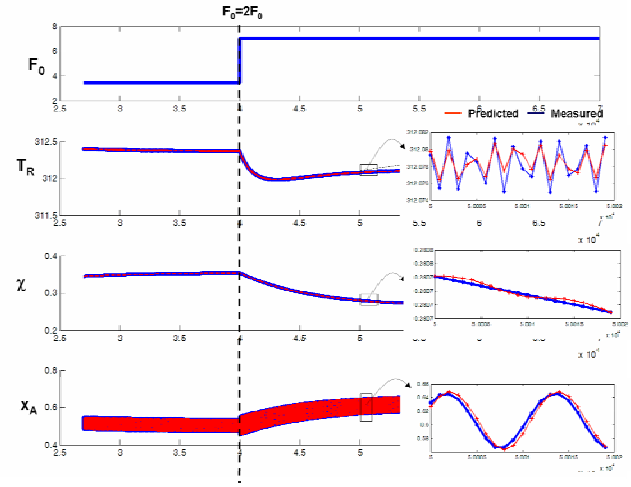


Figure *3* Performance test of identifier. This figure shows sequential least identification method predicts dynamic system behavior well, even in dynamic transition phase.

3.3. Control performance of there different designs

Next, we tested controllability of several designs under specific disturbance scenarios. Dynamic controllability of three reactors with different volumes of 12, 23, 143m^3 were investigated in response to dramatic increase of the feed rate. The doubling feed occurs at t=40,000.

Case 1: Small reactor, V_R=12m^3. When doubling the input feed, this small reactor is loosing the controllability- the control is not capable of handling this upset to process conditions as shown in **Figure 4**. The reactor temperature drops and reaction conversion declines, and product quality is violated.

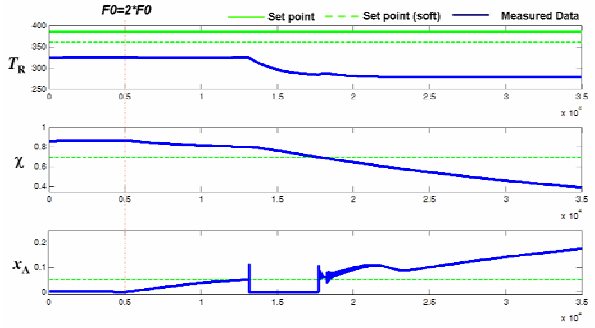


Figure *4* Control performance with VR=12m^3. It fails to keep the values of two control variables (χ, x_A) under the set points.

Case 2: Middle reactor, V_R=23m^3. We tested controllability of a middle size reactor. With this design specification, the controller rejected the disturbance well. The quality constraint was violated briefly in the instance of the disturbance inception, but the controller quickly removed this quality issue and kept values of all process variables under the set points. This design handles disturbance very well with simple control scheme.

Case 3: Large reactor, V_R=143m^3. Finally we tested larger reactor size. Our model demonstrated that the feed flow disturbances do not affect the process dynamics much in this over-dimensioned design. So the large reactor is very robust against the disturbance, but it is not competitive because of cost. It turned out that middle reactor generated more profit than the large reactor. Moreover, large reactors exhibit sluggish response and slow servo performance when set points are changed to adjust to different product specifications. This example supports the notion that arbitrary overdesign of equipment is not a solution to ensuring dynamic process flexibility of high performance processes.

3.4. Optimal integrated design with control

The previous analysis of different design demonstrated the trade off between controllability and servo performance. When considering also costs and profits what would be the best integrated design and control? For maximizing the performance, while at the same time planning flexible operation, we performed the design optimization under uncertainty as follows. For capital configuration, we consider the reactor, the column and the heat exchangers. Also we consider energy consumption for input feed temperature, cooling temperature, the reboiler and condenser, as operating cost. Also, we consider product prices to calculate total annual profit. The master level of this problem is to maximize total profit. To solve optimal design problem, Nelder-Mead simplex method is used in master level of our methodology. We found the best optimal design after 34 iterations.

4. Conclusions & future work

This paper describes conceptual framework for integrating of design and control developed by our group. Our methodology is to recast integrated design and control problem into solvable mathematical programming formulation. The case study described to show suitability of our methodology. In future, more challenging flowsheets will be examined. Also we wish to improve the quality of identification for highly nonlinear processes using more advanced identification such as subspace identification method or nonlinear model predictive control. However these advanced algorithms are more expensive, the trade off between accuracy and performance of algorithms need to be considered.

Acknowledgements

Financial support from NSF Grant CBET-0626162 is gratefully acknowledged.

References

Luyben, W. L. (2007). Chemical reactor design and control. *Wiley-Interscience.*

Malcolm, A., Polan, J., Zhang, L., Ogunnaike, B. A., & Linninger, A. A. (2007). Integrating systems design and control using dynamic flexibility analysis. *AIChE Journal*, 53, 2048-2061.

Mohideen, M. J., Perkins, J. D., & Pistikopoulos, E. N. (1996). Optimal design of dynamic systems under uncertainty. *AIChE Journal*, 42, 2251-2272.

Ramirez, E., & Gani, R. (2007). Methodology for the design and analysis of reaction-separation systems with recycle. 2. Design and control integration. *IECR*, 46, 8084-8100.

Seferlis P, & Georgiadis M. C., The Integration of process design and control. New York: Elsevier; 2004.

Vassilis S., Perkins J. D., & Pistikopoulos, E. N. (2004). Recent advances in optimization-based simultaneous process and control design. *Comp. & Chem. Engr.*, 28(10): 2069-2086.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Comparison between Statistical and Observer-Based Approaches for Fault Detection and Isolation in a Chemical Process

Thiago de Sá Feital[a,b], Uwe Kruger[b], José Carlos Pinto[a], Enrique L. Lima[a]

[a]*Programa de Engenharia Química/COPPE, Universidade Federal do Rio de Janeiro, 68502, Rio de Janeiro, Brazil*
[b]*Department of Electrical Engineering, The Petroleum Institute, P.O. Box 2533, Abu Dhabi, U.A.E.*

Abstract

This paper summarizes the results of a comparison including three different fault detection and isolation (FDI) methods for dynamic systems. The techniques studied have recently been proposed in the literature and are Dynamic Principal Component Analysis (DPCA), Canonical Variate Analysis (CVA) and Subspace Model Identification (SMI). The aim of this study is to contrast the performance of each method in detecting and isolating incipient fault conditions. Utilizing real data from a debutaniser distillation tower, this study yields that the observer approach based on an identified SMI model is most sensitive for fault detection but performs poorly in isolating the fault condition. This method failed to correctly diagnose the fault condition using fault isolation approach. In contrast, DPCA offered a correct picture of this event using variable reconstruction and contribution charts, whilst CVA only yielded satisfactory results using variable reconstruction. For this study, it is therefore concluded that both approaches have complementary strengths and weaknesses.

Keywords: Subspace Methods, Fault Detection and Isolation, Chemical Process

1. Introduction

Based on the ever growing requirement for safe and reliably operating processes in the chemical industry, fault detection and isolation has received considerable attention over the past decades. The research literature has shown two main approaches: (i) data driven multivariate statistical techniques (Chiang *et al.*, 2001; MacGregor *et al.*, 2005) and the use of model-based state-space systems (Isermann, 2005). These approaches yielded similar fault detection and isolation methods. Given the large number of variables that are typically recorded in complex chemical, the use of mechanistic first principal models is difficult in practice, which has led to significant research activities on data-driven methods.

The aim of this article is to investigate the potential of utilizing model-based methods based on the observer design for fault detection and isolation. The required model is provided by subspace model identification, a data driven methodology that has been extensively studied over the past decade (Verhaegen, 1994; Van Overschee and De Moor, 1996). This data-driven and model-based methodology is contrasted with competitive multivariate methods for dynamic systems. Chiang *et al.* (2001) summarized that DPCA and CVA are such methods that identify a dynamic monitoring model in a reduced dimensional latent variable space.

For DPCA and CVA, process monitoring is based on the use of univariate statistics and fault isolation relies on variable reconstruction (Qin and Li, 2001) and contribution charts (Miller *et al.*, 1998). In a similar fashion to the univariate fault detection indices of DPCA and CVA, the paper proposes the use of the T^2 and SPE statistics on the basis of the estimated states and the output estimation error, respectively, for the model-based approach.

To examine the utility of the model- and observer-based approach with DPCA and CVA, the article summarizes the results of an industrial process. The objective of this study is to contrast the performance of each method in detecting fault conditions and in diagnosing the recorded event.

2. Fault Detection and Isolation Methods

This section briefly reviews DPCA, CVA and the proposed model- and observer-based approaches for FDI.

2.1. Dynamic Principal Component Analysis (DPCA)

DPCA (Chiang *et al.*, 2001) relies on a time series arrangement of a set of recorded process variables. An eigendecomposition of the covariance matrix of this augmented variable set forms then the basis for constructing a model plane and a complementary residual subspace. The retained components of this eigendecomposition relate to the most dominant eigenvalues of this decomposition and span the model plane. The residual subspace is consequently spanned by the discarded eigenvectors. Geometrically, the orthogonal projection of the augmented data vector, z, onto the model plane and the residual subspace is $\hat{z} = PP^T z$ and $\tilde{z} = (I - PP^T)z$, respectively, such that the following equality holds $z = \hat{z} + \tilde{z}$.

2.2. Canonical Variate Analysis (CVA)

Similar to DPCA, CVA performs FDI in a model plane and a residual subspace defining state sequences, x, (Chiang *et al.*, 2001). CVA relies on augmented vectors that include time lagged values of the process input and output variables. More precisely, the lagged terms are time-series arrangements of the output variables that relate to future, f, and past measurements and an arrangement of the input variables for past measurements. The state sequences are the first *d* dominant canonical variates, $\hat{x} = J_d p = U_d^T \Sigma_{pp}^{-1/2} p$, where U_d is a matrix storing the first *d* columns of U, $\Sigma_{pp}^{-1/2} \Sigma_{pf} \Sigma_{ff}^{-1/2} = U \Sigma V^T$. Moreover, p is a vector storing the past time-series arrangements of the input and output variables and the different Σ are covariance and cross-covariance matrices. The residual vector of the state-space model in term of the past is given by the *r* remaining state variables, $\tilde{x} = J_r p$.

2.3. Subspace Model Identification (SMI)

The state space model that the observer-based FDI scheme requires is provided by an SMI technique. Based on the numerically stable and efficient QR and SVD decompositions, the MOESP algorithm (Verhaegen, 1994) has gained attention. In addition, Chen (1999) reported that the subspace identification offers various practical advantages, mainly for industrial process systems, over the classical realization from input/output identified models (Chen, 1999). This algorithm relies on future and past arrangements of the input and output variables to form a total of four block Hankel matrices, where the samples are stored row wise. These matrices are then stacked prior

to the application of a QR decomposition. Together with a subsequent SVD decomposition, the quadruple of state space matrices, A, B, C and D can be extracted from the R-matrix of the QR decomposition and the left singular vectors of the SVD decomposition.

2.4. Observer Designs

The observer-based approach for FDI traditionally uses process variables to estimate output variables through a state-space model in a closed-loop design. The provided output estimation error can then be monitored as residuals (Patton and Chen, 1997). As the design of such observers may offer degrees of freedom, many approaches have been proposed to tackle uncertainties and disturbances issues. This paper uses the traditional Luenberger state observer designed by pole placement.

2.5. Univariate Statistical Fault Detection Indices

For fault detection, two traditional univariate statistics are employed for multivariate statistical process monitoring, the Hotelling's T^2 statistic:

$$T^2 = t^T \Lambda^{-1} t \tag{1}$$

which follows an F-distribution and a squared prediction error:

$$SPE = r^T r \tag{2}$$

for which approximate distributions have been proposed in Box (1954). In Equation 1, t denotes an uncorrelated mean-centered variable set and Λ, its covariance matrix.

Table 1 shows the definition of these statistics for each method.

Table 1 – Fault Detection Indices

Technique	**Principal Subspace**	**Residual Subspace**
Based on DPCA	$t^T \Lambda^{-1} t$	$\tilde{z}^T \tilde{z}$
Based on CVA	$\hat{x}^T \Lambda^{-1} \hat{x}$	$\tilde{x}^T \tilde{x}$
Based on observer	$\hat{x}^T \Lambda^{-1} \hat{x}$	$e^T e$

The work in reference Nomikos and MacGregor (1995) allows the determination of confidence limits for each of these statistics with a significance of 0.01 for example.

2.6. Fault Isolation Indices

This task relates to the determination of which process variables are mostly affected by a detected fault condition. The research literature has proposed two different concepts: (i) the use of contribution chart (Kourti and MacGregor, 1996) and (ii) variable reconstruction (Dunia and Qin, 1998). These methods are readily applicable for DPCA (Qin and Li, 2001) and CVA (Lee *et al.*, 2006). The observer-based method relies on the use of generalized observers, as discussed in Isermann (2005). More precisely, each output variable is reconstructed in turn by eliminating the associated column from the observer setting and using the remaining output variables. Table 2 summarizes the isolation indices for each method in which *a* implies "affected", *i*

denotes the *i*th process variable and * represents a vector reconstructed by excluding the *i*th process variable.

Table 2 Isolation Indices

	Variable Contribution		**Variable Reconstruction**
Technique	*Principal Subspace*	*Residual Subspace*	*Residual Subspace*
Based on DPCA	$t_a\lambda_a^{-1}p_{a,i}z_i$	$\tilde{z}_i^T\tilde{z}_i$	$\frac{\tilde{z}_i^{*T}\tilde{z}_i^*}{\tilde{z}^T\tilde{z}}$
Based on CVA	$\hat{x}_a j_{a,i} z_i$	$\tilde{x}_a j_{a,i} z_i$	$\frac{\tilde{x}_i^{*T}\tilde{x}_i^*}{\tilde{x}^T\tilde{x}}$
Based on observer	-	$e_i^T e_i$	$\frac{e_i^{*T}e_i^*}{e^Te}$

3. Industrial Application Example

This section presents a debutaniser process used to contrast the performance of each method in detecting and isolating a fault condition. The process is designed to purify a mixture of hydrocarbons into butane from a fresh feed which is product of a depropaniser distillation. Therefore, the fresh feed is not steady and can interfere in the top and bottom temperatures by an incipient change. The recorded fault data set contains a small drop in the fresh feed followed by a larger one, which some minutes later affects all the temperatures of the process. In such undesired conditions, the concentration of the head-product is altered, increasing the concentration of pentane. In order to manage the quality of this process, it is required a dynamic model and a fast FDI method.

The industrial data were recorded at a sampling interval of 30 seconds, including 8000 samples describing normal operation conditions for the training data set and 4744 samples containing the fault condition for the testing data set. The input variables are: the reflux flow (u_1), fresh feed flow (u_2) and fresh feed temperature (u_3); and the output variables are: mid-tray temperature (y_1), top-tray temperature (y_2), butane product flow (y_3), bottom-tray temperature (y_4) and inlet temperature (y_5).

4. FDI Results using DPCA, CVA and Observer-Design

This section summarizes the application of the presented fault detection and isolation methods to the recorded data sets. The subspace methods, CVA and MOESP, relied on 2 block rows, whilst for the DPCA, Parallel Analysis showed that the number of principal components does not change with the number of lag variables from 5 lag variables on. Table 3 gives the number of selected components for each method, where for CVA and MOESP it was selected according to the Scree test (Chiang et al., 2001). The observer design was based on the selection of poles inside the unit circle.

For fault detection, three parameters were used to assess the performance of each method: type I error (false alarm rate), type II error (misdetection rate) and the detection delay. Table 3 also summarizes the fault detection results and indicates that the observer-based approach outperformed the statistical-based ones in the last two parameters. Ideally, the methods should have values around 1% for the type I error due to the confidence limit chosen. Such result can be explained due to the fact that the observer-based approach has a well-defined dynamic structure, which guarantee quick and accurate responds compared to CVA-based approach. Regarding to DPCA-based

approach, which dynamic structure is based on an ARX model, the detection delay was significant once the influence of the lag variables (fault-free) in the moment of the fault can be high.

Table 3 - Fault Detection Results

			Type I Error (%)		**Type II Error (%)**		**Detection Delay (min)**	
Method	*l*	*n*	T^2	*SPE*	T^2	*SPE*	T^2	*SPE*
DPCA	5	7	0.7	1.3	20.7	24.8	57	51.5
CVA	1	5	0.4	2.3	27.4	22.4	84.5	50.5
OBS	1	5	0	0.1	5.4	14.4	27	34

Next, the performance of each method for fault isolation was contrasted. Table 4 summarizes the results of each method using contribution charts and Table 5, using variable reconstruction approach. Letters *a* and *b* in Tables 4 and 5 stands for the first and the second fault detection indices, respectively.

The fault isolation step via contribution charts showed divergent results among the methods. DPCA-based approach presented the best description of the fault since the residuals were able to detect the minimal increasing in the fresh feed, and the T^2 contributions could reveal the most affected variable. From this result, the root cause of the fault can be clearly identified and the fault diagnosis can be correctly performed as a trouble in the feed stream. CVA-based approach had similar result compared to DPCA only for the model subspace. However, the "faulty" variable (u_2) was missed in such way that impaired the correct fault diagnosis. Finally, the observer-based approach was unable to make a reasonable description of the fault since the proposed fault detection index for the state variables (OBSa) does not allow a kind of contribution and the residuals are restrictive only to the output variables.

Table 4 - Fault Isolation Results Via Variable Contribution

	Variable Contribution							
Methods	u_1	u_2	u_3	y_1	y_2	y_3	y_4	y_5
DPCAa	40.3	3.8	46.2	**734.2**	64.4	34.2	6.1	43.8
DPCAb	16.3	**66.6**	1.5	20.5	1.1	17.5	0.3	2.3
CVAa	0.9	1	4.9	**49.8**	0.5	2.3	21.8	8.4
CVAb	3.7	2.8	4.3	54.7	**157.1**	2.7	37.4	27.9
OBSa	-	-	-	-	-	-	-	-
OBSb	-	-	-	1576	576	263	359	**4626**

The variable reconstruction approach also had good results compared to the contribution charts, concerning the statistical methods. On the other hand, the bank of generalized observers did not show any improvement. Although this approach can make an analysis of the input variables as well as the output variables, the results of these industrial data sets showed that it does not work well. The observers designed for the input variables monitoring were able to predict correctly the output variables so that the influence from the input variables was not significant and the fault isolation turned out to be impractical.

Table 5 - Fault Isolation Results Via Variable Reconstruction

	Variable Reconstruction							
Methods	u_1	u_2	u_3	y_1	y_2	y_3	y_4	y_5
DPCAb	0.89	**0.52**	0.96	0.87	0.98	0.84	1.00	0.94
CVAb	44.80	**32.98**	45.05	36.13	42.71	44.75	34.31	42.48
OBSb	0.000	0.000	0.000	0.019	**0.005**	0.009	0.009	0.010

5. Conclusions

This paper studied the performance of the most traditional fault detection and isolation methods in a chemical process. In order to make a reasonable comparison, univariate statistics for the state variables and the output estimation errors of the observer-based approach were proposed. Results from a real process data showed a very accurate detectability in the case of the observer-based method. The dynamic structure of such method can be handled by pole placement allowing different types of respond. For the fault isolation step, only statistical-based methods were able to make a good description of the fault. Therefore, none of the methods showed a significant advantage in both fault detection and isolation steps simultaneously. A hybrid approach unifying the strengths of such methods will be the focus of future works.

References

C. Lee, S. W. Choi, I.-B. Lee, 2006, Variable reconstruction and sensor fault identification using canonical variate analysis, Journal of Process Control, 16, 747-761.

C. T. Chen, 1999, Linear Systems Theory and Design. Oxford University Press, 3rd Edition.

G. E. P. Box, 1954, Some Theorems on Quadratic Forms Applied in the Study of Analysis of Variance in One-Way Classification, The Annals of Mathematical Statistics 25, 290-302.

J. Chen, R. J. Patton, 1999, Robust Model-Based Fault Diagnosis for Dynamic Systems, Kluwer Academic Publishers.

J. F. MacGregor, H. Yu, S. G. Muñoz 1, J. Flores-Cerrillo, 2005, Data-based latent variable methods for process analysis, monitoring and control, Computers and Chemical Engineering 29, 1217–1223.

L. H. Chiang, E. L. Russell, R. D. Braatz, 2001, Fault Detection and Diagnosis in Industrial Systems, Springer.

M. Verhaegen, 1994, Identification of the deterministic part of MIMO state space models given in innovations form from input-output data, Automatica 30, 61-74.

P. Nomikos, J. F. MacGregor, 1995, Multivariate SPC Charts for Monitoring Batch Processes, Technometrics 37, 1, 41-58.

P. Van Overschee, B. De Moor, 1996, Subspace Identification for Linear Systems – Theory, Implementation and Application, Kluwer Academic Publishers

R. Dunia, S. J. Qin, 1998, A Unified Geometric Approach to Process and Sensor Fault Identification and Reconstruction: The Unidimensional Fault Case, Computers and Chemical Engineering 22, 7-8, 927-943.

R. Isermann, 2005, Model-based fault-detection and diagnosis – status and applications, Annual Reviews in Control 29, 71–85.

R. J. Patton, J. Chen, 1997, Observer-based fault detection and isolation: robustness and applications, Control Engineering Practice 5, 5, 671-682.

S. J. Qin, W. Li, 2001, Detection and Identification of Faulty Sensors in Dynamic Processes, AIChE Journal 47, 7, 1581-1593.

T. Kourti, J. F. MacGregor, 1996, Multivariate SPC Methods for Process and Product Monitoring, Journal Of Quality Technology 28, 4, 409-428.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modelling and Control of the Variable Channel Reactor

Mayank P. Patel,[a] Nilay Shah,[a] Robert Ashe,[b]

[a]*CPSE, Dept. Chemical Engineering, Imperial College, London, SW7 2AZ, UK*
[b]*AM Technology, The Heath Bus. & Tech. Park, Runcorn, Cheshire, WA7 4QX, UK*

Abstract

The full potential of a novel continuous reactor is addressed via the relationship between process design, the ability in handling the constraints and the optimisation framework. A Model Predictive Controller is developed to maximise the yield under hard input and state constraints. Simulations show that the designed control system gives high reaction yield and ensures that the temperatures inside the reactor are within a predefined limit. The challenges of an efficient start-up phase and the transition to the optimal operating point are also investigated by means of an event-driven approach.
Keywords: Process Intensification, process control, MPC, HEX reactor start-up

1. Introduction

The synthesis of fine chemicals and pharmaceuticals are generally carried out in a batch or semi-batch processes due to flexibility and versatility of the equipment over the continuous counterpart [Roberge *et al.*; 2005]. However, poor transfer of heat generated from reactions limits the application of batch equipment to products of low volumes and short life-times.

A new concept of small-scale continuous reactor, the Variable Channel Reactor (VCR) is being developed by AM Technology. The VCR is well incumbent under the umbrella of process intensification, as it simultaneously tolerates high reactant concentrations whilst minimising solvent usage. It is evident that reaching the optimal operating point when starting from a cold, empty reactor is a non-trivial matter.

Here we propose a control and event-driven start-up strategy with the objective of optimising the reactant conversion with constraints on both control and state variables.

2. The Variable Channel Reactor

The VCR is a modular construction of reactor plates. Through initial process design, inherent control of the temperature profile is achieved by altering the geometry of the channel along the process fluid flow path, as seen in Figure 1. The reaction power is thus controlled by varying the volume to heat transfer area. In this paper, a simple second order exothermic reaction is considered where C is the product of interest.

$$A + B \longrightarrow C + \text{heat}$$

The VCR assembled as multiple consecutive channels may, from a modelling point of view, be approximated as a 1-D continuous tubular reactor with injections of reactant B along the reactor with cooling jackets around the tube. Initial experiments indicate perfect mixing conditions after a few channels, implying that the Arrhenius law can be used for modelling the reaction kinetics. From first principles, partial differential equations (PDEs) of the distributed system are discretised into elements which characterise the states of reactor temperature T and concentration C of components i.

The spatial derivatives are approximated with a first order backward finite difference method (BFDM) to a system of ordinary differential equations (ODEs);

Material balance: $$\frac{\partial C_i}{\partial t} = -\bar{v} \cdot \frac{1}{L} \cdot \frac{\partial C_i}{\partial z} + D \cdot \frac{1}{L^2} \cdot \frac{\partial^2 C_i}{\partial z^2} + \sum_{j=1}^{\text{NoReac}} \left(v_{i,j} \cdot r_j \right) \qquad i=1,\ldots, \text{NoComp}$$

Energy balance: $$\bar{\rho} \cdot \bar{C}_P \cdot \frac{\partial T}{\partial t} = -\bar{\rho} \cdot \bar{C}_P \cdot \bar{v} \cdot \frac{1}{L} \cdot \frac{\partial T}{\partial z} + k \cdot \frac{1}{L^2} \cdot \frac{\partial^2 T}{\partial z^2} + Q - Q_{rxn}$$

where, $$Q = \frac{S}{H} \cdot U \cdot (T_{wall} - T) \; ; \qquad Q_{rxn} = \sum_{j=1}^{\text{NoReac}} (\Delta H \cdot r)$$

Reaction rate: $$r_j = A_j \cdot \exp^{\left(-E_a / R \cdot T\right)} \cdot \prod_{i=1}^{\text{NoComp}} C_i^{a_{i,j}} \qquad j = 1,\ldots, \text{NoReac}$$

The terms Q and Q_{rxn} refer to the heat removal rate of the reactor element and the heat generation rate from the reaction respectively. The assumption of constant wall temperatures, T_{wall}, considers the VCR as 'defacto' isothermal. It is worth noting that the derived model is highly nonlinear due to temperature influence on the reaction rate.

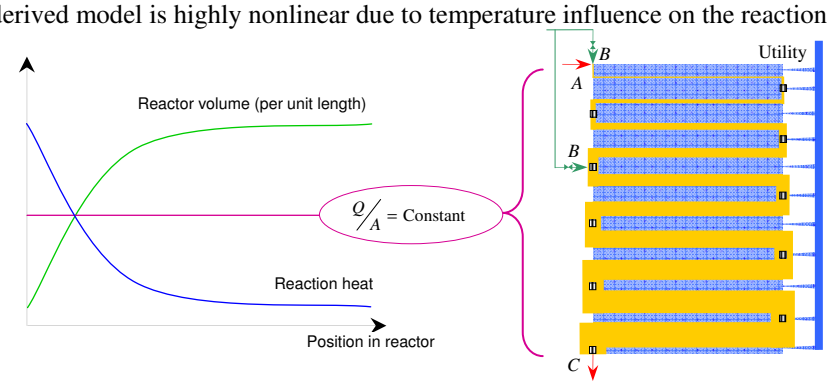


Figure 1. *Left:* The VCR concept. *Right:* A schematic of the VCR, with reactant and utility flows, i.e. 0.25 g/s and 290 g/s respectively. Temperature probes are within the vertical channels.

3. Process Operation

The main focus is on the number and locations of the injections and the cooling zones, as these significantly affect reactor performance and control conditions [Fogler; 1992]. Steady state temperature and conversion profiles are plotted for an exothermic reaction in Figure 2. Reactant *B* is distributed with 60% at the inlet and 40% in channel 9 (positioned at 40% along the reactor's length). The cooling capacity also changes with the latter injection. A maximum temperature, T_{max}=30 °C is in place to avoid product degradation above this limit and which inadvertently limits production capacity.

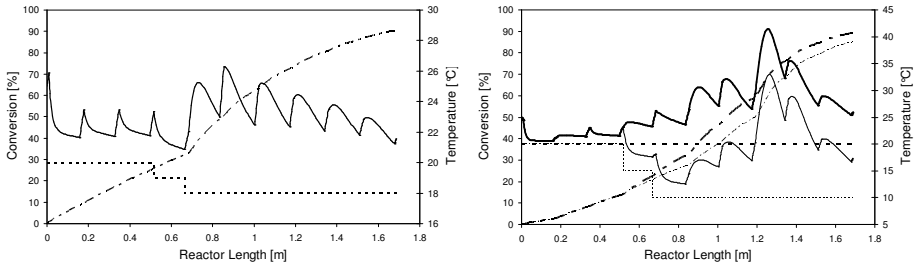


Figure 2. *Left:* Simulated process temperature profile (solid), conversion (dash-dot) and cooling (dashed) with injection at two sites giving a temperature maxima of T=26.3 °C. *Right:* Same as *left*, but here five injections are used with one (thick) and two (thin) cooling zones.

The importance in accurately representing the uncertainty in parameters such as reaction kinetics, micro-mixing conditions, heat transfer capacity as well as nonlinearities in the process dynamics, is crucial to give confidence in the model output [Shah *et al.*; 1999].
The overall control objective is an optimisation of the reactant conversion, γ, defined as the ratio between the amount of product formed to a theoretical conversion of 100%.

$$\max_{u\in U} \gamma(x,u,d) = \frac{C_C}{C_A + C_C} \quad \text{so that} \quad Wx \le \Omega$$

where x are the state variables; temperatures and concentrations along the reactor, u are the three control signals, U is the admissible control set, d are exogenous disturbances with the state constraints defined as W and Ω.

4. The control strategy and design

The configuration indicated in Figure 2 is applied for simplicity, i.e. two injections of reactant *B* and two cooling zones. In this paper, three main control signals will be investigated; the injection flow of reactant *B*, u_1, and the inlet temperature of both cooling utilities, u_2 and u_3. To guarantee stoichiometric balance, the total flow rate of reactant *B* is fixed (at 0.25 g/s) and the control signal u_1 denotes the fraction injected at the first point. The remainder, $(1-u_1)$ is injected at the second point. The total flow rate of reactant *B* is thus constant at all times as are all other inputs.
Two control approaches are available for the VCR; decentralised and centralised. In decentralised control, control loops are designed independently with automatic modes to give feedback such as PID controllers. In centralised control, the feedback control is based on a multivariable process model. As a result all control inputs and measurements are computed simultaneously. An accurate process model is thus required to implement centralised control, whereas decentralised control does save on modelling effort and is easier to tune. The centralised controller in most cases improves control performance, due to its ability to consider the cross-couplings between control inputs.

4.1. Model Predictive Control

The reaction used in the simulations is the preparation of *cis*-5-norbornene-endo-2,3-dicarboxylic anhydride, a fast Diels-Alder exothermic single liquid-phase reaction. The kinetic reaction parameters at ambient temperature are; A_j=1.1·10^5 L/(mol·s), E_a=35700 kJ/kmol and ΔH=−96600 kJ/kmol.
A linear MPC controller is developed based on notations from Maciejowski [2002] and Morari *et al.* [2004]. The nonlinear model is linearised around a working point (x^0,u^0) and is sampled with h=1 s to a discrete-time system. The VCR is discretised in 20 elements. There are six states within each element, the reactor and two utility temperatures, with the concentrations of the reactants and product, giving the linear model 120 states.
The controlled outputs z are chosen to be the outlet concentrations of the two reactants. The cost function to be minimised is employed from Maciejowski [2002] and the reference trajectories are set to r=[0 0]^T, corresponding to γ=100%.

$$V(k) = \sum_{i=1}^{H_p} \left\| \hat{z}(k+i|k) - r(k+i|k) \right\|_{Q(i)}^2 + \sum_{i=0}^{H_u-1} \left\| \Delta\hat{u}(k+i|k) \right\|_{R(i)}^2$$

The temperatures inside the reactor are subjected to hard constraints due to product degradation, $T_i \le T_{max}$=30 °C, where T_i is the temperature in the i:th element of the model. There are also constraints on the control signals, $0.1 \le u_1 \le 0.9$ and 5 °C $\le u_2; u_3 \le$ 30 °C, with the changes $|\Delta u_1| \le 0.1$ and $|\Delta u_2; \Delta u_3| \le 1$.
The prediction and control horizons are set to H_p = 200 and H_u = 10 respectively, as a

result the prediction time is $H_p \cdot h$=200 seconds. The weighting matrices for the controlled variables Q and the control actions R are chosen as Q=$10^{-5}\cdot\mathbf{I}$ and R=[1000 0; 0 1; 0 1], so to remain below the temperature constraint of 30 °C.
The VCR is subjected to an unmeasured inlet disturbance d namely the inlet concentrations. The temperatures of the reactor walls, the inlet and outlet cooling utilities are the available state measurements, y.
The MPC controller and the nonlinear model are implemented in Matlab and gPROMS. An extended Kalman filter (EKF) is designed from a tutorial given by Dochain [2003]. The filter calculates the gain matrix $P(\hat{x})$ as the solution of the dynamical Riccati matrix equation, thus minimising the variance of the estimation error. R_1 and R_2 are the variances of the process and measurement noise, respectively.

$$\frac{\mathrm{d}\hat{x}}{\mathrm{d}t} = -f(\hat{x},u,d) + K(\hat{x})(y-\hat{y}) \qquad ; \qquad K(\hat{x}) = \frac{P(\hat{x})C_y^T}{R_2}$$

$$\frac{\mathrm{d}P}{\mathrm{d}t} = -\frac{PC_y^T \cdot C_y P}{R_2} + PA^T(\hat{x}) + A(\hat{x})P + R_1$$

4.2. Event-driven start-up control
Upon start-up, temperatures within the VCR should guarantee that a reaction starts and that the reactants are consumed, for example at the first reactant B injection point T_1. To satisfy these two conditions, the start-up sequence is divided into four modes, schematically shown in Figure 3. With an open-loop start-up control procedure, it is possible to find the best operating point, whilst obeying the temperature constraints, However, this is difficult to isolate a suitable route in taking the process from initial rest, through to start-up to optimal operation. Therefore an event-driven start-up as described in Figure 3 is implemented with feedback control. For each step of the start-up, the already developed MPC controller due to its multivariable nature and its capacity to handle state constraints is applied.

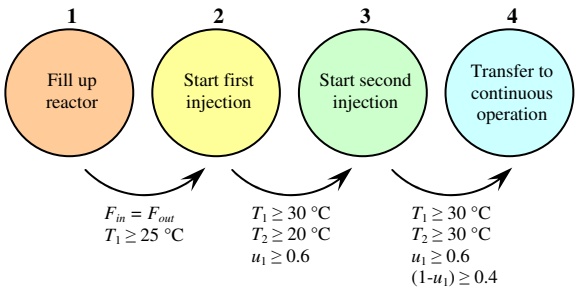


Figure 3. A finite state machine illustrates the steps during start-up.

5. Simulation results

The process starts at rest in a non-optimal working point. The rejection property of the EKF and the MPC controller is tested with a ramp increase of 5% in the concentration of both reactants during t=50-100 s, see Fig. 4 and 5. This increase increases the heat released from reaction and the associated risk of violating the temperature constraint.
The signal u_1 increases from 0.6 to 0.65 within the first 50 seconds, and thus redistributing reactant B from the second injection. Temporary changes in stoichiometry cause a drop in the product yield during the first 70 s. The flow from the second injection point, i.e. 40 % along the reactor length, to the exit is around 20 s which

explains the time delay. The MPC controller optimises conversion by increasing the temperatures inside the reactor by increasing the cooling temperature u_2 and u_3 by 0.8 °C and 1.5 °C, respectively. However, since the cooling temperature has a limited impact on the reactor temperature in the first part of the reactor, the reactant injection distribution was used to increase the temperature there. After t=50 s, the temperatures throughout have reached the prerequisite constraint of 30 °C. To maximise the yield, the temperatures are kept just below this constraint level, which indeed is the main limitation to the reactor's performance.
During the disturbance, the disturbance rejection property of the EKF and the MPC controller reacts by decreasing u_1 that is, redistributing more reactant *B* to the second injection and increasing the cooling rates, see Figure 5. With the estimation, the MPC controller can ensure that the temperature constraints are only slightly violated during the ramp and afterwards in steady-state the temperatures stay on the constraint. This is the main advantage of multivariable model-based control [Qin and Badgwell; 2003].

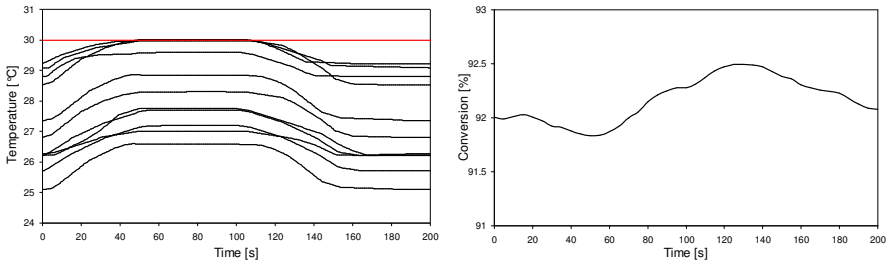


Figure 4. *Left:* Temperatures within each channel containing temperature probes whilst being under a ramp disturbance. *Right:* Outlet conversion under the 5% ramp disturbance.

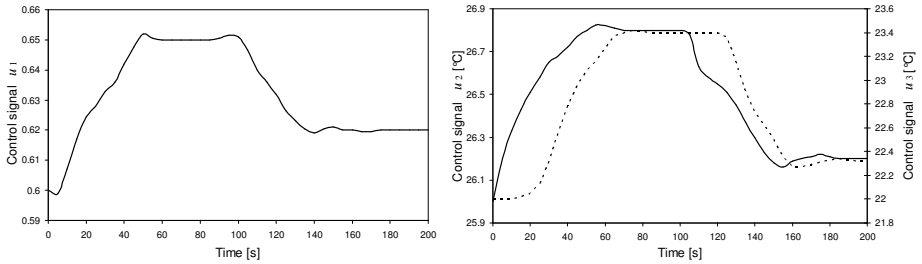


Figure 5. Control signals for the VCR. *Left:* Injection flow of reactant *B*, u_1. *Right:* Cooling utilities to the channels 1-7, u_2 (solid; right axis); and to channels 9-21, u_3 (dashed; left axis).

The closed-loop start-up control procedure is simulated with the non-linear process model is shown in Figure 6. The switch between step 1 and 2 in Figure 3 occurs at t=0, with reactant *A* is flowing through the reactor at steady-state.
Reactant *B* is injected to a channel 1 at t=0 s. As the inflow increases to its outflow the controller increases the utility to follow the temperature reference. After 3 seconds, the temperature at the first injection has reached 26 °C, meaning the reaction has started.
At t=22 s the controller switches from step 2 to 3. The inlet flow rate at the injection point in channel 9 is now included as a control variable. The temperature reference of 20 °C is tracked, and as the second feed increases, the cooling is intensified.
At t=36 s transition between step 3 and 4 occurs. At this step, the controller switches from start-up mode to continuous operation mode. The MPC now aims to optimise for product conversion, by increasing the reactor temperature to the highest allowable temperature before product degradation. In turn, some of reactant *B* is redistributed from

the second injection to the first, which decreases the need of cooling and saving energy. The transfer to optimal operation in step 4 allows the reactor to reach a conversion of 92% at the outlet, compared to 90.6% after step 3.

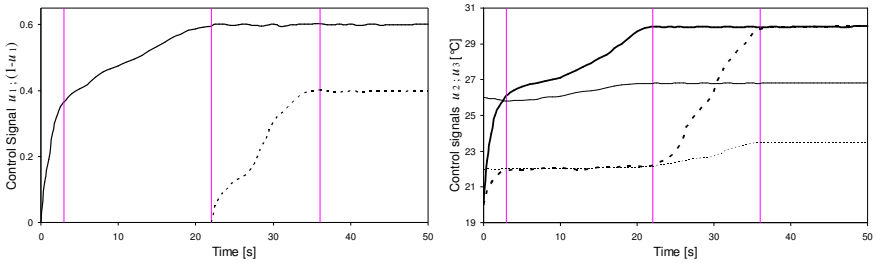


Figure 6. *Left:* During start-up, control signal of injection flow of reactant *B*, u_1 (solid) and 1- u_1 (dashed). *Right:* Temperatures of cooling utilities, u_2 (thick solid); and, u_3 (thick dashed), with temperature profiles within channels 1 and 9, T_1 (thin solid) and T_9 (thin dashed) respectively.

6. Conclusions

The concept and operation of the Variable Channel Reactor (VCR) has been presented. A combination of good mixing and high heat transfer capacity makes the VCR suitable for exothermic reactions and process intensification. The flexible configuration with cooling zones and multiple injections give possibilities for improved process control.

A centralised control is chosen over a decentralised approach, since it takes advantage of multiple inputs multiple outputs (MIMO) models, and offers explicit handling of input and state constraints. Thus, a linear MPC controller is designed, using the injection flow of reactant *B* and the inlet temperatures of cooling utilities to maximise conversion whilst obeying temperature constraints for improved product quality. An extended Kalman filter (EKF) was applied for estimation of unmeasured concentrations. Disturbance estimation was used to reduce the impact from variations in the feed concentrations. The simulations show that high reaction yield can be achieved, without violating the safety constraints. The disturbance and parameter estimation in the EKF increases the robustness of the process control system. By reducing the impact from external disturbances, the risk of unnecessary shutdowns is decreased.

The MPC controller was then applied to the start-up of the VCR which transfers the process from initial state to an operating point. Multiple steady states and multiple inlet ports along the reactor form a process with challenging dynamics. Safe and robust start-up is achieved through an event-driven sequence, i.e. divided into several steps, each associated with a transition condition that needs to be satisfied. The rule of thumb is that the feed of reactant *B* should only begin when the temperature is favourable for the reaction to start. This reduces the sensitivity to uncertainties and aids for good control.

References

D. Dochain, 2003, *Journal of Process Control*, 13, 801-818.

S. Fogler, 1992, Elements of chemical reaction engineering, Prentice Hall.

J. Maciejowski, 2002, Predictive Control with Constraints. Pearson Education Ltd.

M. Morari, J. Lee, C. Garcia, 2004, Model Predictive Control, Prentice Hall.

S. Qin and T. Badgwell, 2003, *Control Engineering Practice*, 11, 73–764.

D. Roberge, L. Ducry, N. Bieler, P. Cretton, B. Zimmermann, 2005, *Chemical Engineering & Technology*, 28(3), 318-323.

N. Shah, N. Samsatli, M. Sharif, J. Borland, L. Papageorgiou, 1999, 4th Intl. Conf. Foundations Computer Aided Process Design, Breckenbridge, USA.

10th International Symposium on Process Systems Engineering - PSE2009
Lenita S. L. Fernandesa, João B. A. Pauloa, Jackson A. Oliveira
© 2009 Elsevier B.V. All rights reserved.

Development of a Strategy to Monitor and Control the Oil-Water Interface Level of a Liquid-Liquid Separator for Treatment of Wastewater Using an Image-Based Detector

Lenita S. L. Fernandes[a], João B. A. Paulo[a], Jackson A. Oliveira[a]*
[a] *DEQ - Departamento de Engenharia Química – CT, Universidade Federal do Rio Grande do Norte Campus Universitário, Lagoa Nova, 59078-970, Natal/RN – Brazil*
**Email: jackson@eq.ufrn.br*

Abstract

The objective of this work is to present the development of a strategy to monitor and control the organic solvent-water interface level on the MDIF® (a mixer-settler based on phase inversion) process. This is one of the most important variables for this equipment. However, the measurement of the organic solvent-water interface level (in line) is still a hard task. There are few sensors able to measure it in laboratory scale systems due to dimensional compatibility. In the present study, a sensor for detecting the organic solvent-water interface level was developed based on the acquisition and treatment of images obtained dynamically through a camera (webcam). In addition, a control strategy was implemented in feedback mode in order to maintain the referred interface level in a desired condition. A control and data acquisition program was built in Compaq Visual Fortran language to perform the following tasks: acquisition and treatment of images for identification of water-solvent interface; decisions and implementations of control signals; and recording of data in files. Some experimental runs in open-loop were carried out using the MDIF® and random pulse disturbances were applied on the input variable (water outlet flow). The responses of interface level permitted the process identification by transfer models. From these models, the parameters for a PID controller were tuned by direct synthesis and tests in closed-loop were performed. Preliminary results for the feedback loop demonstrated that the sensor and the control strategy developed in this work were suitable for level control.

Keywords: MDIF®; process control; image-based sensor; interface level.

1. Introduction

The process imaging use with applications in process engineering has increased significantly more recently. Process information can be extracted from images to estimate important variables as part of a control scheme. Regarding the image sensors, there are three applications field with interest in image detection techniques and analysis: communication and entertainment; medical, security and scientific; and industrial sectors. In the particular case of the chemical process industry, the imaging technology can be found in a significant range of applications, such as: in three dimensional multiphase flow; in the development and control of combustion systems; in detection and measurement of particulate size and shape, mixture uniformity, amount of fluidization, process efficiency, and various factors related to product quality; in particle separation by flotation cells and columns; in thermal imaging for monitoring flow profiles, and so on (Scott and McCann, 2005). Although this technology is already in industrial use, it is still an area in development. The main objective of this work is to present a based-image sensor developed to measure, and subsequently to control, the organic solvent/water interface of a system mixer-settler based on phase inversion (MDIF®) (Fernandes Jr. et al., 2006). The idea is to promote a safer and more efficient

operation for this continuous liquid-liquid separator which is used to treat continuously wastewater of oil industries.

2. Experimental Unit - MDIF®

MDIF® (a mixer-settler based on phase inversion) is a continuous system composed basically by three sections: mixing, settling and separation (Hadjiev et al., 2004). As shown in Figure 1, in a mixing chamber (1), the water contaminated with oil (2) is put in contact with an organic solvent (3) by agitation in order to extract the oil from the water. When this mixture passes through a perforated plate (4), water drops (5) are formed and carry small oil droplets internally (6). The water, which is originally the continuous phase, becomes the dispersed phase (as drops), while the organic solvent that is added to extract the oil becomes the continuous phase of the system. The water drops descend toward to the settling chamber (7) filled with organic solvent, allowing a transfer of the oil droplets from the water drops to the continuous phase (organic solvent). In the separation section (8), the water drops coalesce and form a dense phase (treated water) containing oil in acceptable accordance with the environmental legislation in force, which, in Brazil, establishes maximum total oil and grease (TOG) of 20 mg/L. An important variable during the MDIF® operation is the organic solvent-water interface level (9) in the separation section. In order to measure this interface level, it was developed an image-based sensor, which consists on image detection by a camera (webcam) (10) and its treatment by computational routine. Using a feedback control loop, this measurement is compared to the desired level and a control action is taken according to the PID law, via computational program (11). This action is sent to manipulate the water outlet flow through a control valve (12). The interface level control is essential both to avoid the dragging of the solvent during the water removal and improve the extraction efficiency of the oil by the solvent.

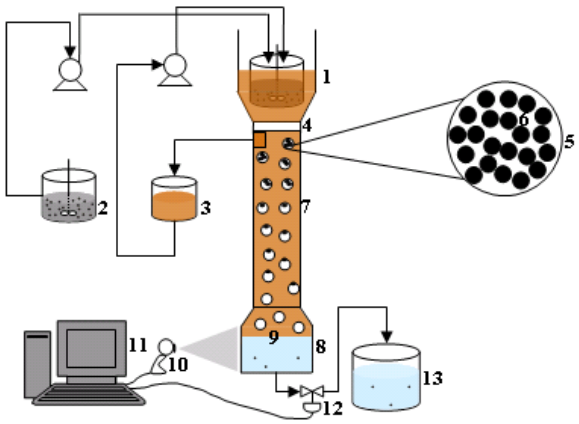


Figure 1 – MDIF® schematic representation.

3. Development of a image-based sensor

The image-based sensor was developed with the objective of detecting the organic solvent/water interface level in the MDIF® separation section. As it mentioned previously, the oil-water interface level control on the separation chamber is extremely important by two reasons: first, with regard to security and quality - the organic solvent-water interface level must be kept far from the outlet to avoid dragging organic to the treated water; second, regarding to the process performance - the organic solvent -water interface level must not ascend in the settling chamber to avoid reducing the residence time of the drops. Moreover, as the MDIF® is in laboratory scale, it is difficult to find an interface sensor with reliable dimensions to this system, and conventional level

sensors are not applied on interface detection. In this study, the image capture was carried out by a simple camera (webcam) and the interface level was estimated using a computational routine, according to the following steps:

1. The camera shows images in real time in preview mode and a static image is captured in grey scale every 3 seconds.
2. Color arrays with 200 discretization points at vertical direction from captured image are obtained in four horizontal positions and a mean is taken, as shown in the Figure 2.
3. From the color arrays mean, a theoretical curve (sigmoid function), represented by the Equation 1, is fitted dynamically through a parameters estimation method based on Particle Swarm Optimization (PSO) (Kennedy and Eberhart, 1995). The first derivative of this curve is used to compute the interface level through the calculation of maximum point. This approach allows to signalize the colors transition between organic solvent phase and water phase (Figure 3).

In this method, the theoretical curve used to fit the color data of the image is given by the expression:

$$f(x) = [a - (a - b).e^{-p_1.e^{-p_2.x}}] \quad (1)$$

where a, b, p1 and p2 are parameters; x indicates the image vertical coordinate; and f(x) is the resultant color in gray scale (0-255).

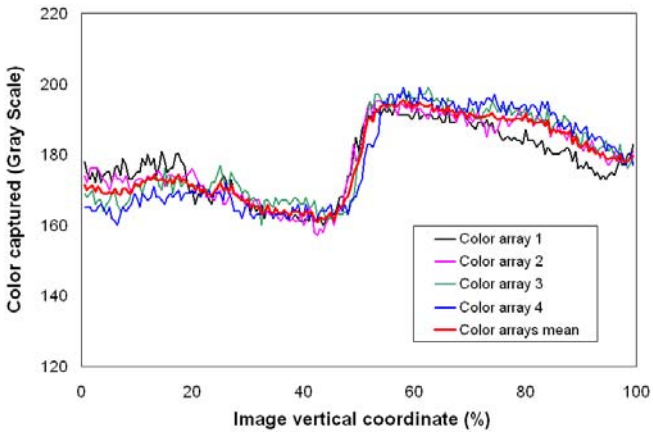


Figure 2 – Color spectra measured from the image in gray scale.

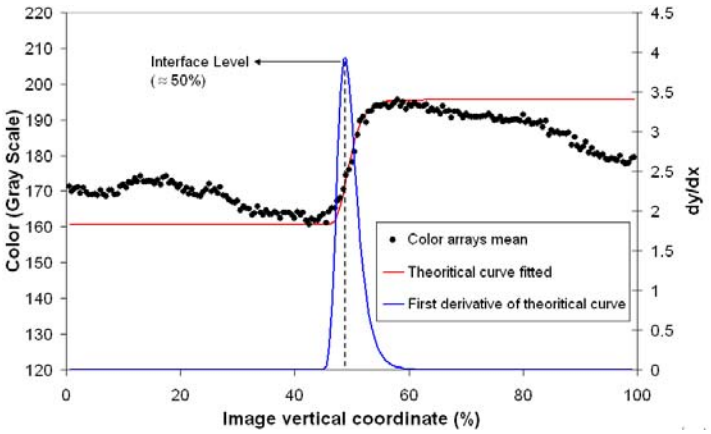


Figure 3 – Sigmoidal fitting with the color spectra mean and its first derivative.

4. Process identification and control strategy

In order to identify the dynamic model capable to describe the behavior of the MDIF® process regarding the interface level, some experiments were performed in open loop from different initial steady state. By applying random pulse disturbances with different magnitude and duration, it was possible to analyze the interface level response from the reaction curves for distinct operational conditions of flowrate. In Figure 4, it can be seen one of studied cases where an attempt of reaching the process identification was led using two transfer models: first-order and second-order functions. The parameters for each function were estimated using the Particle Swarm Optimization method (Kennedy and Eberhart, 1995). The fitted curves allow to show that the models presented a similar behavior. Therefore, it is not possible to discriminate the better model which describes the process (Figure 4). In spite of this, the transfer functions were used to find the PID (proportional-integral-derivative) controller parameters in feedback structure by direct synthesis approach (Ogunnaike and Ray, 1994), admitting a first-order closed-loop trajectory (q) with time constant (τc) of 1.0 minute. The Table 1 lists the adjusted functions for the experiment presented in Figure 4 and the parameters obtained by direct synthesis for PI and PID controllers.

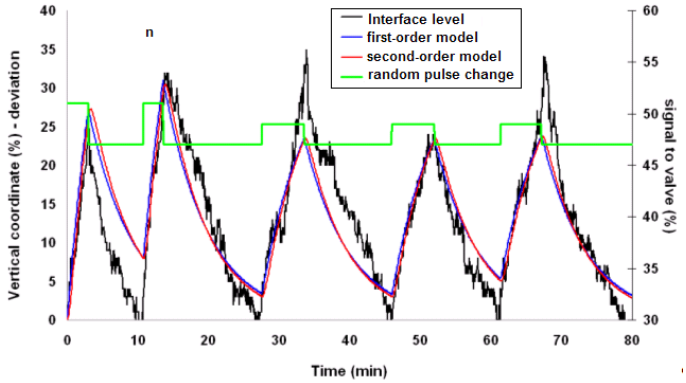


Figura 4 – Dynamic curve for the interface level as response to random pulse changes – experimental and approximate curves.

Tabela 1 – Transfer functions and tuning for feedback PID controller

First-order trajectory: $q(s) = \frac{1}{(s+1)}$

System	Approximate Model	Parameters			Controller
		Kc	τ_D	τ_I	
First-order	$G(s) = \frac{18.1}{(6.32s+1)}$	0.35		6.32	PI
Second-order	$G(s) = \frac{17.84}{(5.78s+1)(0.34s+1)}$	0.34	0.32	6.12	PID

A control and data acquisition program was implemented with the Compaq Visual Fortran compiler (CVF 6.5) as shown in Figure 5. In this program, the following tasks are accomplished: acquisition and treatment of images for detecting the interface level; decisions and implementations of control signals based on PID controller; and recording of all data in file. In addition, as displayed in the program window (Figure 5), the users

can notice the interface level in real-time and to operate the process in manual control mode, if necessary. Other important fact is that the control strategy can be modified in order to improve the controller. In this preliminary study, only PID controller was implemented in this program. However, other control schemes can be adapted easily as routines for the program.

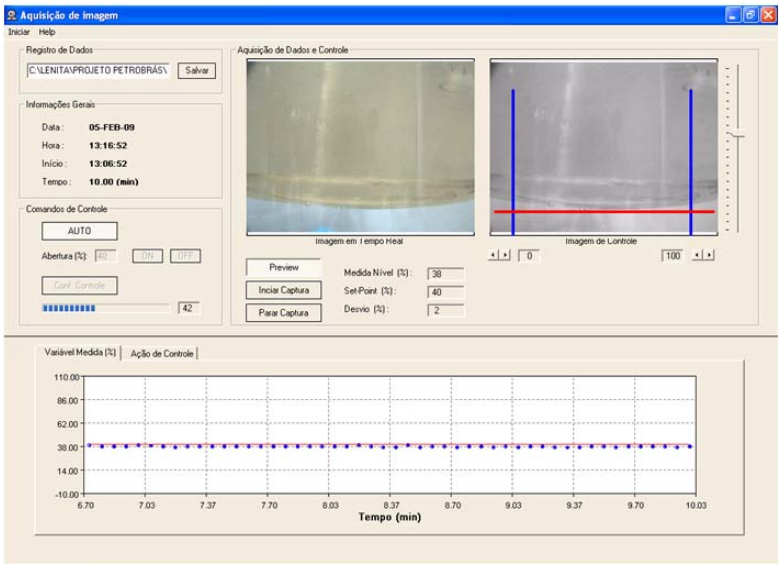


Figure 5 – Main window of the program implemented to monitor and control the MDIF® process.

5. Results and Discussion

In order to investigate the response of the closed-loop system to step changes in the set-point and in the input disturbances (Ogata, 1997), some experiments were carried out with the MDIF® process. For reason of space limit in this document only two experiments are presented here. The particular case of the set-point response for the interface level using the PID controller tuned by direct synthesis (Table 1) is shown in Figure 6. It is possible to note that the PID controller provides a response with short rise time and without offset. Although the response has presented short overshoot and damped oscillations, the controller has demonstrated satisfactory for MDIF® process.

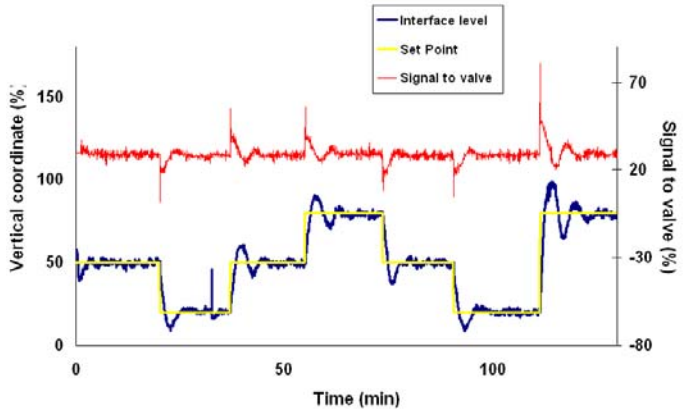


Figure 6 – Closed-loop set-point change responses for the interface level using the PID controller tuned by direct synthesis, as shown in Table 1.

In the case of response to input disturbance, the experiments were performed applying step perturbation at the process water feed through manipulating the pump flowrate. As shown in Figure 7, the step changes were done at time values signalized by dashed lines with magnitude represented by the green points. The response for this case provides overshoots when the changes are implemented. As larger the perturbation, more intense it is the overshoot. Even so, the controller is able to bring back the system to the desired set-point without offset. So, the controller presented suitable results regarding the regulatory control condition.

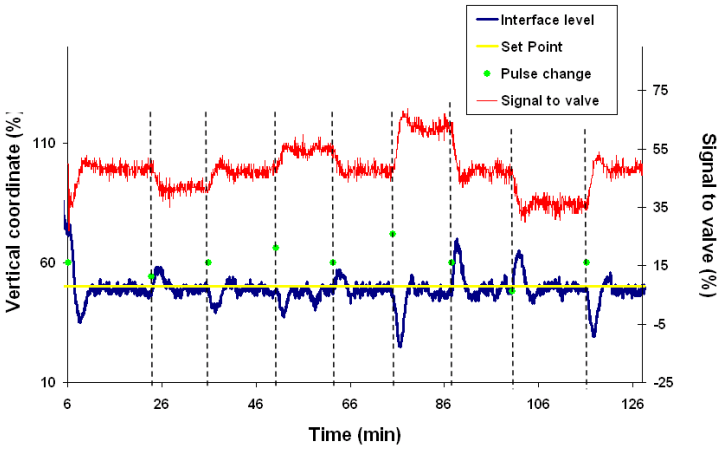


Figure 7 – Closed-loop response to the input disturbance using the PID controller tuned by direct synthesis, as shown in Table 1.

6. Conclusions

The results and analysis obtained showed that the strategy of image detection and treatment was able to measure the organic solvent/water interface level from image color spectra for the MDIF® system. The procedures to the process identification and PID controller design resulted in a closed-loop scheme that provided satisfactory responses both to servo and regulatory control. As this is a preliminary study, other controller tuning rules could be tested in order to enhance the control performance. From a practical point of view, the image-based sensor presented here may be applied as part of a control system used in other processes, where the control of interface level is required.

References

K. Ogata, 1997, Modern Control Engineering, 3ed. USA: Prentice Hall.

D. M, Scott , H. Mccann, 2005, Process imaging for automatic control. USA: Taylor & Francis group, LLC.

B. A. Ogunnaike, W. H. Ray, 1994, Process Dynamics, Modeling, and Control. Oxford University Press.

D. Hadjiev; L. Limousy; N. E. Sabiri, 2004. The design of separators based on phase inversion at low velocities in the nozzles. Separation and Purification Technology, n. 38, p. 181–189.

W. E. Fernandes Jr; J. B. A. Paulo; N. A. Moraes; A. F. Lima; G. M. Lacerda, 2006. Tratamento de águas produzidas por meio de nova tecnologia (MDIF®): aplicação para águas contendo baixas concentrações em óleo. Boletim Técnico da Petrobras, Rio de Janeiro, v.49.

J. Kennedy, R. Eberhart, 1995. Particle Swarm Optimization, in Proceedings of the IEEE Int. Conf. on Neural Networks, pp. 1942–1948.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Dynamic degrees of freedom for tighter bottleneck control

Elvira Marie B. Aske[a,b], Sigurd Skogestad[a]

[a]*Department of Chemical Eng.ineering, Norwegian University of Science and Technology, 7491 Trondheim, Norway.*
[b] *StatoilHydro, Reaserch & Development, 7005 Trondheim, Norway.*

Abstract

To realize maximum throughput, tight control of the bottleneck unit(s) is necessary. Dynamic degrees of freedom can be used to obtain tighter bottleneck control. Here, "dynamic" means that the degree of freedom has no steady-state effect on plant operation, like most inventories (levels). Nevertheless, temporary changes of inventories can allow for dynamic changes in the flow through the bottleneck that keeps the process closer to its bottleneck constraint and increase the throughput.

Keywords: Throughput maximization, bottleneck, inventory, ratio control

1. Introduction

In many cases, prices and market conditions are such that optimal operation is the same as maximizing plant throughput. In this case, the optimum lies at constraints, and in order to maximize throughput, the flow through the bottleneck(s) should be at its maximum at all times (Aske *et.al*, 2008). If the actual flow through the bottleneck is not at its maximum at any given time, then this gives a loss in production that can never be recovered. Tight bottleneck control is therefore important for maximizing throughput and avoiding losses.

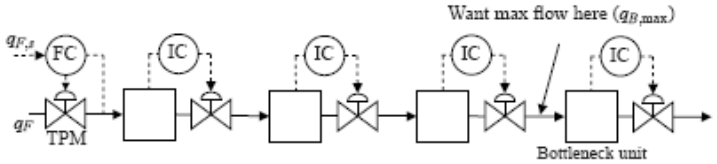
In existing plants, the most common approach for controlling the throughput is to set the feed flow at the inlet of the plant and use inventory control in the direction of flow (Price et al., 1994). One reason for this is that most of the control structure decisions are done at the design stage (before the plant is built), where one usually assumes a fixed feed rate. However, tight bottleneck control requires that the throughput manipulator (TPM) is located close to the bottleneck (Skogestad, 2004). The term "close to the bottleneck" means that there is a short effective delay from the input (TPM) to the output (bottleneck flow).

Ideally the TPM should be located at the bottleneck, but this may not be desirable (or even possible) for other reasons. First, if the TPM is moved, the inventory loops must be reconfigured to ensure self-consistency (Aske and Skogestad, 2009). Second, there may be dynamical reasons for avoiding a so-called on-demand control structure with inventory control opposite the direction of flow (Luyben, 1999). Third, if a bottleneck(s) moves in the plant due to disturbances, then single-loop control requires relocation of TPM and reconfiguration of inventory loops. Thus, in practice one is often left with a fixed throughput manipulator, usually the feed rate. This usually leads to a large effective delay ("long loop") because the bottleneck is usually located inside the plant, and this leads to an economic loss because of a large required back off from the bottleneck constraints.

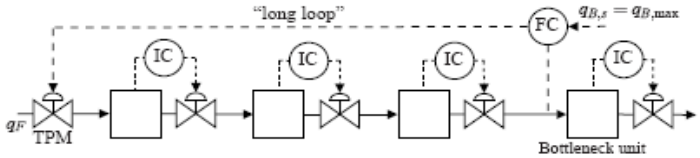
There are also related issues in business systems for using inventories as degree of freedom. Supply chains are sometimes modelled as continuous processes and Schwartz *et al.* (2006) used simulation to study decision policies for inventory management.

2. Alternative strategies for bottleneck control

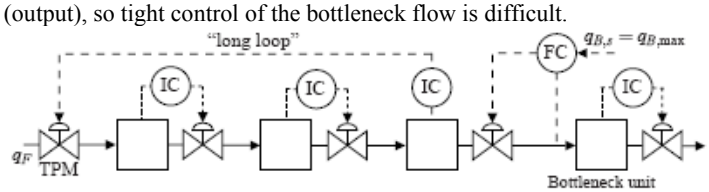
Assume that the objective is to maximize the flow through the bottleneck and that the feed rate is available as a degree for freedom (TPM). Figure 1 show four ways of achieving this using simple single-loop control structures.



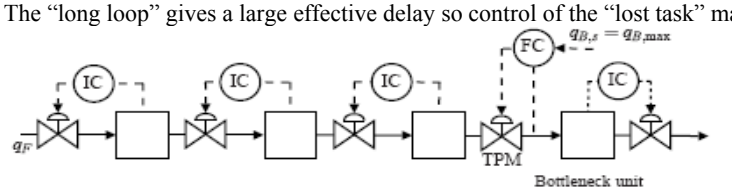
(a) **Traditional configuration** (manual control of feed rate). The feed rate is the degree of freedom for manipulating throughput (TPM), and inventory control is in the direction of flow. To maximize the flow through the bottleneck, the operators change the feed valve manually based on information about the plant operation and experience.



(b) **Alternative 1: Single-loop control of bottleneck flow using the feed rate.** *Problem:* The "long loop" gives a large effective delay from the feed flow (input) to the bottleneck flow (output), so tight control of the bottleneck flow is difficult.



(c) **Alternative 2: Move TPM from feed to bottleneck**. This achieves tight control of the bottleneck flow. The inventory loops are not reconfigured, so the feed rate now needs to take over the "lost task" which in this case is control of the inventory upstream of the bottleneck. *Problem:* The "long loop" gives a large effective delay so control of the "lost task" may be poor.

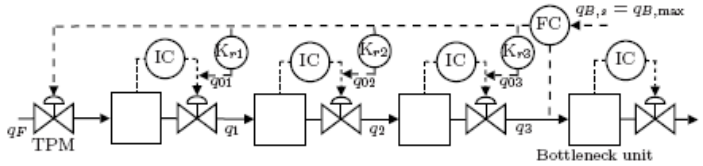


(d) **Alternative 3: Reconfigure inventory control**. The TPM is moved to the bottleneck and at the same time all the upstream inventory loops are reconfigured to be in the opposite direction of flow. Tight bottleneck control (of both flow and local inventory) may be achieved. *Problem:* Reconfiguration of inventory loops is usually very undesirable from a practical point of view.

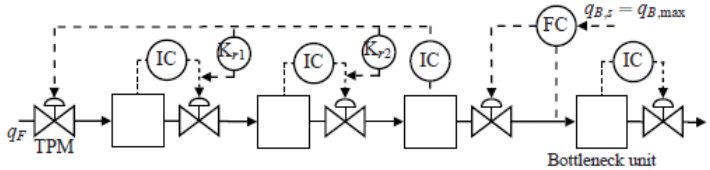
Figure 1: Simple single-loop control structures for maximizing bottleneck flow in serial process. IC stands for inventory controller (e.g. level controller).

In summary, none of the alternatives in Figure 1 are desirable. To improve control and keep the flow through the bottleneck closer to its maximum at all times, we would like to have additional degrees of freedom, and the only ones that are normally available are the inventories (holdups) in the buffer tanks, which can be used to make

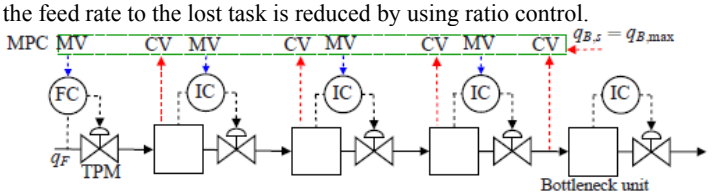
dynamic flow changes. The word ”dynamic” is used because most inventories have no steady-state effect on plant operation.
The main idea is as follows: To change the flow through the bottleneck, for example, to increase it, we temporarily reduce the inventory in the upstream holdup volume. However, this inventory needs to be kept within bounds, so if we want to increase the bottleneck flow permanently, we need to increase the flow into this part of the process and so on, all the way back to the feed (throughput manipulator). The simplest (but not generally optimal) approach is to use a “ratio” control system where all flows upstream the bottleneck are increased simultaneously by the same relative amount. The idea is illustrated in Figure 2.



(a) **Alternative 1D: Single-loop plus ratio control.** The idea is to control the bottleneck flow by simultaneously changing all the flows upstream of the bottleneck by the same relative amount. The advantage is that the effective delay from the feed to the bottleneck may be significantly reduced and even eliminated in some cases. However, the dynamic flow changes are counteracted by the inventory controllers. In particular, note that the feed flow is the only degree of freedom that has a steady-state effect on the bottleneck flow. The strategy may also be viewed as a “ratio feedforward controller” from the feed flow to the downstream flows.



(b) **Alternative 2D: Move TPM to bottleneck and add ratio control to “lost task”**. The TPM is moved to the bottleneck and the “lost task” (inventory upstream the bottleneck) is controlled by the feed rate. The use of ratio control is the same as for Alternative 1D. The effective delay from the feed rate to the lost task is reduced by using ratio control.



(c**) Alternative 4: Multivariable controller**. A multivariable controller (e.g. MPC) uses the feed rate and the inventory controller set points as manipulated variables (MVs). The controlled variables (CVs) are the bottleneck flow and inventory constraints.

Figure 2: Structures for controlling bottleneck flows that use inventories as dynamic degrees of freedom (with no reconfiguration of the inventory loops). Alternative 1D is studied in this paper. IC stands for inventory controller (e.g. level controller) and K_{ri} is a constant gain (ratio controller).

The most obvious is to adjust the inventory set point I_s, but it is more direct in terms of flow changes to adjust the bias q_0. The two approaches are not very different, because a change in q_0 can equivalently be implemented as a set point change by choosing $I_s = -q_0/K(s)$, where $K(s)$ is the feedback controller. In this paper, we choose to use the bias q_0 as the dynamic degree of freedom for ratio control. The important point to note is that

there are no dynamics in K_r. This means that all the flows q are changed simultaneously when q_F changes.

3. Example: Four distillation columns in series

Consider four distillation columns in series, as shown in Figure 3. The four columns represent the liquid upgrading part of a gas processing plant and consist of a deethanizer, a depropanizer, a debutanizer and a butane splitter. Assume that the butane splitter is the bottleneck unit. The throughput is manipulated at the feed to the first column. The idea is to use the column inventories (sump or condenser drum holdup) as dynamic degrees of freedom to obtain tighter bottleneck control.

The distillation column models are implemented in Matlab/Simulink. Each of the four columns is modelled as multicomponent distillation with one feed and two products, constant relative volatilities, no vapor hold-up, constant molar flows, total condenser and liquid flow dynamics represented by the Francis weir formula. All columns use the "LV-configuration" where distillate (D) and bottoms flow (B) are used for inventory control (M_D and M_B). To stabilize the column composition profile, all columns have temperature control in the bottom section by manipulating the boilup. The column inventories M_D and M_B are controlled with P-controllers with gain $K_c = 1/\tau_V$. Here we use "smooth" level control where we set $\tau_V = V_{tank}/q_{out}$ (Skogestad, 2006) where q_{out} is the flow out of the volume (D or B). The temperature controllers (TC) are tuned with SIMC PI-tuning (Skogestad, 2003) with $\tau_c = 0.5$ min.

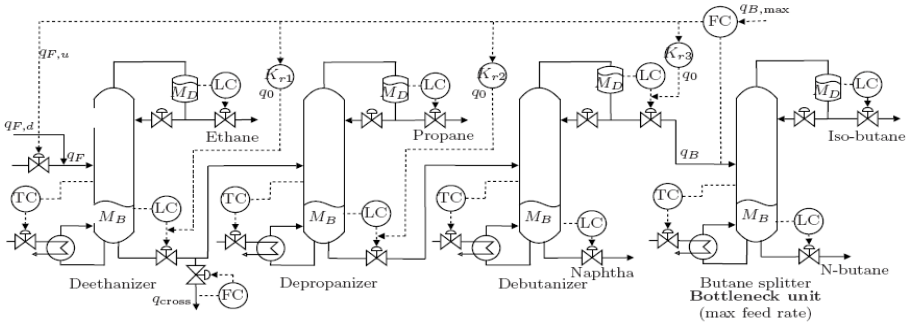


Figure 3: Distillation process: Four columns in series, here shown with throughput controlled by using single-loop with ratio control (Alternative 1D).

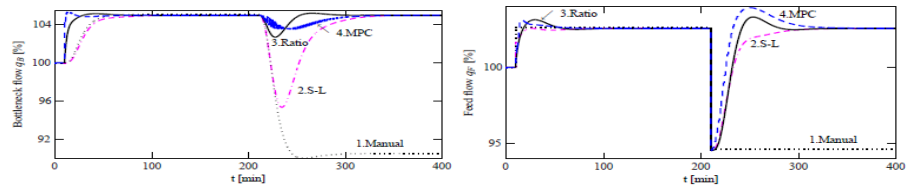
Two disturbances are considered. First, at *t = 10 min*, we make a 5% increase in the bottleneck flow set point ($q_{B,s}$). Second, *at t = 210 min*, there is an 8% unmeasured decrease in the feed rate to the deethanizer (q_F). The net feed flow is $q_F = q_{F,u} + q_{F,d}$, where $q_{F,u}$ is the flow contribution from the controller (initially $q_{F,d} = 0$ and $q_F = q_{F,u} = 100$, but then $q_{F,d} = -8$ at *t = 210*).

Four different control structures are tested for maximizing throughput:

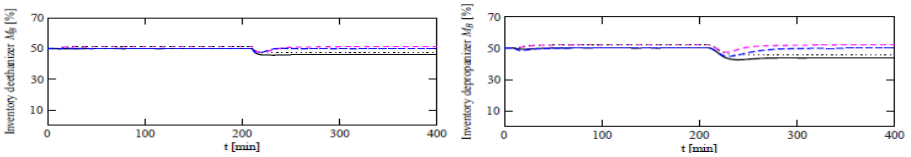
1. **Manual:** Traditional (manual) control of the throughput. We assume that a skilled operator can immediately change the feed rate to the value corresponding to the new bottleneck flow set point. However, we assume that the operator does not notice the unmeasured feed flow disturbance, so no adjustment is therefore done for the feed rate disturbance.
2. **Single-loop:** Single-loop control where the bottleneck flow is controlled using the feed rate (Alternative 1).We want smooth tuning to avoid overshoot and "aggressive" use of the feed valve. Therefore, the bottleneck flow controller (FC) is tuned with SIMC tunings with $\tau_c = 3\theta$ for smooth tuning (Skogestad, 2006).

3. **Single-loop with ratio:** Use of the inventories as dynamic degrees of freedom by adding a bias (q_0) to the inventory controller outputs as in Figure 3 (Alternative 1D). In this case there is no effective delay and the bottleneck flow controller (FC) is tightly tuned with a short integral time, which are typical FC tuning parameters.
4. **Multivariable:** MPC with the feed rate and the inventory set points as MVs and the bottleneck flow and level constraints as CVs (Alternative 4). The built-in MPC toolbox in Matlab is used and tuned with a low penalty on the use of inventories (MV moves) and a high penalty on the deviation from the bottleneck flow set point (CV set point).

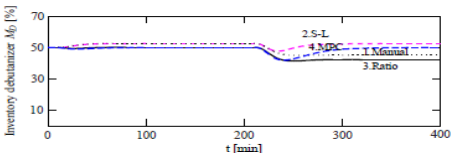
The four control structures are evaluated in terms of how tightly the bottleneck flow (q_B) is controlled in spite of disturbances. The resulting bottleneck flow (q_B), the net feed flow (q_F) and the inventories used as dynamic degrees of freedom (deethanizer M_B, depropanizer M_B and debutanizer M_D) for the four different control structures are displayed in Figure 4. The bottleneck control is significantly tighter with ratio control and MPC where inventories are used as dynamic degrees of freedom. The inventories are quite tightly controlled with surprisingly small variations. This follows because the disturbances introduced here are small compared to what the IC's are tuned to handle.



(a) Response in bottleneck flow q_B (CV) (b) Responses in feed flow q_F (MV)



(c) Responses in deethanizer bottoms level M_B (d) Responses in depropanizer bottoms level M_B



(e) Responses in debutanizer distillate level M_D

Figure 4: Bottleneck control of the distillation process for four different control structures. 1) Manual control (dotted), 2) Single-loop control (dash-dotted), 3) Single-loop with ratio (solid), 4) MPC using both feed rate and inventories as MVs (dashed).

4. Summary: Implications for design of inventory tanks

The effect of using inventories as dynamic degrees of freedom on the design of inventory tanks is summarized. The derivations are given in details in Aske (2009).

4.1. Tank size

A desired change in tank throughput ΔqB results in a volume variation ΔV and we have

$$|\Delta V| = \tau_G \cdot |\Delta qB| \quad (1)$$

where τ_G is the time constant for "refilling" the tank. In practice, τ_G is the time for the flow rate into V to reach 63% of its steady-state change following a step in flow rate out of the (closest) upstream inventory. This is for the normal case when the TPM is upstream the bottleneck; the same formula applies also when it is downstream. For design purposes, the flow change $|\Delta q_B|$ is the (steady-state) flow change through tank resulting from the largest expected throughput (bottleneck flow) change.

Equation (1) is useful for sizing the tank (inventory volume). In words, the expected volume variation for an inventory used for bottleneck control is approximately the expected variation in flow through the unit multiplied by the time constant for the flow dynamics for "refilling" V from the upstream inventory.

4.2. Level control tuning

The level control tuning involves the closed-loop time constant (τ_V) for the level control loop in the inventory tank. We get

$$|\Delta V| = \tau_V \cdot |\Delta q_d| \quad (2)$$

where Δq_d is the flow rate change through the tank in question. Equation (2) can be used to tune the level controller, and then gives the well-known formula for smooth (averaging) level control. To see this, note that for a nominally half-full tank we must require $|\Delta V_{peak}| < 0.5\ V_{tank}$ to avoid overfilling or emptying. If we furthermore assume that the maximum expected change in flow through the tank is 50% of the nominal flow, then $q_d = 0.5\ q$. This gives $\tau_V < V_{tank}/q$, which is the well-known value for smooth level control, (e.g. Skogestad (2006)).

References

Aske, E., 2009, Design of plantwide control systems with focus on maximum throughput. Ph.D. thesis, Norwegian University of Science and Technology, Trondheim, Norway, available at www.nt.ntnu.no/users/skoge/publications/thesis/2009_ aske/.

Aske,E., Skogestad, S., 2009. Self-consistent inventory control. Ind. Eng. Chem.Res.Submitted.

Aske, E., Strand, S., Skogestad, S., 2008. Coordinator MPC for maximizing plant throughput. Comput. Chem. Eng. 32 (1-2), 195–204.

Luyben, W., 1999. Inherent dynamic problems with on-demand control structures. Ind. Eng. Chem. Res. 38 (6), 2315–2329.

Price, R. M., Lyman, P. R., Georgakis, C., 1994. Throughput manipulation in plantwide control structures. Ind. Eng. Chem. Res. 33, 1197–1207.

Schwartz, J., Wang, W., Rivera, D., 2006. Simulation-based optimization of process control policies for inventory management in supply chains. Automatica 42, 1311–1320.

Skogestad, S., 2003. Simple analytic rules for model reduction and PID controller tuning. J. Proc. Control 13, 291–309.

Skogestad, S., 2004. Control structure design for complete chemical plants. Comput. Chem. Eng. 28, 219–234.

Skogestad, S., 2006. Tuning for smooth pid control with acceptable disturbance rejection. Ind. Eng. Chem. Res. 45, 7817–7822.

Skogestad, S., 2007. The dos and don'ts of distillation column control. Trans.IChemE, Part A 85 (A1), 13–23.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Global Optimization Approach for the Estimation of Domains of Attraction

Luis G. Matallana, Aníbal M. Blanco, J. Alberto Bandoni
PLAPIQUI (UNS - CONICET), Camino "La Carrindanga" – km 7
(8000) Bahía Blanca, ARGENTINA

Abstract

In order to completely characterize an asymptotically stable equilibrium point, some information about the size and shape of its Domain of Attraction is required. In this contribution a global optimization approach is proposed to estimate domains of attraction of general nonlinear dynamic systems. The technique is illustrated by a two states system that presents a very rich nonlinear behavior and is then applied to a typical continuous stirred tank reactor.

Keywords: Nonlinear dynamic systems, Domains of attraction, Global optimization

1. Introduction

The Domain of Attraction (DOA) of an asymptotically stable equilibrium point represents the portion of the state space where trajectories that converge to the equilibrium point originate. For the general nonlinear case, the DOA is a complicated set that does not admit analytical representation. Many techniques have been proposed so far to address the estimation of DOAs. In particular, the Lyapunov stability theory provides the basis of a family of techniques whose rationale is to approximate the DOA by a level set of a Lyapunov function. Several approaches have been proposed to identify the best level set of a given Lyapunov function by solving an optimization model. This problem is very challenging for the general case since the resulting optimization model is nonlinear semi-infinite. Formulations that make use of results on deterministic global optimization were recently proposed (Hachicho, 2007). The technique allows the identification of the best possible level set of a rational Lyapunov function that constitutes an estimation of the DOA of the equilibrium under study for nonlinear dynamic systems of polynomial type. In this contribution, an extension of the formulation presented in Hachicho (2007) is proposed. An optimization model is formulated, which includes additional constraints to avoid possible dummy solutions and makes use of a global optimization approach to address nonlinear systems of the general type.

2. Estimation of DOAs

Consider the following autonomous nonlinear dynamic system:

$$d\mathbf{x}/dt = \mathbf{f}\left(\mathbf{x}(t)\right) \tag{1}$$

being the origin of the state space, $\mathbf{x} = \mathbf{0}$, an asymptotically stable equilibrium point. The DOA of the origin is defined by the set of initial conditions $\mathbf{x}_0$, such that the trajectories initiated at $\mathbf{x}_0$ (denoted by $\mathbf{x}(\mathbf{x_0}, t)$) converge to the origin as time increases:

$$\mathrm{DOA}(\mathbf{0}) = \left\{\mathbf{x}_0 : \mathbf{x}(\mathbf{x_0}, t) \to 0 \text{ as } t \to \infty\right\} \tag{2}$$

Consider that a function $V(\mathbf{x})$ exists such that:

$$V(\mathbf{x}) \succ 0 \text{ in } \mathrm{R}(\mathbf{0}) \tag{3a}$$

$$dV(\mathbf{x})/dt \prec 0 \text{ in } \mathrm{R}(\mathbf{0}) \tag{3b}$$

where R(**0**) is a region around the origin. Equation (3a) means that $V(\mathbf{x})$ is positive definite in region R(**0**) and (3b) that its time derivative is negative definite in that region. If (3a) and (3b) hold for system (1) then $V(\mathbf{x})$ is known as a Lyapunov function of (1). Let $V(\mathbf{x}) = c$ be the level set of $V(\mathbf{x})$ at value c, this is, the projection on the state space of $V(\mathbf{x})$ at value c. Consider the region S(**0**)={**x**: $V(\mathbf{x}) - c \leq 0$} is contained in R(**0**):

$$\mathrm{S}(\mathbf{0}) \subset \mathrm{R}(\mathbf{0}) \tag{4}$$

then S(**0**) belongs to the DOA of the origin and can be considered an estimation of DOA(**0**). A powerful technique based for the design of "maximal" Lyapunov functions has been proposed in Vannelli and Vidyasagar (1985). A maximal Lyapunov "candidate" is a rational function with the following structure:

$$V(\mathbf{x}) = \frac{N(\mathbf{x})}{D(\mathbf{x})} = \frac{\sum_{i=2}^{n} R_i(\mathbf{x})}{1 + \sum_{i=1}^{n-2} Q_i(\mathbf{x})} \tag{5}$$

Where $N(x)$ and $D(x)$ are polynomials defined through the $R_i(\mathbf{x})$ and $Q_i(\mathbf{x})$ homogeneous functions of degree i and n ≥ 2. If the system of nonlinear equations is represented in terms of a series of homogeneous functions of degree i, $F_i(\mathbf{x})$:

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}[\mathbf{x}(t)] = \sum_{i=1}^{\infty} F_i(\mathbf{x}) \tag{6}$$

a Lyapunov function can be obtained by the procedure proposed in Vannelli and Vidyasagar (1985) and then an estimation S(**0**) of DOA(**0**) can be calculated by finding the largest positive value of c that verifies (4). The calculation of the maximum level set of a given Lyapunov function can be stated as:

$$\begin{aligned} &\max_{c,\mathbf{x}} \; c \\ &s.t. \;\; V(\mathbf{x}) = c \qquad \text{a)} \\ &\qquad \dot{V}(\mathbf{x}) \prec 0 \qquad \text{b)} \end{aligned} \tag{7}$$

In order to handle the negative definiteness constraint (7b), Hachicho (2007) proposed to reformulate problem (7) as:

$$\begin{aligned} &c = \min_{\mathbf{x}} \; V(\mathbf{x}) \\ &s.t. \;\; \dot{V}(\mathbf{x}) = 0, \quad \mathbf{x} \neq \mathbf{0} \end{aligned} \tag{8}$$

The rationale behind problem (8) is to find the minimum level set of function $V(\mathbf{x})$ which is contained in the level set $dV(\mathbf{x})/dt = 0$. The solution of algebraic model (8) is a single point in the state space, which corresponds to the tangential contact of level sets $V(\mathbf{x}) = c$ and $dV(\mathbf{x})/dt = 0$. In order to graphically illustrate this situation consider dynamic system (9) and Lyapunov function (10).

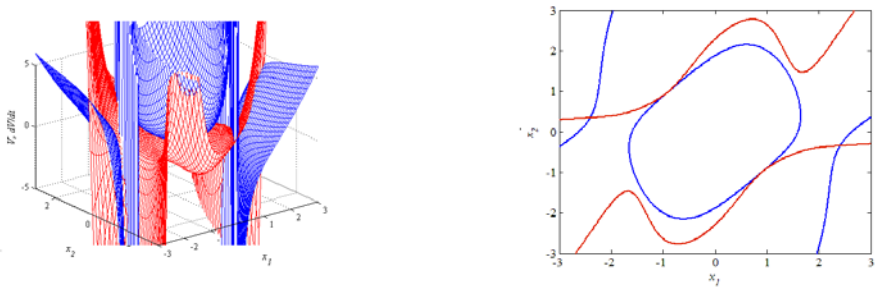
$$\begin{aligned}\frac{dx_1}{dt} &= -x_2 \\ \frac{dx_2}{dt} &= x_1 - x_2\left(1-x_1^2\right)\end{aligned} \tag{9}$$

$$\begin{aligned}N(\mathbf{x}) &= 0.593x_1^2 - 0.364x_1x_2 + 0.437x_2^2 - 0.1253x_1^4 + 0.2885x_1^3x_2 - 0.0537x_1^2x_2^2 \\ &\quad + 0.0581x_1x_2^3 - 0.0196x_2^4 \\ D(\mathbf{x}) &= 1 - 0.0001x_1 + 0.0001x_2 - 0.2685x_1^2 + 0.3217x_1x_2 - 0.1163x_2^2\end{aligned} \tag{10}$$

In Fig. 1a) it is shown the Lyapunov function and its time derivative. In Fig. 1b) the level sets of interest are depicted. The closed curve $V(\mathbf{x}) = 2.2086$ (blue line) contained in the level $dV(\mathbf{x})/dt = 0$ (red line) constitutes the desired solution of problem (8) for the system under study (9)/(10). It should be noted that, due to symmetry, there are two points in the state space for the same $V(\mathbf{x}) = c$ for this system. Since both solutions are equivalent regarding the resulting estimation only one of them is reported.

3. Proposed Approach

Problem (8) is a nonlinear optimization model and therefore it may have many local solutions. A local solution may not correspond to a proper estimation of the DOA for the system under study. Consider for example the different solutions of problem (8) for system (9)/(10) shown in Fig. 2. In Fig. 2 (a to d) different level sets for the Lyapunov function are shown in blue line. In all cases the level set of the time derivative at value zero is shown in red line. The solution depicted in Fig. 2 a) is a solution of problem (8) with a value of c lower than that of the desired estimation (Fig. 1 b)). The solution corresponds to an intersection between the level set $V(\mathbf{x}) = 0.1$ and the level set $dV(\mathbf{x})/dt = 0$. It can be noticed that the intersection takes place far from the origin and therefore a very small portion of the level set of $V(\mathbf{x})$ is in fact an estimation of the actual DOA (the small circle around the origin). The solution of Fig. 2 b) verifies tangential contact between the level sets. The situation is pretty the same as in the previous case. The tangency between level sets takes place far from the origin and the estimation of the DOA is also poor. It can be observed in this case, that although the constraints of problem (8) hold, as well as a tangency condition, both $V(\mathbf{x})$ and $dV(\mathbf{x})/dt$, verify a transition of definiteness in an intermediate portion of the state space. Particularly $V(\mathbf{x})$ is positive definite in the small circle around the origin, then becomes negative definite in an intermediate region and recovers positive definiteness far from the origin. Similarly $dV(\mathbf{x})/dt$ is negative definite close the origin, becomes positive definite in an intermediate region and recovers negative definiteness far from the origin. While the small circle is an estimation of the DOA since it is fully contained in the region of negative definiteness of $dV(\mathbf{x})/dt$, the fact that the actual solution occurs beyond the transition of definiteness of the functions is clearly undesirable. Figs. 2 c) and d) are also solutions of problem (8) that present tangency between the level sets $V(\mathbf{x}) = c$ and $dV(\mathbf{x})/dt = 0$. However, none of them are actual DOAs since the level sets of $V(\mathbf{x})$ are not fully contained in the region of negative definiteness of $dV(\mathbf{x})/dt$. Since all solutions of Fig. 2 are valid solutions of problem (8) but none is the desired estimation of the actual DOA, they are considered dummy solutions of for system (9)/(10). From the above qualitative description it can be concluded that for a certain level set $V(\mathbf{x}) = c$ be an optimal estimation of the DOA, it should verify the following conditions:



a) Surfaces:
$V(\mathbf{x})$ (blue), $dV(\mathbf{x})/dt$ (red)

b) Level sets:
c = 2.2086, x_1 = -0.9079, x_2 = 0.9887

Fig. 1: Lyapunov function and time derivative for system (9)/(10)

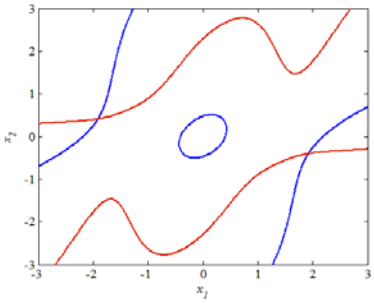
1. The level sets $V(\mathbf{x}) = c$ and $dV(\mathbf{x})/dt=0$ should verify tangential contact
2. The solution should belong to the portion of the state space around the origin, previous a transition of definiteness of $V(\mathbf{x})$ and $dV(\mathbf{x})/dt$ takes place
3. Provided that conditions 1 and 2 hold, the level set $V(\mathbf{x}) = c$ should be the global minimum

Condition 1 helps in avoiding dummy solutions such as the one in Fig. 2a). Condition 2 ensures that solutions of the type of Fig. 1b) are preferred over a solution like the one of Fig. 2b). Finally, the global minimum condition 3 favors the solution of Fig. 1b) over the tangential solutions of Figs. 2 c) and d). While problem (8) explicitly implies global optimality (condition 3), neither condition 1 nor condition 2, are considered. Therefore, dummy solutions such as those of Fig. 2 may result from problem (8). The proposed formulation is presented in Eq. (11) where ε is an auxiliary variable. While condition 1 can be straightforwardly imposed as a geometric constraint (11 b), condition 2 does not posses obvious mathematical expressions to be included within the optimization problem. In this contribution it is proposed to include a strict positive constraint on the denominator of function (5) to account for condition 2 (11 c). In Fig. 3 it is shown in gray the region determined by $D(\mathbf{x}) > 0$. Such condition contains the desired solution (Fig. 3 a) but excludes the dummy solution shown in Fig. 2b as can be appreciated from Fig. 3b).

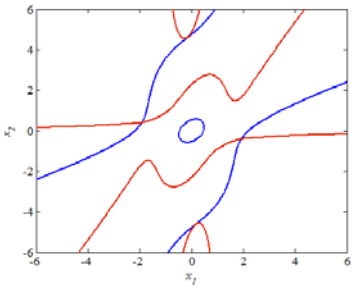
$$
\begin{aligned}
& c = \min_{\mathbf{x}} \; V(\mathbf{x}) \\
& s.t. \;\; \dot{V}(\mathbf{x}) = 0, \quad \mathbf{x} \neq \mathbf{0} && \text{a)} \\
& \nabla\left(V(\mathbf{x}) - c = 0\right) = \varepsilon \nabla\left(\dot{V}(\mathbf{x}) = 0\right) && \text{b)} \\
& D(\mathbf{x}) > 0 && \text{c)}
\end{aligned}
\tag{11}
$$

4. Application Example

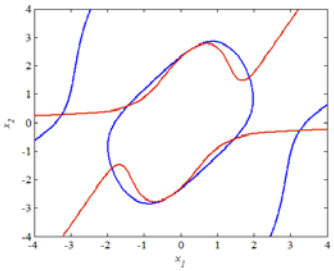
Formulation (11) is applied in this section to the determination of the DOA of a typical continuous stirred tank reactor. Consider the dynamic system of a typical CSTR described by mass and energy balances (12). Definitions and data are provided in Table 1. An estimation of the DOA of the steady state is calculated by solving problem (11) with Lyapunov function (13) ($x \equiv C_A$, $y \equiv T$).



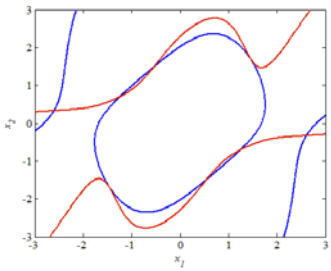
a) c = 0.1, x_1 = -1.9183, x_2 = 0.4052



b) c = 0.1287, x_1 = 0.1505, x_2 = -4.6180

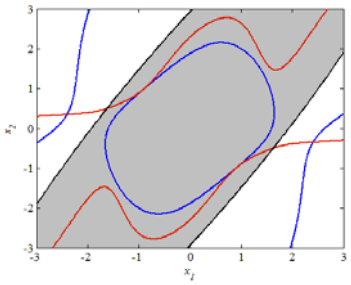


c) c = 4.7652, x_1 = -0.2962, x_2 = -2.5906

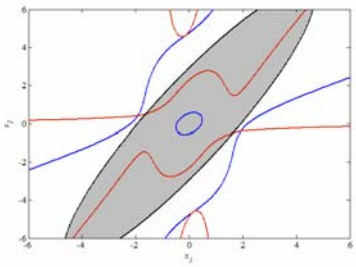


d) c = 2.8, x_1 = -1.4649, x_2 = -1.6832

Fig. 2: Dummy solutions for system (9)/(10)



a) Desired solution included



b) Dummy solution excluded

Fig. 3: Region D(**x**) > 0

$$\frac{dC_A}{dt} = \frac{F}{V}\left(C_{Af} - C_A\right) - k_0\, exp\left(\frac{-\Delta E}{RT}\right)C_A$$
$$\frac{dT}{dt} = \frac{F}{V}\left(T_f - T\right) + \left(\frac{-\Delta H}{\rho C_P}\right)k_0\, exp\left(\frac{-\Delta E}{RT}\right)C_A - \frac{UA}{V\rho C_P}\left(T - T_j\right) \quad (12)$$

$$N(x,y) = 325.4977x^2 + 71.4479xy + 4.2514y^2 + 69.4087x^3 + 22.9472x^2y + 2.1447xy^2 + 0.0404y^3 + 1.7325x^4 + 0.7648x^3y + 0.0591x^2y^2 0.0008xy^3 + 0.0005y^4$$
$$D(x,y) = 1 + 0.7341x + 0.0809y + 0.0877x^2 + 0.0195xy + 0.0007y^2 \quad (13)$$

Table 1: Nomenclature and data

Symbol	Description	Value
C_A	Concentration of A in reactor	2.3589
T	Reactor temperature	368.0629
k_0	Pre-exponential factor	9703*3600 h^{-1}
$-\Delta H$	Reaction heat	5960 kcal/kgmol
ΔE	Activation energy	11843 kcal/kgmol
ρCp	(Density)(heat capacity)	500 kcal/(m^3°C)
UA/V	(Hat transfer coefficient)(Area)/(Reactor Volume)	150 kcal/(m^3°C h)
R	Ideal gas constant	1.987 kcal/(kgmol K)
F/V	Residence time	1 h^{-1}
C_{Af}	Feed Concentration	10 kgmol/m^3
T_f	Feed Temperature	298 K
T_j	Jacket temperature	298 K

The problem was solved in the GAMS platform with the global optimization solver BARON (GAMS, 2008). The level sets of the Lyapunov function and its time derivative are shown in Fig. 4. The obtained estimation is the closed curve in blue around the equilibrium. $dV(\mathbf{x})/dt=0$ is shown in red and region $D(\mathbf{x}) > 0$ in gray.

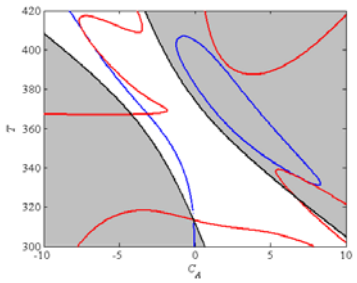


Fig. 4: Level sets $V(\mathbf{x}) = 639.5575$ and $dV(\mathbf{x})/dt = 0$ for system (12)/(13)

5. Conclusions

A global optimization formulation was proposed to estimate DOAs of stable equilibriums of systems of dynamic equations based on Lyapunov functions. In order to avoid potential dummy solutions, a tangency condition and a constraint on the denominator of $V(\mathbf{x})$ were added to previously published global optimization approaches. The rationale behind the proposed methodology was qualitatively described by means of a very rich example and also applied to a classic chemical engineering system.

References

GAMS (2008). A Users' Guide /The Solvers Manual.GAMS Development Corporation.

Hachicho, O. (2007). A novel LMI-based optimization algorithm for the guaranteed estimation of the domain of attraction using rational Lyapunov functions. Journal of the Franklin Institute. 344, pp. 535-552.

Vannelli, A. and M. Vidyasagar (1985). Maximal Lyapunov Functions and Domains of Attraction for Autonomous Nonlinear Systems. Automatica, 21, 69-80.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Probabilistic modelling and stochastic dynamic optimization for managing abnormal situations in plant-wide operations

Yu Yang, Jong Min Lee

Chemical and Materials Engineering, University of Alberta, Edmonton, Alberta, T6G 2G6, Canada

Abstract

This study proposes a novel dynamic real-time optimization framework that detects abnormalities at their onset and calculates a new operating condition to contain their further development. The proposed approach utilizes operational data and constructs a probability transition matrix that describes transition probabilities among representative operating modes under uncertainties. A quantitative measure of risk is defined to strike a balance between risk and profit in decision making. An integrated plant example consisting of a reactor, a storage tank, and a separator with a recycle loop is presented to demonstrate the efficacy of the approach.

Keywords: self organizing map, approximate dynamic programming, real-time optimization, plant-wide control.

1. Introduction

With an ever increasing need for improving process efficiency and product quality in today's complex manufacturing environment, real-time optimization (RTO) is considered an essential technique for the operation of complex chemical plants. Steady-state model-based RTO has been the most popular technique due to its simplicity (Cutler and Perry, 1983). However, steady-state RTO may leave significant room for performance improvement because it ignores the dynamic degrees of freedom such as unused capacity of a buffer tank and provides infeasible operating strategies during certain time periods. On the other hand, full-scale dynamic plant-wide models were used for rigorous online dynamic optimization based on a nonlinear programming method (Biegler et al., 2001). This approach may not be scalable to practical plant-wide optimization problems due to its excessive computational requirements.

One common difficulty associated with the model-based methods is obtaining a reliable plant-wide dynamic model in the form of algebraic/differential equations, apart from its relevance to online optimization. Another key issue of a model-based approach is that uncertainties from various sources can seriously degrade the performance. For example, model uncertainty can lead to RTO solutions that are far from the true plant optimum, or even an infeasible solution with respect to actual process constraints. Furthermore, once such solutions implemented and early symptoms go unnoticed, abnormal situations may arise and lead to significant periods of off-spec products or plant shutdowns.

Dynamic programming (DP) offers a unified approach to solving dynamic optimization problems under uncertainty (Bertsekas, 2005). The DP algorithm decomposes a

stochastic dynamic optimization problem into a sequence of single-period sub-problems that are solved recursively backward in time. Central to the methodology is the "value" function, which is obtained via solving Bellman's equation. Since DP algorithms compute and store a table consisting of one value per state, the size of a state space typically grows exponentially in the number of state variables, referred to as the curse of dimensionality.

Recently, the approach of approximate dynamic programming (ADP) has received much attention in the community of operations research (Si et al., 2004) and has been introduced to process systems engineering field (Lee and Lee, 2006). In most ADP methods, the value table is constructed only for a small subset of states and a function approximator is used to estimate values for unvisited states. The sample points are determined by running simulations with some known or random policies. The goal of this contribution is to develop a novel ADP-based RTO framework suitable for a large-scale plant under uncertainty.

2. Background

2.1. Markov Decision Processes and Dynamic Programming

Markov decision process (MDP) is a general modeling framework for multistage optimal control problems under uncertainty. At each of a series of discrete time steps, the state of a system (x_t) is observed and used to select an action (u_t), which then causes the system to change state and emit a numerical reward, $r(x_t, u_t)$. In general, the system dynamics is probabilistic for systems under uncertainty:

$$p_t(x_{t+1} | x_t, u_t) \tag{1}$$

which gives the probability of being in state x_{t+1} at time t+1 when the system is in state x_t and the decision u_t is implemented at time t.

The objective is to find a policy that maximizes the total reward received. The trade-off between immediate and delayed reward is handled by a discount factor $\gamma < 1$. Then the *value* of following a policy π from a state x is defined as the expectation of the sum of the subsequent rewards, each discounted geometrically by its delay:

$$V^{\pi}(x) = E\left\{ \sum_{t=0}^{\infty} \gamma^t r\left(x_t, u_t\right) \middle| x = x_0 \right\} \tag{2}$$

An optimal policy, π^*, achieves $V^{\pi^*}(s) \geq V^{\pi}(s)$ $\forall \pi, s$. DP involves iteratively solving the following Bellman equation:

$$V^*(x_t) = \max_{u_t} E\left\{ r(x_t, u_t) + \gamma V^*(x_{t+1}) \right\} \tag{3}$$

Once V^* is obtained, the optimal decision is easy to find through the following single-stage optimization

$$u_t^* = \arg\max_{u_t} E\left\{ r(x_t, u_t) + \gamma V^*(x_{t+1}) \right\} \tag{4}$$

2.2. Approximate Dynamic Programming

In ADP, the optimal (or *nearly optimal*) value function is obtained by iteratively improving the initial estimate of optimal value function based on a set of sampled state points visited by simulation or real-operation. One of the simplest methods in ADP is temporal difference learning of order 0 (TD(0)), which simply performs an update as follows:

$$\tilde{V}(x_t) \leftarrow r(x_t, u_t) + \gamma \tilde{V}(x_{t+1}) \tag{5}$$

where $\tilde{V}$ is approximation of the optimal value function.

Another advantage of ADP in solving the Bellman equation is that it can be extended to the case where transition probability or process model is unknown. This is possible by encoding the optimal value function with state and action pair as follows:

$$Q^*(x_t, u_t) = E\left\{ r(x_t, u_t) + \gamma \max_{u_{t+1}} Q^*(x_{t+1}, u_{t+1}) \right\} \tag{6}$$

The optimal value function and the optimal action-value function have the relationship of $V^*(x) = \max_u Q^*(x, u)$. The optimal decision given x_t is simply obtained by solving

$$u_t^* = \arg\max_{u_t} Q^*(x_t, u_t) \tag{7}$$

3. Proposed RTO Framework

3.1. Construction of Markov Chain Model

The first step in the proposed approach is to model underlying behavior of plant-wide dynamics under uncertainties as a series of "representative" discrete states evolving over time without any predetermined equations. In order to characterize such representative states, we cluster process data using self-organizing map (SOM) (Kohonen, 2001). SOM creates an ordered set of states in one or more dimensions. This offers topological structure of data, which means data points that are closer to each other in high dimensional space will be projected to the neighboring clusters (or nodes) in low dimensional space.

Once the historical data set is divided into representative nodes, the probability transition matrix P and one step reward are computed. It is assumed that each unit can take a finite number of set-point values. Under a particular action vector u (i.e., a particular set of set points), the transition probability from a discrete node i to j is estimated as

$$P_{ij}^u = \frac{N_{ij}^u}{\sum_{k=1}^{n_T} N_{ik}^u} \tag{8}$$

where N_{ij}^u is the number of switches from the states belonging to the node i to those belonging to j under action u and n_T is the total number of discrete nodes. The one step reward R_{ij}^u, the reward incurred during the transition from i to j under action u, is estimated as

$$R_{ij}^u = \frac{\sum_{x_{ij}} r(x_{ij}, u)}{M} \tag{9}$$

where x is the state point that made a transition from node i to j under action u and M is the total number of transitions.

A special node, F, is also designed to represent "failure states" such as shutdown, reaction termination, or constraint violation. This node behaves like an absorbing state,

meaning that the system stays at the same state with probability of one. A negative one step reward should be assigned to this node due to the loss of performance.

Once the representative states and their transition probabilities are obtained, the optimal value for each cluster is obtained by iteratively solving the Bellman equation:

$$V^*(i) = \max_u E\left\{R_{ij}(i,u) + \gamma V^*(j)\right\} \tag{10}$$

3.2. Model Refinement

A group of data points with distinctive values from the majority of data points inside each node can occur, which will degrade the performance of RTO. In order to identify such "outliers," Q value of each process state and action pair is first estimated using

$$Q(x_t, u_t) = r(x_t, u_t) + \gamma V^*(j) \tag{11}$$

where j is a representative node containing x_{t+1}. The outliers are defined as the data points with Q values falling outside $[Q_1\text{-}1.5(Q_3\text{-}Q_1), Q_3\text{+}1.5(Q_3\text{-}Q_1)]$ (Tukey, 1977). Q_1 and Q_3 denote the first and the third quartiles, respectively.

As the outliers are detected successively in each node, simple hyper-spheres are constructed to represent the area of outliers for fast classification in real-time applications. The center of each hyper-sphere is the current outlier, and the radius is the distance between that outlier data and its nearest "normal" data. Since the probabilistic model is not valid for outliers any more, a model-free approach is employed to estimate a new value for each outlier. Once the radius and centroid of hyper-sphere is defined, the data points belonging to the area are counted and their action-values $Q(x, u)$ are computed by the following model-free update scheme:

$$Q(x_t, u_t) \leftarrow Q(x_t, u_t) + \alpha\left\{r(x_t, u_t) + \gamma Q(x_{t+1}, u_{t+1}) - Q(x_t, u_t)\right\} \tag{12}$$

where α is a step-size parameter less than 1. The model refinement step can be performed both off-line (i.e., using the historical data only) and online (i.e., when the optimizer is put to work). The transition matrix, optimal value for each node, and action-values for outliers are updated online as a new observation is available at each time step.

3.3. Quantitative Risk Term

The proposed method designs a special node, F, to represent failure states and assigns a negative value V_F. The policy derived from the learned value and action-value functions with V_F will strike a balance between total reward and the risk of entering undesirable states. In this work, V_F is assigned by

$$V_F = \frac{(D+1)\underline{r} - D\overline{r}}{\gamma(1-\gamma)} \tag{13}$$

where D is a user-specified parameter, $\underline{r}$ and $\overline{r}$ are lower and upper bounds of single-stage reward, respectively. The detailed discussion on how to derive Eq. (12) is provided in Yang and Lee (2009).

4. Case Study

An integrated process studied in Tosukhowong et al. (2004) is considered to illustrate the efficacy of the proposed framework. The process is composed of a CSTR, a storage tank, and a flash tank with a material recycle stream as shown in Fig. 1. A fresh feed

stream F_0 consisting of pure component 1 is fed to the reactor, where two irreversible reactions take place to produce the desired product 2 and undesired product 3. Detailed model descriptions can be found in Tosukhowong et al. (2004).

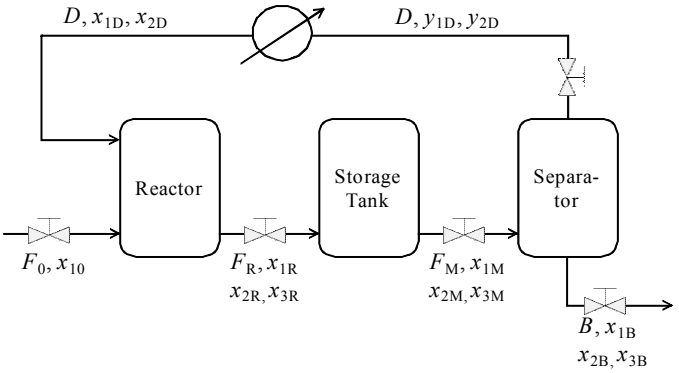


Figure 1. Schematic of the reaction-storage-separation network.

Three proportional-only (P) controllers are designed to control liquid level of each unit, and inappropriate assignment of a set point trajectory to each unit may lead to instability or significantly reduced performance under uncertainty in the inlet flow and kinetic parameters. Use of more sophisticated local controllers than P-controller may enhance the plant-wide performance. However, the question here is how to improve the plant-wide performance given the existing local controllers by coordinating operating conditions of each unit.

The objective of plant-wide optimization is to maximize the production of desired product formulated as

$$\max \sum_{t=1}^{\infty} 0.9^{t-1} B(t) x_{2B}(t) \tag{14}$$

The time interval of dynamic optimization is set as 40 minutes as suggested in Tosukhowong et al. (2004). In each episode of simulation, a sequence of random set points was used for each unit for six hours. Every 100-200 minutes, F_0 and the kinetic parameters were alternating between upper and lower bounds ($1.45 \leq F_0 \leq 1.85$, $0.01 \leq k \leq 0.03$) with their nominal values 1.667 and 0.0167, respectively.

From the data, the SOM structure with 7×3 was selected considering the number of outliers. With the 21 clusters, the single-stage reward and the transition matrices were calculated. V_F was estimated as -1478 with γ and D set as 0.9 and 4, respectively.

For comparison with steady state RTO and dynamic RTO proposed in Tosukhowong et al. (2004), 300 episodes of simulation scenarios different from the training ones were executed. As inappropriate set points will lead to total drainage of storage tank and plant shutdown, the ratio of failing optimization are compared in Fig. 2. Fig. 3 shows one of the sample realizations in the test. The production of x_{1B} is interrupted (a negative value) at 250th sample time under the dynamic RTO.

5. Conclusions

The proposed approach is based on a Markov chain model and approximate dynamic programming and provides a more robust and risk-averse policy compared to other

deterministic model-based approaches. This method can be applied to any dynamic optimization problems when a process model is difficult to obtain, which is likely for large-scale problems, and uncertainties affect operations significantly.

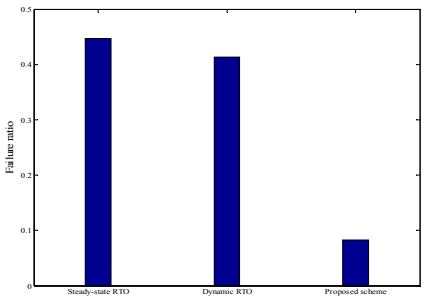


Figure 2. Ratio of operation leading to total drainage of the storage tank.

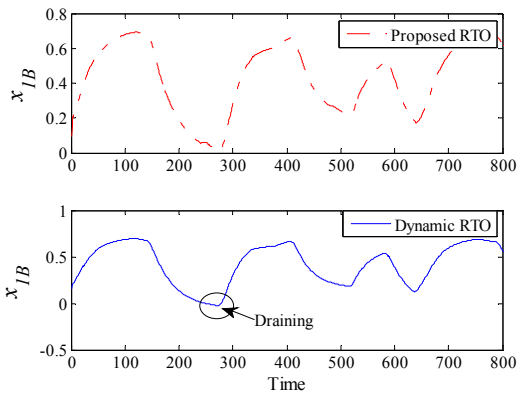


Figure 3. Constraint violation (total drainage) based on the dynamic RTO.

References

D.P. Bertsekas, 2005, Dynamic Programming and Optimal Control, 3rd Ed., Vol. 1, Athena Scientific.

C.R. Cutler, Perry, R.T., 1983, Real time optimization with multivariable control is required to maximize profits. Computers and Chemical Engineering, 7, 663-667.

T. Kohonen, 2001, Self-Organizing Maps, Springer.

J.H. Lee, J. M. Lee, 2006, Approximate dynamic programming based approach to process control and scheduling. Computers and Chemical Engineering, 30, 1603-1618.

J. Si, A.G. Barto, W.B. Powell, 2004, Handbook of Learning and Approximate Dynamic Programming, Wiley-IEEE.

T. Tosukhowong, J.M. Lee, J.H. Lee, J. Lu, 2004, An introduction to a dynamic plant-wide optimization strategy for an integrated plant. Computers and Chemical Engineering, 29, 199-208.

J.W. Tukey, 1977, Exploratory Data Analysis, Addison-Wesley.

Y. Yang, J.M. Lee, 2009, Probabilistic modeling and dynamic optimization for performance improvement and risk management of plant-wide operation, Computers and Chemical Engineering, submitted.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Multi-scenario-based robust nonlinear model predictive control with first principle models

Rui Huang[a], Sachin C. Patwardhan[b], Lorenz T. Biegler[a]

[a] *Department of Chemical Engineering, CMU, Pittsburgh, PA, USA 15213.*
[b] *Department of Chemical Engineering, IIT Bombay, Mumbai, India, 400076.*

Abstract

A robust nonlinear model predictive control (NMPC) algorithm is developed based on a multi-scenario formulation. The uncertainties are characterized by different scenarios so that the calculated control action is feasible over the entire uncertainty region. We show that this multi-scenario formulation is Input-to-State practically Stable (ISpS), and can be easily extended to the recently proposed advanced-step NMPC (as-NMPC), which is able to reduce the online computational delay. We demonstrate the advantages of this strategy on a large-scale air separation unit.

Keywords: NMPC, Robust strategy, large-scale applications

1. Introduction

Nonlinear model predictive control (NMPC) has recently received a great amount of attention, especially for energy and chemical processes that require dynamic real-time optimization (D-RTO). Moreover, a recent vehicle for D-RTO deals with advanced-step Nonlinear Model Predictive Control (as-NMPC) (Zavala & Biegler, 2008), which incorporates very large dynamic optimization problems and can reduce the on-line feedback delay to only a few CPU seconds, even with industrial-size first principle dynamic models.

On the other hand, ensuring stability of closed-loop systems with faults and model mismatch remains an important problem in NMPC. While nominal stability and feasibility have been extensively addressed in the past decades (Magni & Scattolini, 2007, Limon et al, 2008), NMPC with asymptotic stability does not guarantee robust stability and large unmeasured disturbances can lead to poor performance for D-RTO. In these cases, robust design strategies are necessary to account for uncertainties explicitly in the controller formulation. A well known strategy to guarantee the robust stability is the min-max NMPC formulation, which computes the best control policy based on the worst expected realization of the uncertainties. However this formulation dramatically increases the computational cost of the on-line NMPC problem, despite some recent remedies (Diehl et al, 2008, Fontes & Magni, 2003). Moreover, since the control action is based on the worst case, the performance is compromised, e.g. output variables present large offset.

In this work, we propose a robust NMPC design strategy based on multi-scenario NLP formulation and extend it to as-NMPC. The robust stability can be established through the recent input-to-state practical stability (ISpS) framework (Limon et al, 2008). Section 2 presents the multi-scenario formulation for NMPC and as-NMPC and analyze their robust stabilities; section 3 shows that the proposed method is suitable for real world systems by considering a large-scale air separation unit with plant-model mismatches; section 4 concludes the paper.

2. NMPC based on Multi-scenario formulation and Stability Analysis

2.1. Multi-scenario Formulation

In this work, the dynamics of an uncertain plant will be described by the following discrete-time model,

$$x(k+1) = f(x(k), u(k), \phi(k)),\ k \geq 0 \tag{1}$$

where $x(k) \in \mathfrak{R}^{n_x}$ is a system state vector, $u(k) \in \mathfrak{R}^{n_u}$ is a control variable vector and $\phi(k) \in \mathfrak{R}^{n_\phi}$ is an uncertainty signal which models disturbances and plant-model mismatches at time steps $k > 0$. For the sake of clarity, we only consider the state-independent uncertainty signals in this work, i.e. $\phi(k)$ does not depend on $x(k)$ and $u(k)$. However with slight modification of notations, the analysis here can be extended to the case with state-dependent uncertainties. In general, the control and state of the plant are required to fulfill the constraints, e.g. $x(k) \in \mathbb{X},\ u(k) \in \mathbb{U}$, and only the partial information of the uncertainty $\phi(k) \in \mathbf{\Omega}$ is available, e.g. its feasible region Ω. Without losing generality, we assume that the given plant (1) has an equilibrium point at the origin, i.e. $f(0,0,0) = 0$.

Given $x(k)$, the current state value at time step t_k, the NMPC formulation can be described in the following discretized form:

$$\min_{z(j),v(j)}\ J(x(k)) := F(z(N)) + \sum_{j=0}^{N-1} \varphi(z(j), v(j), \theta(j)) \tag{2a}$$

$$s.t.\ z(j+1) = f(z(j), v(j), \theta(j)),\ z(0) = x(k) \tag{2b}$$

$$z(j) \in \mathbb{X},\ z(N) \in \mathbb{X}_f,\ v(j) \in \mathbb{U},\ \theta(j) \in \mathbf{\Omega},\ j = 0, \ldots N-1 \tag{2c}$$

where N is the finite time horizon, $\theta \in \Omega$ is the uncertainty parameter in the controller. The computed control $v(j) \in \mathfrak{R}^{n_u}$ and predicted state $z(j) \in \mathfrak{R}^{n_x}$ are enforced to satisfy the constraints $v(j) \in \mathbb{U}$, $z(j) \in \mathbb{X}$, $z(N) \in \mathbb{X}_f$. The cost function $J(x(k))$ at t_k comprises stage cost $\varphi(\cdot,\cdot,\cdot)$ and terminal penalty cost $F(\cdot)$. When the solution sequence $(z^*(l), v^*(l))$ is available, the first control $u(k) = v^*(0)$ is injected into the plant. In general, $x(k+1)$ is different from the predicted value because of the plant-model mismatches, (e.g.$\theta \neq \phi$). At the next time step t_{k+1} when $x(k+1)$ is measured or estimated, we move the time horizon one step further and solve problem (2) again.

If the presence of uncertainties does not cause any loss of feasibility (e.g. no constraints), the NMPC (2) enjoys inherent robustness (Magni & Scattolini, 2007). However if the constraints are enforced and uncertainties are present, the calculated control v from (2) may be infeasible for the plant (1). To avoid this difficulty, we solve a multi-scenario problem:

$$\min_{z_l(j),v(j)}\ V(x(k)) := \sum_{l=1}^{M} w_l J_l(x(k))$$

$$= \sum_{l=1}^{M} w_l \left\{ F_l(z_l(N)) + \sum_{j=0}^{N-1} \varphi(z_l(j), v(j), \theta_l) \right\} \tag{3a}$$

$$s.t.\ z_l(j+1) = f(z_l(j), v(j), \theta_l),\ z_l(0) = x(k),\ l = 1, \ldots M \tag{3b}$$

$$z_l(j) \in \mathbb{X},\ z_l(N) \in \mathbb{X}_f,\ v(j) \in \mathbb{U},\ \theta_l \in \mathbf{\Omega},\ j = 0, \ldots N-1, \tag{3c}$$

with M different uncertainties $\theta_l, l \in 1, \ldots, M$, where l is the index of scenarios and w_l are weights for each scenario, satisfying $0 \leq w_l \leq 1,\ \sum_{l=1}^{M} w_l = 1$. In addition, consider $B_{\delta l}$ as M closed balls, $B_{\delta l} \triangleq \{\vartheta_l \in \Omega | |\vartheta_l - \theta_l| \leq \delta_l\},\ \forall l \in 1 \ldots M$, centered around q_l with radius δ_l. The balls $B_{\delta l}$ are defined such that an NMPC formulated with

θ_l as nominal model parameter is robustly stable within this ball. We will prove that the control sequence $v(\cdot)$ calculated from (3) remains feasible and stable for all the uncertainty regions $B_{\delta_l}, \forall l$. If $\phi(j) \in \bigcup_{l=1}^{M} B_{\delta l}, \forall j$, then the closed-loop system is robustly stable.

2.2. Robust Stability Analysis

For the analysis of the robust stability of multi-scenario NMPC, we make use of the regional input-to-state practical stability (ISpS) property developed in Limon et al, 2008. Here, we only summarize key definitions and results.

Definition 1. *Given a set $\Gamma \subseteq \Re^{n_x}$, including the origin as an interior point, the system (1) is said to be ISpS in Γ w.r.t ϕ, if Γ is a robust positive invariant set for (1) and there exist a $\mathcal{KL}$ function β, a $\mathcal{K}$ function γ and a constant $c \geq 0$ such that,*

$$|x(k)| \leq \beta(|x_0|, k) + \gamma(|\phi_{(k-1)}|) + c, k \geq 0, \ \forall x_0 \in \Gamma, \ \phi(k) \in \mathbf{\Omega} \tag{4}$$

with $|\phi_{(k-1)}| \triangleq sup_{k\geq 0}\{|\phi(k-1)|\}$. Moreover, a function $V(\cdot)$ is an ISpS-Lyapunov function for system (1) in Γ if there exist $\mathcal{K}$ functions α_1, α_2, α_3, σ and constants $c_1, c_2 \geq 0$ such that:

$$\alpha_2(|x|) \leq V(x) \leq \alpha_2(|x|) + c_1, \ \forall x \in \Gamma$$
$$\triangle V(x, \phi) = V(f(x, u, \phi)) - V(x) \leq -\alpha_3(|x|) + \sigma(|\phi|) + c_2, \forall x \in \Gamma, \ \forall \phi \in \mathbf{\Omega} \tag{5}$$

Lemma 1. *If (1) admits an ISpS-Lyapunov function in Γ, then it is ISpS in Γ w.r.t ϕ.*

In order to deal with the multi-scenario formulation, the common robust stability assumptions (Magni & Scattolini, 2007, Zavala & Biegler 2008, Limon et al, 2008) are modified as follows.

Assumption 1 *(Assumptions for Multi-scenario Formulation and Robust Stability)*

a) There exists a constant $c_3 \geq 0$, that the terminal penalty $F_l(\cdot)$ satisfy $\alpha_F(|x(k)|) \leq F_l(x(k)) \leq \gamma_F(|x(k)|) + c_3, \ \forall l, \ \forall x \in \mathbb{X}_f \subseteq \Gamma, \ \alpha_F(\cdot)$ and $\gamma_F(\cdot)$ are $\mathcal{K}$ functions.

b) There exists a local control law $h_f(z_l(N))$ defined on $\mathbb{X}_f \subseteq \Gamma$, such that $f(z_l(N), h_f(z_l(N)), \theta_l) \in \mathbb{X}_f, \forall z_l(N) \in \mathbb{X}_f$, and $F_l(f(z_l(N), h_f(z_l(N)), \theta_l)) - F_l(z_l(N)) \leq -\varphi(z_l(N), h_f(z_l(N)), \theta_l), \ \forall l$.

c) $\alpha_p(|x(k)|) \leq \varphi(x(k), u(k), \theta(k)) \leq \alpha_q(|x(k)|) + c_4, \forall \theta(k) \in \mathbf{\Omega}$, for some constant $c_4 \geq 0$ with $\alpha_p(\cdot)$, $\alpha_q(\cdot)$ as $\mathcal{K}$ functions.

d) Given $u(k)$ and $z_l(k+1) = f(x(k), u(k), \theta_l)$ for each scenario l, there exist future mismatches at the next sampling time:

$$\epsilon_l(x(k+1)) := J_l(x(k+1)) - J_l(z_l(k+1)) \ \forall l \tag{6}$$

with $\mathcal{K}$ functions α_4, σ_2 and positive Lipschitz constants L_ϵ such that $|\epsilon_l(x(k+1))| \leq L_\epsilon(\alpha_4(|x(k)|) + \sigma_2(|\theta_l - \phi(k)|))$.

e) For a constant $c_5 \geq 0$, $\mathcal{K}$ functions α_5 and $\theta_l \in \Omega, \ \forall l$, there exists a constant $K \geq 0$ such that:

$$\sum_{l=1}^{M} w_l \Big[-\varphi(x(k), u(k), \theta_l) + K(\alpha_4(|x(k)|) + \sigma_2(|\theta_l - \phi(k)|)) \Big]$$
$$\leq -\alpha_5(|x(k)|) + \sigma_1(|\phi(k)|) + c_5 \tag{7}$$

Robust stability of multi-scenario NMPC can now be established by the following theorem.

Theorem 1 *(Robust ISpS stability of Multi-scenario NMPC) If the plant uncertainty parameter vector lies in the union of balls centered around the uncertainty parameters in the controller, i.e.* $\phi(j) \in \bigcup_{l=1}^{M} B_{\delta l}, \forall j$*, then under Assumption 1, with* $K \geq L_\epsilon$*, the cost function* $V(x(k)) = \sum_{l=1}^{M} w_l \big[F_l(z_l(N)) + \sum_{j=0}^{N-1} \varphi(z_l(j), v(j), \theta_l)\big]$ *is an ISpS-Lyapunov function for plant (1) in* Γ *and the resulting closed loop system is ISpS stable.*

Proof. First let $u(k)$ be the first element of the control sequence calculated from (3) based on $x(k)$ at t_k, There exists a $\theta_{min}(k) \in \mathbf{\Omega}$ so that $\varphi(x(k), u(k), \theta_l) \geq \varphi(x(k), u(k), \theta_{min}(k)), \forall \theta_l \in \mathbf{\Omega}$ for the fixed $x(k)$ and $u(k)$, then

$$\begin{aligned} V(x(k)) &= \sum_{l=1}^{M} w_l \left\{ F_l(z_l(N)) + \sum_{j=0}^{N-1} \varphi(z_l(j), v(j), \theta_l) \right\} \\ &\geq \sum_{l=1}^{M} w_l \sum_{j=0}^{N-1} \varphi(z_l(j), v(j), \theta_l) \\ &\geq \sum_{l=1}^{M} w_l \varphi(x(k), u(k), \theta_{min}(k)) \geq \alpha_p(|x(k)|) \end{aligned} \tag{8}$$

the last two inequalities follow from letting $j = 0$ and Assumption 1.c).

Then we consider the multi-scenario NMPC with horizon N+1, its objective function is

$$V(x(k), N+1) = V(x(k), N) + \sum_{l=1}^{M} w_l \big[- F_l(z_l(N)) +$$

$$F_l(f(z_l(N), h_f(z_l(N)), \theta_l)) + \varphi(z_l(N), h_f(z_l(N)), \theta_l)\big] \leq V(x(k), N) \tag{9}$$

the inequality is from Assumption 1.b). So that

$$\begin{aligned} V(x(k), N+1) &\leq V(x(k), N) \leq \ldots \leq V(x(k), 0) \\ &= \sum_{l=1}^{M} w_l F_l(x(k)) \leq \gamma_F(|x(k)|) + c_3 \end{aligned} \tag{10}$$

the last inequality is from Assumption 1.a).

Finally, we are able to compare the cost function $V(x(k+1))$ and $V(x(k))$ from the two neighboring problems with horizon length N,

$$\begin{aligned} &V(x(k+1)) - V(x(k)) \\ &= \sum_{l=1}^{M} [w_l J_l(z_l(k+1))] - V(x(k)) + V(x(k+1)) - \sum_{l=1}^{M} [w_l J_l(z_l(k+1))] \\ &= \sum_{l=1}^{M} w_l [J_l(z_l(k+1)) - J_l(x(k))] + \sum_{l=1}^{M} w_l \epsilon_l(x(k+1)) \end{aligned} \tag{11}$$

From Assumption 1.b), we know $J_l(z_l(k+1)) - J_l(x(k)) = F_l(f(z_l(N), h_f(z_l(N)), \theta_l)) - F_l(z_l(N)) + \varphi(z_l(N), h_f(z_l(N)), \theta_l) - \varphi(x(k), u(k), \theta_l) \leq -\varphi(x(k), u(k), \theta_l), \forall l$. Consequently

$$\begin{aligned} &V(x(k+1)) - V(x(k)) \leq \sum_{l=1}^{M} w_l [-\varphi(x(k), u(k), \theta_l) + \epsilon_l(x(k+1))] \\ &\leq \sum_{l=1}^{M} w_l \Big[- \varphi(x(k), u(k), \theta_l) + L_\epsilon(\alpha_4(|x(k)|) + \sigma_2(|\theta_l - \phi(k)|)) \Big] \end{aligned} \tag{12}$$

where (12) follows from Assumption 1.d). With $K \geq L_\epsilon$ and Assumption 1.e),

$$V(x(k+1)) - V(x(k)) \leq -\alpha_5(|x(k)|) + \sigma_1(|\phi(k)|) + c_5 \tag{13}$$

The desired result follows as (8), (10) and (13) allow *x(k)* and *V* to satisfy (4) and (5). ?

Note from the analysis, the robust stability is guaranteed when $\phi(j) \in \bigcup_{l=1}^{M} B_{\delta l}, \forall j$, we also see that the it will improve if the majority of the uncertainty parameters in the controller are close to that in the plant. In practice, since the plant uncertainty parameter is unknown, the multi-scenario NMPC can be initiated with a few guessed parameters, if the controller appears to be unstable, more scenarios should be added to make sure that the plant uncertainty ϕ is in the neighborhood of one uncertainty parameter θ_l, hence the robust stability is achieved.

2.3. Multi-scenario for as-NMPC

Although the multi-scenario formulation can improve robust stability, it may introduce large computational delay online. In order to reduce the online computational delay, we modify the algorithm of the advanced-step NMPC (Zavala & Biegler, 2008) to incorporate the multi-scenario formulation, as follows:

1. Background calculation: at t_k, predict the future state $z_l(k+1)$ with $x(k)$ and $u(k)$ for each scenario *l*. Solve problem (3) with $z_l(k+1)$ as the initial condition for each scenario *l*. The NLP solver takes a Newton step, which can be expressed as:

$$K(z_l(k+1))\triangle s = -\varphi(z_l(k+1)) \tag{14}$$

where s is the solution vector and Δs represents the Newton step from the current solution towards the new point, $K(\cdot)$ and $\varphi(\cdot)$ are the KKT matrix and KKT condition at current solution, respectively.

2. Online update: at t_{k+1}, obtain $x(k+1)$ and use (14) to get the updated solution: $K(z_l(k+1))\triangle s = -\varphi(x(k+1))$. Extract control action $u(k+1)$ and inject into plant.

3. Iterate: set $k \leftarrow k+1$, and go to step 1.

Note that the KKT matrix in (14) has already been factorized in solving (14), so only a single back-solve is required in Step 2; this is at least two orders of magnitude faster than solving the optimization problem (3). Along the same lines, we can derive a similar but weaker ISpS property for multi-scenario as-NMPC, where the online update leads to only a small deterioration of the stability property.

3. Simulation examples

We now revisit the large-scale air separation unit reported in Huang et al, 2009 with plant-model mismatches. In this work, we modify the ASU model to introduce a tray efficiency, which is considered as an uncertain parameter. In this simulation, we assume that the true value of tray efficiency is $\alpha = 0.8$ and the ASU starts from a nominal steady-state. The oxygen (POX) and nitrogen production (PNI) rate set-points are reduced by *30%* through a ramp change from $t = 30$ to $t = 60$ minutes, while temperatures (Tl30 & Th15) are required to remain constant to control the product purities. After this, they undergo a ramp increase from $t = 1000$ to $t = 1030$ back to their original values. The NMPC formulation with $\alpha = 0.9$ or $\alpha = 1$, and multi-scenario NMPC with two scenarios $\alpha = 0.9$ and $\alpha = 1$ yield infeasible control actions. Hence, we choose equally weighted three scenarios with $\alpha_1 = 0.8$, $\alpha_2 = 0.9$ and $\alpha_3 = 1$ in the multi-scenario as-NMPC. Figure 1 shows that it is able to track the output variables with stable control actions and product purities are satisfied. Moreover multi-scenario as-NMPC yields nearly the same performance as the ideal multi-scenario NMPC. Without plotting results, the same NMPC formulation is robustly stable for cases where tray efficiency equals to 0.85, 0.95 and 1. After discretization, the optimization problem with three scenarios leads to a NLP contains 350,940 variables and 350,700 constraints. The simulation was done on an Intel Quad Core @ 2.4GHz server running Linux. For each horizon, IPOPT took 20-30 CPU minutes for the background calculation, while

sensitivity update took around 2 CPU seconds. Note online computation time is reduced by at least 600 times with virtually no loss in performance.

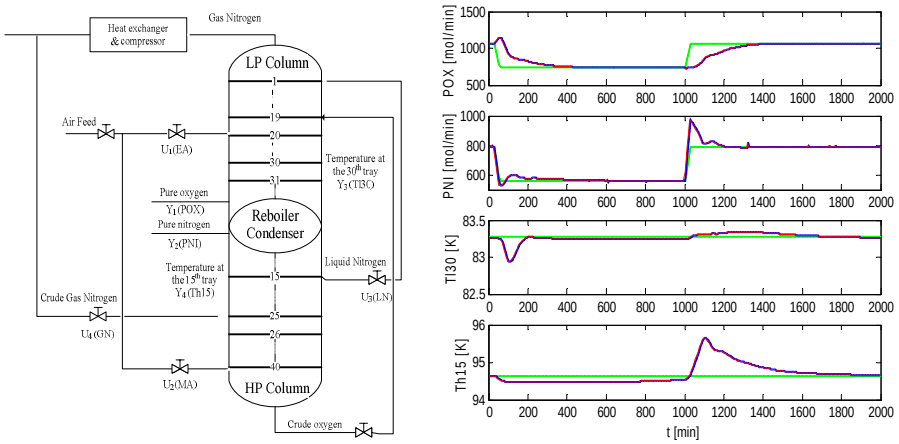


Figure 1. ASU flowsheet and output variable profiles. Dash-dotted- lines are setpoints, superimposed dashed and solid lines are multi-scenario NMPC and as-NMPC, respectively.

4. Concluding Remarks

In this study a robust NMPC strategy based on multi-scenario formulation is presented. This strategy solves multiple problems with different uncertainty parameters to guarantee the satisfaction of constraints. We establish the input-to-state practical stability of this formulation for both NMPC and as-NMPC. The advantages of the proposed method are demonstrated by simulation of a large-scale ASU. In addition the multi-scenario as-NMPC presents nearly the same performance as the NMPC counterpart, but dramatically reduces the online computation delay. For the future work, we propose to study the strategies that allow the background solution taking more than one sampling time while still providing instantaneous on-line update. Besides parallel computing with distributed computer cluster will be studied to future reduce the background computational effort. Moreover we will explore methods to reduce the number of scenarios needed to guarantee the robust stability, e.g. back-off constraints.

References

M. Diehl, J. Gerhard, W. Marquardt, M. Mönnigmann, 2008, “Numerical solution approaches for robust nonlinear optimal control problems”, Computers & Chemical Engineering, 32: 1279.

F. Fontes & L. Magni, 2003, “Min-max model predictive control of nonlienar systems using discontinuous feedbacks”, IEEE Transactions on Automatic Control, 48: 1750.

R. Huang, V. Zavala & L. Biegler, 2009, “Advanced step noninear model predictive control for air separation units”, J. of Process Control, 19: 678

D. Limon, T. Alamo, D. Raimondo, D.M. Peña, J. Bravo, E. Camacho, 2008. “Input-to-state stability: an unifying framework for robust model predictive control”, In Assessment and Future Directions of Nonlinear Model Predictive Control, Springer, in press

L. Magni & R. Scattolini, 2007, “Robustness and robust design of mpc for nonlinear discrete-time systems”, Assessment and Future Directions of Nonlinear Model Predictive Control, 239-254.

V. Zavala & L. Biegler, 2008, “The advanced step nmpc controller: Optimality, stability and robustness”, Automatica, 45: 86.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Approximation of Arrival Cost in Moving Horizon Estimation Using a Constrained Particle Filter

Rodrigo López-Negrete[a], Sachin C. Patwardhan[b], Lorenz T. Biegler[a]

[a] *Department of Chemical Engineering, CMU, Pittsburgh, PA, USA 15213.*
[b] *Department of Chemical Engineering, IIT Bombay, Mumbai, India, 400076.*

Abstract

Moving horizon estimation (MHE) is increasingly being used for state estimation in model predictive control schemes. Approximation of the arrival cost in MHE formulation is an open issue in this domain. In this work, we propose to use a constrained particle filter for estimating the moments of the prior probability density function of the initial state and approximating the arrival cost. The efficacy of the proposed approach is illustrated in a simulation case study on a benchmark problem.

Keywords: Moving Horizon Estimation, Arrival Cost, Constrained Ensemble Kalman Filter

1. Introduction

Moving horizon estimation is a preferred approach for state estimation particularly when used in connection with model predictive control schemes (Zavala and Biegler, 2008). The main advantage of MHE is that it can handle bounds on state and parameter estimates or irregular measurements in a systematic manner. Moreover, it provides a unifying framework for state estimation in systems described by differential algebraic equations or hybrid systems. Recent focus of research in this area has been on development of computationally efficient algorithms for on-line implementation (Zavala and Biegler, 2008). An important aspect of MHE formulation is estimation of the arrival cost. In a recent review paper, Bakshi and Rawlings (2006) have indicated that best choice of the arrival cost remains an open issue in the domain of MHE research. As the true arrival cost is difficult to determine, it is often estimated using linearization approximation as in the extended Kalman filtering (EKF) formulation. This method of approximating the arrival cost, however, does not consider the bounds on state estimates in a systematic manner.

Recently, a new class of recursive Bayesian filtering technique, called particle filtering, has attracted attention of many researchers (Rawlings and Bakshi, 2006). The objective of the recursive Bayesian state estimation problem is to find the mean and variance of the random state variable using its probability density function conditioned on the past measurements. While dealing with nonlinear systems, the exact analytical solution to the recursive propagation of the posterior density is difficult to obtain. Thus, it becomes necessary to develop approximate and computationally tractable sub-optimal solutions to the sequential Bayesian estimation problem. The particle filter is a numerical method for implementing an optimal recursive Bayesian filter through Monte-Carlo simulation. Thus, even when the state estimation error has multimodal and non-Gaussian distributions, a particle filter formulation can accurately estimate its moments. Ensemble Kalman filter (EnKF) developed by Evenson (2003) is a form of particle filter that retains the computational simplicity of EKF without requiring linearization or

normality of the conditional density function. However, as proposed, EnKF formulation cannot deal with bounds on states or parameters. Recently, Prakash et al. (2007) have developed constrained version of ensemble Kalman filter (C-EnkF), which provides a systematic and statistically consistent approach to deal with bounds on the state estimates. In the present work, we exploit this feature of C-EnKF for approximating the arrival cost in the MHE formulation. The efficacy of the proposed approach is illustrated by conducting simulation studies on a benchmark problem in the state estimation literature.

2. Review of MHE and C-EnKF

2.1. Preliminaries

Consider a nonlinear system represented by the following nonlinear state space equations:

$$\mathbf{x}(k) = F\left[\mathbf{x}(k-1), \mathbf{u}(k-1)\right] + \mathbf{w}(k-1) \tag{1}$$

$$\mathbf{y}(k) = H\left[\mathbf{x}(k)\right] + \mathbf{v}(k) \tag{2}$$

In the above process model, $\mathbf{x}(k) \in R^n$ is the system state vector, $\mathbf{u}(k) \in R^m$ is known system input, $\mathbf{w}(k) \in R^p$ is the state noise with known distribution, $\mathbf{y}(k) \in R^r$ is the measured state variable and $\mathbf{v}(k) \in R^r$ is the measurement noise with known distribution. The index k represents the sampling time. The operators $F[.]$ and $H[.]$ represent the nonlinear process model and nonlinear measurement model, respectively. It is further assumed that the initial state of the system $\mathbf{x}(0)$ is a random vector with known probability distribution.

2.2. Moving Horizon Estimation

In most physical systems, states/parameters are bounded, which introduces constraints on state/parameter estimates. One major limitation of recursive nonlinear filters is that these formulations cannot handle such constraints. Moving horizon estimation (MHE) (Bakshi and Rawlings, 2006) formulation provides a systematic approach to handling of bounds on states/parameters or any other algebraic constraints. The MHE problem can be stated as (Bakshi and Rawlings, 2006)

$$\underset{\mathbf{x}(k-N:k)}{Min} \; \Pi\left[\mathbf{x}(k-N)\right] + \sum_{i=k-N+1}^{k} \left\{\log\left[p\left(\mathbf{w}(i-1)\right)\right] + \log\left[p\left(\mathbf{v}(i)\right)\right]\right\} \tag{3}$$

$$\text{subject to model equations (1) and (2) and} \quad \mathbf{x}_L \le \mathbf{x}(i) \le \mathbf{x}_U \tag{4}$$

where, $p(.)$ denotes the probability density function of the argument. The term $\Pi\left[\mathbf{x}(k-N)\right]$, called the *arrival cost*, represents the negative log of the conditional density $p\left[\mathbf{x}(k-N) \mid \mathbf{Y}_{k-N}\right]$ where $\mathbf{Y}_{k-N} = \left\{\mathbf{y}(1), ...\mathbf{y}(k-N)\right\}$. The arrival cost represents the information in the prior measurement sequence $\mathbf{Y}_{k-N}$ that is not considered in the horizon at time k (Bakshi and Rawlings, 2006). The main difficulty with the arrival cost term is that, except for the linear unconstrained case, the conditional density $p\left[\mathbf{x}(k-N) \mid \mathbf{Y}_{k-N}\right]$ cannot be easily characterized using closed form algebraic expressions.

In most practical situations, the probability density functions of the state noise and the measurement noise can be well approximated using Gaussian density functions. In addition, if we further assume that the *a-priori* estimate is normally distributed

$\mathbf{N}(\hat{\mathbf{x}}(k-N), \mathbf{P}(k-N))$ such that covariance $\mathbf{P}(k-N)$ is invertible, then the MHE objective function can be reformulated as follows

$$\underset{\Omega_k}{Min}\ \Phi\left[\mathbf{x}(k-N)\right] + \sum_{i=k-N+1}^{k}\left[\mathbf{w}(i-1)^T Q^{-1}\mathbf{w}(i-1) + \mathbf{v}(i)^T \mathbf{R}^{-1}\mathbf{v}(i)\right] \quad (5)$$

$$\Omega_k \equiv \left\{\mathbf{x}(k-N), \mathbf{w}(n-N+1), \ldots \mathbf{w}(k-1)\right\}$$

$$\Phi\left[\mathbf{x}(k-N)\right] = \boldsymbol{\varepsilon}(k-N)^T \mathbf{P}(k-N)^{-1}\boldsymbol{\varepsilon}(k-N) + \Phi^*\left[\mathbf{x}(k-N)\right] \quad (6)$$

$$\boldsymbol{\varepsilon}(k-N) = \mathbf{x}(k-N) - \hat{\mathbf{x}}(k-N)$$

where $\mathbf{w}(k) \sim \mathbf{N}(\bar{\mathbf{0}}, \mathbf{Q})$, $\mathbf{v}(k) \sim \mathbf{N}(\bar{\mathbf{0}}, \mathbf{R})$ and $\Phi^*\left[\mathbf{x}(k-N)\right]$ represents the optimal cost at time *(k - N).* Similar to extended Kalman filter, an approximation for the covariance $\mathbf{P}(k-N)$ can be constructed by solving the Riccati equations with local linearization of operators F[.] and H[.] (Bakshi and Rawlings, 2006; Zavala et al., 2008).

2.3. Constrained Ensemble Kalman Filter

The filter is initialized by drawing N samples $\hat{\mathbf{x}}^{(i)}(0\,|\,0)$ from the given distribution. At each time step, N samples $\left\{\mathbf{w}^{(i)}(k), \mathbf{v}^{(i)}(k) : i = 1,2,\ldots N\right\}$ are drawn randomly using the distributions of the state noise and the measurement noise. These sample points together with particles $\left\{\hat{\mathbf{x}}^{(i)}(k-1\,|\,k-1) : i = 1,2,\ldots N\right\}$ are then propagated through the system dynamics to obtain transformed particles as follows:

$$\mathbf{x}^{(i)}(k\,|\,k-1) = F\left[\mathbf{x}^{(i)}(k-1\,|\,k-1), \mathbf{u}(k-1)\right] + \mathbf{w}^{(i)}(k-1) \quad (7)$$

for i = 1,2,…N. The transformed sample points that violate the bounds are projected on the constraint boundary to obtain a cloud of predicted particles

$\left\{\hat{\mathbf{x}}_c^{(i)}(k\,|\,k-1) : i = 1,2,\ldots N\right\}$) where $\hat{\mathbf{x}}_c^{(i)}(k\,|\,k-1) = \mathbf{P}_r\left[\hat{\mathbf{x}}^{(i)}(k\,|\,k-1)\right]$ (8)

and $\mathbf{P}_r[.]$ represents the projection operator. The sample covariance is estimated as

$$\bar{\mathbf{x}}_c(k\,|\,k-1) = (1/N)\sum_{i=1}^{N}\hat{\mathbf{x}}_c^{(i)}(k\,|\,k-1) \quad (9)$$

$$\mathbf{P}(k\,|\,k-1) = (1/N)\sum_{i=1}^{N}\boldsymbol{\varepsilon}^{(i)}(k\,|\,k-1)\left[\boldsymbol{\varepsilon}^{(i)}(k\,|\,k-1)\right]^T \quad (10)$$

where $\boldsymbol{\varepsilon}^{(i)}(k\,|\,k-1) = \hat{\mathbf{x}}_c^{(i)}(k\,|\,k-1) - \bar{\mathbf{x}}_c(k\,|\,k-1)$. The updated state estimates are then obtained by solving the following set of constrained optimization problems

$$\underset{\mathbf{x}(k)}{Min}\ \left[\boldsymbol{\xi}^{(i)}(k)^T \mathbf{P}(k\,|\,k-1)^{-1}\boldsymbol{\xi}^{(i)}(k) + \mathbf{e}^{(i)}(k)^T \mathbf{R}^{-1}\mathbf{e}^{(i)}(k)\right] \quad (11)$$

$$\boldsymbol{\xi}^{(i)}(k) = \mathbf{x}(k) - \mathbf{x}_c^{(i)}(k\,|\,k-1) \quad ; \quad \mathbf{e}^{(i)}(k) = \mathbf{y}(k) - H\left[\mathbf{x}(k) + \mathbf{v}^{(i)}(k)\right]$$

subject to constraints $\mathbf{x}_L \le \mathbf{x}(k) \le \mathbf{x}_U$. The solutions of the above set of optimization problems yields updated particles $\left\{\hat{\mathbf{x}}^{(i)}(k\,|\,k) : i = 1,2,\ldots N\right\}$. The updated state estimate is then computed as follows

$$\bar{\mathbf{x}}(k\,|\,k) = (1/N)\sum_{i=1}^{N}\hat{\mathbf{x}}^{(i)}(k\,|\,k) \quad (12)$$

2.4. Generation of Truncated Distribution of Initial State

When states have bounds, it becomes necessary to generate particles that are consistent with these bounds. This can be achieved by using the concept of truncated distributions. A truncated distribution is a conditional distribution that is conditioned on the bounds on the random variable. In particular, the truncated univariate normal distribution can be obtained as follows:

$$N_T\left[0,1 \mid a \le \zeta \le b\right] = \frac{1}{\sigma\sqrt{2\pi}} \frac{\exp\left(-\zeta^2/2\right)}{\phi(b) - \phi(a)} \tag{13}$$

Where $\phi(a)$ and $\phi(b)$ are the cumulative normal distribution functions evaluated at the lower and upper bounds, respectively. Now, consider a situation in which the distribution of random variable vector $\mathbf{x}$ is approximated by truncated multivariate Gaussian density function, denoted as $\mathbf{N}_T\left[\bar{\mathbf{x}}, \mathbf{P}\right]$, and defined over bounds $\left(\mathbf{x}_L, \mathbf{x}_H\right)$. In this work, we make use of the following approach for the generation of samples from the truncated normal distribution (Prakash et al., 2007). Since $\mathbf{P}$ is a symmetric and positive definite matrix, Cholesky factorization of $\mathbf{P}$ yields a lower triangular matrix $\mathbf{S} = \sqrt{\mathbf{P}}$. Now, a transformed random vector $\mathbf{T}$ is defined such that $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{ST}$. With this transformation, we can now define n-truncated univariate normal distributions $N_T^{(i)}\left[0,1 \mid t_{L,i} < t_i \le t_{H,i}\right]$ where the limits of i^{th} truncated distribution are defined as follows:

$$t_{L,i} = \left(\mathbf{x}_{L,i} - \bar{\mathbf{x}}_i - \sum_{r=1}^{i-1} s_{ir} t_r\right) / s_{ii} \quad ; \quad t_{H,i} = \left(\mathbf{x}_{H,i} - \bar{\mathbf{x}}_i - \sum_{r=1}^{i-1} s_{ir} t_r\right) / s_{ii} \tag{14}$$

where s_{ij} represents (i,j)'th element of matrix **S**. The above transformation requires that we draw samples recursively. Thus, first t_1 can be drawn from $N_T^{(1)}\left[0,1 \mid t_{L,1} \le \zeta \le t_{H,1}\right]$. Then, after setting limits $\left(t_{L,2}, t_{H,2}\right)$, t_2 can be drawn from $N_T\left[0,1 \mid t_{L,2} \le \zeta \le t_{H,2}\right]$, and so on. Thus, a sample from the truncated normal distribution $\mathbf{N}_T\left[\bar{\mathbf{x}}, \mathbf{P}\right]$ can be expressed as follows

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{S}T \quad ; \quad T = \left[t_1 \quad t_2 \quad \ldots \quad t_n\right]^T \tag{15}$$

In the Constrained EnKF formulation, the above mentioned procedure is used for generating the initial ensemble.

3. Estimation of Arrival Cost using C-EnKF

There are two conceptual difficulties associated with the arrival cost approximation described in Section 2.2.

(a) If we decide to use normal distribution to approximate the arrival cost, it is important to use truncated normal distribution in the presence of bounds on states.

(b) The covariance update step involves linearization of nonlinear operators F[.] and H[.] in equations (1) and (2), respectively. Depending on the nature of local nonlinearity, this further introduces error in the approximation of the arrival cost.

A possible remedy to the problem of approximating arrival cost can be to construct the conditional density $p\left[\mathbf{x}(k-N \mid \mathbf{Y}_{k-N})\right]$ using a particle filter formulation that systematically handles the constraints. Since particle filters use Monte Carlo simulation approach, there is no assumption involved on the distribution of the conditional density function. If we allow infinite sample size, then, in principle, we can construct an accurate approximation of the moments of the conditional probability density function even if it is multi-modal and non-Gaussian. Moreover, operators F[.] and H[.] can be directly used for estimation of the moments without requiring the linearization step.

In this work, we retain the form of arrival cost approximation given by equation (6). The second moment of $p\left[\mathbf{x}(k-N) \mid \mathbf{Y}_{k-N}\right]$ necessary for constructing this term is approximated by employing C-EnKF over interval *[k – N - 1, k - N]*. Defining posteriori estimation error as

$$\boldsymbol{\varepsilon}^{(i)}(k-N\,|\,k-N)=\hat{\mathbf{x}}^{(i)}(k-N\,|\,k-N)-\bar{\mathbf{x}}(k-N\,|\,k-N), \quad (16)$$

the covariance $\mathbf{P}(k-N\,|\,k-N)$ can be estimated as follows

$$\mathbf{P}(k-N\,|\,k-N)=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{\varepsilon}^{(i)}(k-N\,|\,k-N)\left[\boldsymbol{\varepsilon}^{(i)}(k-N\,|\,k-N)\right]^{T} \quad (17)$$

The arrival cost is the approximated by setting $\mathbf{P}(k-N)=\mathbf{P}(k-N\,|\,k-N)$ in equation (6). It may be noted that some modifications are necessary in C-EnKF formulation described in Section 2.3 while using it for approximation of arrival cost. In the conventional formulation, the particles $\left\{\hat{\mathbf{x}}^{(i)}(k-N\,|\,k-N):i=1,2,...N\right\}$ are propagated while performing computations over interval $\left[k\text{-}N,\ k\text{-}N+1\right]$. In the proposed formulation, however, we only propagate the cloud of estimation errors $\left\{\boldsymbol{\varepsilon}^{(i)}(k-N\,|\,k-N):i=1,2,...N\right\}$ while performing the computations over the interval $\left[k\text{-}N,\ k\text{-}N+1\right]$. Thus, while estimating $\mathbf{P}(k-N+1)$, new particles at the beginning of the interval $\left[k\text{-}N,\ k\text{-}N+1\right]$ are generated as follows

$$\hat{\mathbf{x}}^{(i)}(k-N\,|\,k-N)=\tilde{\mathbf{x}}(k-N)+\boldsymbol{\varepsilon}^{(i)}(k-N\,|\,k-N)\ \ for\ \ i=1,2,...N \quad (18)$$

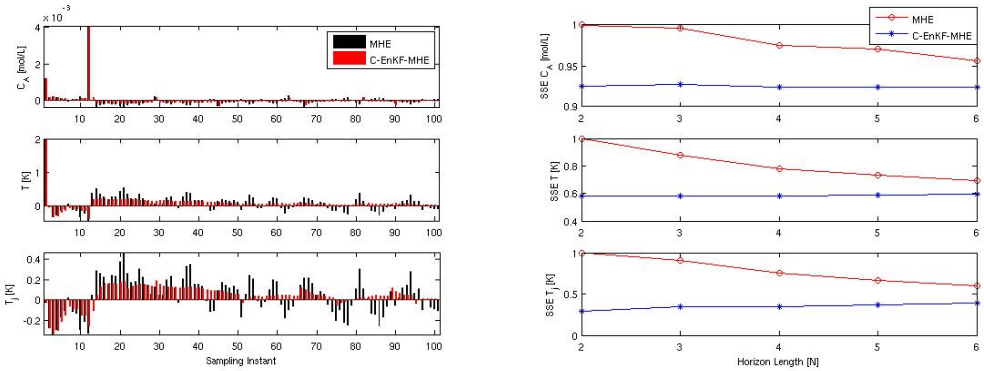
where $\tilde{\mathbf{x}}(k-N)$ represents the estimate generated by MHE problem solved over interval [*k-N, k*]. This modification facilitates *alignment* of the particles with the estimates generated using MHE.

4. Simulation Example

The performance of the C-EnKF based MHE algorithm is illustrated and compared with the MHE using a simulation example taken from the literature (Qu and Hahn, 2008). This example consists of a non-isothermal continuously stirred tank reactor model which includes reactor concentration (C_A), reactor temperature (T), and cooling jacket temperature (T_J) dynamics. The cooling water temperature is considered as the only measured variable. The horizon length used in MHE has 3 sampling times, with 20 samples used at each time point for the C-EnKF. State and measurement noise were considered zero mean Gaussian. The model parameters, initial values, and noise covariance matrices were taken from (Qu and Hahn, 2008)., The MHE problem was formulated using the simultaneous dynamic optimization formulation where the dynamic model is approximated with orthogonal collocation, and the resulting NLP problems were solved using IPOPT (see Zavala et al., 2008).

The evolution of the estimation errors is shown in **Erro! Fonte de referência não encontrada.**(a) where it is shown that the estimation errors for the C-EnKF-MHE are almost always smaller that the MHE. Figure 1(b) shows the sum of squared errors (SSE) for each state as a function of horizon lengths (from 2 to 6 sampling times) for both methods. Here, the SSE values have been normalized by the maximum SSE, which corresponds to the MHE using a horizon length of 2. From this figure it is easy to see that the performance of C-EnKF-MHE is significantly better than MHE for all the choices of horizon lengths. Moreover, there is hardly any change in the performance of C-EnKF-MHE when the horizon length is increased from 2 to 6. This improvement may be attributed to the fact that the approximation of the state error covariance matrix in the C-EnKF-MHE is statistically more consistent with bounds on states. Furthermore, at each sampling time the samples for the states used in the C-EnKF are not being redrawn. Instead, because the state samples are propagated over time, no assumption is needed on the distribution for the state estimation error. Moreover, unlike the

conventional MHE, the covariance estimation step does not involve linearization of the model as the covariance is approximated by Monte Carlo simulations. Additional computations arising from C-EnKF can be an issue for online implementation. However, when using the C-EnKF to approximate of the arrival cost, it is possible to achieve good estimator performance using even smaller horizon lengths. Working with shorter horizon lengths significantly reduces the on-line iterative computations. Also, since the measurements at the beginning of the horizon are known, the C-EnKF calculations can be done in the background between any two sampling times, thereby reducing the computation load at any sampling time.



(a) (b)

Figure 1: (a) State estimation errors. (b) Normalized SSE with as a function of horizon length.

5. Conclusions

A C-EnKF based scheme for updating the arrival cost in MHE is presented in this paper. It generates an approximation of the state error covariance matrix that is consistent with bounds on the state variables and without making assumptions on the distribution of the state estimation error. The simulation results demonstrate that the proposed formulation performs significantly better than the conventional MHE formulation. Moreover, the simulations reveals that MHE with short horizon lengths can be made to work well if the arrival cost estimation is improved.

References

V. M. Zavala, C. D. Laird, and L. T. Biegler. A Moving Horizon Estimation Algorithm Based On Nonlinear Programming Sensitivity. *Journal of Process Control*, In Press, 2008.

J. B. Rawlings and B. R. Bakshi (2006), Particle Filtering and Moving Horizon Estimation. *Computers and Chemical Engineering*, **30**, 1529–1541.

G. Evensen (2003) The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation, *Ocean Dynamics*, 53, 343–367.

J. Prakash, S. C. Patwardhan and S. L. Shah (2008a), Constrained State Estimation Using the Ensemble Kalman Filter, *Proc. of the 2008 American Control Conference*, Seattle, June 2008.

C. C. Qu and J. Hahn. Computation of Arrival Cost for Moving Horizon Estimation via Unscented Kalman Filtering. *Journal of Process Control*, In Press, 2008.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

SLUG CONTROL STRUCTURES FOR MITIGATION OF DISTURBANCES TO OFFSHORE UNITS

Diego Di D. Pinto,[a] Ofélia Q. F. Araújo,[a] José Luiz de Medeiros,[a] Giovani C. Nunes[b]

[a]*Escola de Química, Universidade Federal do Rio de Janeiro. Av. Horácio Macedo, 2030, Ilha do Fundão, Rio de Janeiro, RJ, Brazil, CEP 21941-909*
[b]*PETROBRAS, ENGP/TPP/PMF, Av. Chile, 65 – Centro, Rio de Janeiro - RJ, Brazil, CEP 20035-900*

Abstract

Petroleum has been used by society in many ways, but mainly as an energy source. To meet the increasing energy demand, oil companies seek for new technologies to optimize production in a field. For this purpose, gas is employed for artificial lift, therefore enhancing productivity of old fields which suffer from insufficient pressure to force the oil out to the surface. Often, gas lift is associated to intermittent flow regime, slug flow that promotes intense disturbances to downstream offshore processes. In the present work, three control strategies for a gas-liquid separator are proposed: (i) feedback averaging level control; (ii) feedback + feedforward level control employing an ARX estimator; and (iii) override control on the production choke valve. The control schemes attempt to minimize downstream slug effects by exploiting the separator's hold-up, while imposing limits to liquid level and maintaining pressure under satisfactory performance.

Keywords: gas-lift, phase separators, slug flow, process control, ARX.

1. Introduction

Most nations rely on petroleum to meet their energy demand, requiring increased efficiency in the extraction process. In offshore oil fields, risers are used to transport a multiphase mixture (composed by oil, gas, water and sand) from the wellheads to separation tanks on producing platforms. In case of inexistence of separation and pumping facilities near the wellheads, on the sea bottom, this multiphase mixture must be propelled to the sea surface at expenses of the reservoir pressure. For deep water, a common situation is that the reservoir is not sufficiently pressurized to promote the flow of oil at satisfactory rates, requiring artificial ascension of oil.

The method is based on reducing the production column hydrostatic pressure by reducing its average density. The reduction of de density is given by injecting an elevating fluid, usually natural gas. Reducing this hydrostatic pressure will allow the reservoir pressure to force the oil out to the surface. The injection of the gas must be controlled; otherwise, it can promote pressure surges in the ducts and related equipments (Robinson *et al*, 2001). In some cases, the continuous gas lift, associated to high injection rate, is not economically efficient, requiring reduced injection rates responsible for intermittent flow regime (Santos *et al*, 2001). Figure 1 shows a gas lift process where: (i) gas is injected in the annular region; (ii) gas passes to the production

tube through a check valve located slightly before the reservoir; (iii) in the production tube, the gas is mixed with the oil reducing its density; (iv) as the density is reduced, the base hydrostatic pressure decreases; (v) being this pressure lower than the reservoir pressure, the column fluid ascension occurs pushed by the reservoir pressure; (vi) since the gas injection rate is low, the column ascension will carry the gas out from the column, increasing the density and, hence, the bottomhole pressure what will close the check valve, stopping the gas inflow; (vii) production ceases until enough gas is injected into the annular column, restarting the cycle.

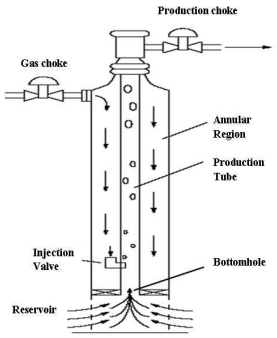


Figure 1: Gas Lift Process.

In consequence, high flow rates of gas have to be accommodated in the system of risers for further recovery by gas-liquid separators at the platform, before re-compression and re-injection. In this scenario, and depending on the flow conditions, a flow regime called *riser slugging* can develop in the pipelines. This regime, naturally highly transient, introduces severe oscillations of pressure and flow rates into the system. On the other hand, space and load constraints for offshore units favor compact equipments which increase sensitivity of downstream processes, mostly in the early stage of phase separation. Multiphase separators are, hence, a key-step in offshore facilities, with a two-fold process objective: (a) phase separation; and (b) dampening *riser slugging*, as in surge tanks. Faanes and Skogestad (2003) defined a *buffer tank* as a unit at which the holdup is explored to promote smooth operation. While in some level applications it is desirable to maintain tight regulation, such as in a reactor to achieving maximum capacity, this is seldom the case for level control in surge tanks, as it betrays the anti-surge requirement.

In the present work, three control strategies for a gas-liquid separator are proposed: (i) feedback averaging level control; (ii) feedback + feedforward level control employing an ARX estimator; and (iii) override control on the production choke valve. The control schemes attempt to minimize downstream slug effects by exploiting the separator's hold-up, while imposing limits to liquid level and maintaining pressure under satisfactory performance.

2. Process Model

For testing the control strategies and developing the stochastic predictor, a process model is developed based on Aamo *et al*, 2005. The dynamic behavior of the gas lift derives from the interaction between the annular and tubular volumes above the gas injection point, described by mass balances shown in Eqs. 1, 2 and 3, where H and W,

are, respectively, hold-up (kg), mass flow rate (kg/s). Subscripts *A*, *G*, *L*, *I*, *C*, *R* and *P*, stand respectively for annular region, gas in the production tube, liquid in the production tube, injected gas in the production tube, gas in the production choke, reservoir, production choke.

$$\frac{dH_A}{dt} = W_G - W_I \tag{1}$$

$$\frac{dH_G}{dt} = W_I - W_C \tag{2}$$

$$\frac{dH_L}{dt} = W_R - W_P \tag{3}$$

The valve responsible for injecting gas into the annular region is taken as a *swing-check-valve*. The choke valve for feeding gas to the annular region and the production choke valve are modeled as gate valves, assuming that the oil reservoir and the offshore separator are at constant pressure. The oil single phase flow section between the reservoir and the gas injection point is solved in pseudo-stationary mode with the reservoir and injection point pressures, assuming incompressible flow with friction factor given by Chilton's Equations, which comprise all hydraulic regimes.
The gas phase behavior is simplified to an isothermal ideal gas. Furthermore, temperatures in the gas lift subsystems are simplified to annular temperature equal to fed gas temperature, and production tube temperature identical to the oil temperature in the reservoir.
Figure 2a shows gas flow profiles for gas lift simulated at conditions shown on Table 1. Fixing these conditions and varying the gas injection rate leads to the identification of two distinct operating regions: (i) continuous gas lift or stable region; and (ii) intermittent gas lift or unstable region, as displayed on Figure 2b.

Table 1: Gas Lift Conditions.

	Oil	Gas
Pressure, Temperature	160 bar, 108°C	120 bar, 60°C
Density , Viscosity	850 kg/m^3, 0.012 kg/m.s	Ideal Gas
Choke Valve (opening fraction)	0.8	0.5

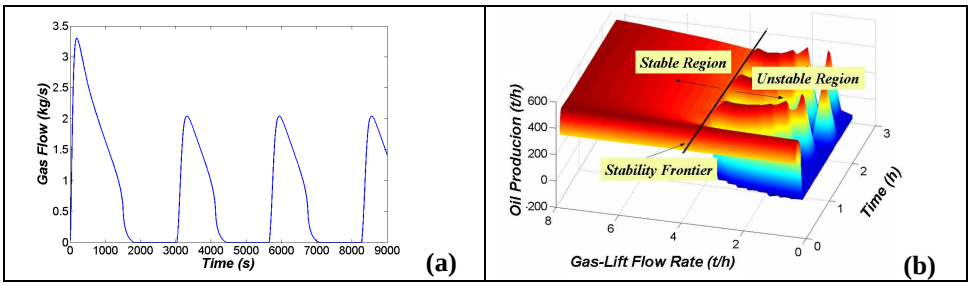


Figure 2: Simulation Results. (a) Gas Flow Profile for Gas Lift Simulated at Table 1 Conditions, (b) Sensitivity Analysis to Injection Gas Flow Rate

The gas lift model was expanded to include the operation of the gas-liquid separation. Mass balance for the liquid and gas phases yield Equations 4 and 5, where *L*, *G*, *C*, *D*

and V are respectively the liquid flow rate, the gas flow rate, separator's length, separator's diameter, and the volume. Subscripts I, o, T and L stand for, respectively, input rate, output rate, total, and liquid.

$$\frac{dh_L}{dt} = \frac{L_i - L_0}{2C\sqrt{h_L(D-h_L)}} \tag{4}$$

$$\frac{dP}{dt} = \frac{P(L_i + G_i - L_0 - G_0)}{V_T - V_L};\ V_T = C\frac{\pi D^2}{4} \tag{5}$$

3. Control Strategies

The control objectives are to maintain separator's pressure and its level in an appropriate range. Since pressure in the separator affects flow rate through the valves, it will impact the dynamics of the gas lift, and decrease the oil production. The level controller protects the separator against overflow or drainage. Also, due to the eventual occurrence of severe slug caused by the intermittent gas lift, level controller must provide buffer capacity, protecting downstream operation. For this purpose, three strategies are tested: (i) averaging level control exploiting the liquid hold up of the tank; (ii) feedback and feedforward control using an ARX predictor of separator feed flow rate; (iii) override control on the production choke valve and feedforward control. All the three strategies are shown schematically in Figure 3.

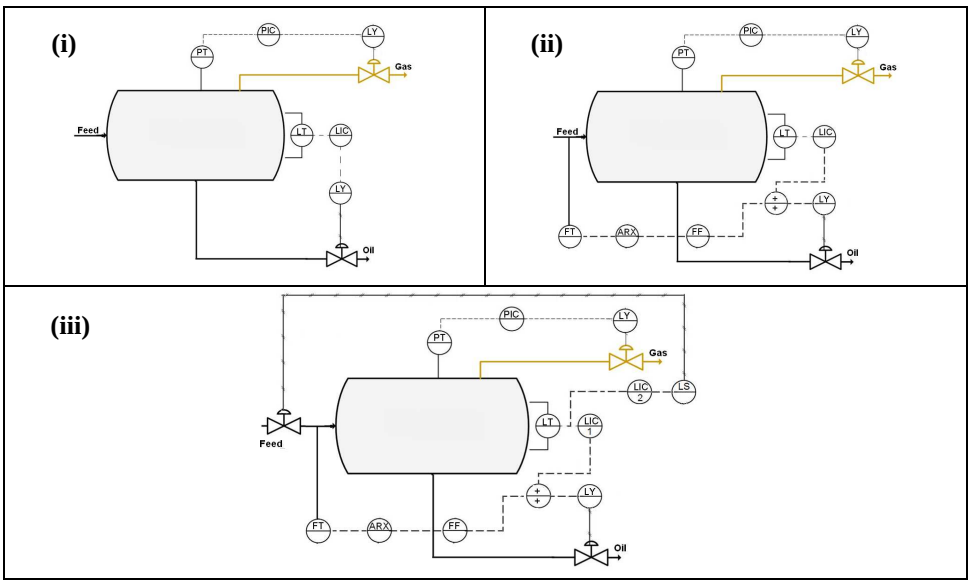


Figure 3: Control Strategies for Gas-Liquid Separators.

3.1. Averaging Level Control

Cheung and Luyben (1980) explore different level control strategies, such as the strategy proposed by Shunta and Fhervari (1976), the wide-range controller, with adaptation of PI controller tuning parameters (K_C and τ_I) according to Equation 6:

$$K_C = f\ K_{C0}; \quad \tau_I = \frac{\tau_{I0}}{f} \tag{6}$$

In this work, an alternative adaptation law proposed *f*, by Araújo *et al* (2007), is adopted, as given in Equation 7.

$$f = 1 + \frac{1}{1+e^{(\lambda_1(E_1 - abs(e))}} + \alpha\left(1 + \frac{1}{1+e^{(\lambda_2(E_2 - abs(e))}}\right) + \beta\left(1 + \frac{1}{1+e^{(\lambda_3(E_3 - abs(e))}}\right) \tag{7}$$

3.2. Feedforward + Feedback Control

For this strategy, an identification of the gas lift process is required as separators inflow is frequently not available from measurements. An ARX 3x3 with 3 inputs (gas injection rate, choke opening fraction, and separator tank pressure) and 2 outputs (oil and gas flows) is presented in Pinto (2009) and is used to inferring oil and gas flows. Faanes and Skogestad (2003) propose a model for buffer tank level response (*h(s)*) to inflow rate (*d(s)*, Equation 8) that is herein applied to a separator to obtain the feedforward control law (Equation 9). with K = -16 and τ = 343.53 s. The feedback controller is identical with same as showed at strategy 3.1.

$$h(s) = \frac{1}{\tau s + 1} d(s) \tag{8}$$

$$G_{FF}(s) = \frac{K}{\tau s + 1} \tag{9}$$

3.3. Override Control

In this strategy, under normal operating conditions, level controller (LIC1) is adapted according to Equation 7. Under severe slug, this controller is overridden though a low value switcher (LS) to a level controller (LIC2, with set-point equals to a high level limit) which acts reducing opening of the production choke valve.

Figure 4 compares the three strategies.

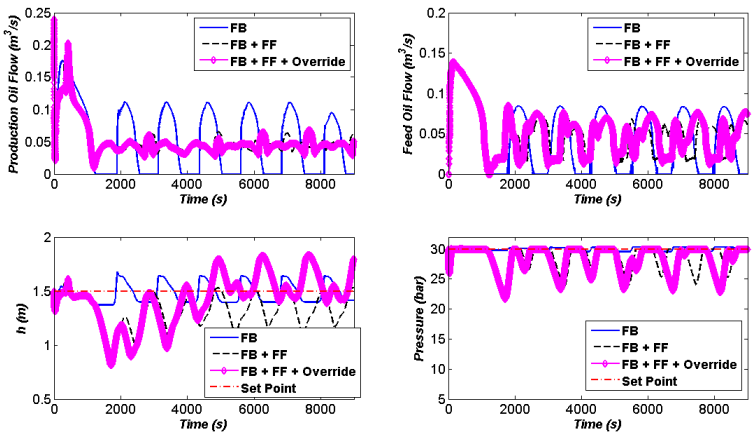


Figure 4: Comparison of Control Strategies.

Feedback level controller with adaptation law given by Equation 7 is able cope with slugs but is unable to reduce flow perturbations propagating downstream. With a feedforward combined to the adaptive feedback, it is possible to minimize the slug effects, while maintaining level and pressure in the separator within an acceptable working range. The override control is used as a security layer, preventing the separator's level from reaching dangerous values. It's possible to observe in Figure 4 that the results from both feedforward and override control (strategies ii and iii) are almost the same differing only on the level control, where the addition of the override strategy shows improves performance.

4. Conclusions

To mitigate the impact of slug flow downstream of a gas-liquid separator, this work approaches 3 control strategies. These strategies are presented to work in layers: a feedback layer, for correcting adaptively level deviations, a feedforward layer to anticipate control action based on an ARX predictor of the disturbance, the inlet liquid flow), and adopting a new level controller through override of the feedback controller.
It's possible to conclude that strategies (ii) and (iii) improve performance of an adaptive feedback controller. It is worth noting that the override layer only interferes when the separator's level reaches dangerous values.

References

AAMO, O. M., EIKREM, G.O., SIAHAAN, H.B., FOSS, B.A., 2005, Observer design for multiphase flow in vertical pipes with gas-lift – theory and experiments, Journal of Process Control, 15, 247-257.

ARAÚJO, O.Q.F., MEDEIROS.J. L., NUNES, G.C., 2007, Modeling and Control Solutions for Riser Slugging in Offshore Oil Fields, Proceeding of European Congress of Chemical Engineering.

CHUNG, T.F., LUYBEN, W.L. ,1980, Nonlinear and Nonconventional Liquid Level Controllers, *Ind. Eng. Chem. Fundam.*, 19, 93-98.

FAANES, A., SKOGESTAD, S., 2003. Buffer Tank Design for Acceptable Control Performance. *End. Eng. Chem. Res.*, v. 42, 2198-2208.

PINTO, D. D., 2009, Estratégias de Controle Contra Intermitência Severa na Alimentação de Separadores Offshore, MSc. Thesis, Escola de Química, Federal University of Rio de Janeiro.

ROBISON, C. E. , SCHNATZMEYER, M. A., Bayh III, R. I. , 2001, Self-regulating lift fluid injection tool and method for use of same. U.S. Patent No. 336459.

SANTOS, O. G., BORDALO, S. N. , ALHANATI, F. J.S.. 2001, Study of the dynamics, optimization and selection of intermittent gas-lift methods—a comprehensive model, Journal of Petroleum Science and Engineering, 32, 231 – 248.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Published by Elsevier B.V.

Embedded Control and Monitoring Systems in Production Machine Networks

Sirkka-Liisa Jämsä-Jounela,[a] Mikko Huovinen [b]

[a]*HelsinkiUniversity of Technology, Department of Biotechnology and Chemical Technology, P.O.Box 6100, FI-02015 Espoo, Finland*
[b]*Tampere University of Technology, Department of Automation Science and Technology, P.O.Box 692, FI-33101 Tampere, Finland*

Abstract

The ongoing globalization trend is tightening competition and setting new, higher efficiency requirements in the process industries. In order to enhance the efficiency of the production chain, new functionalities and information networking must be incorporated into the production equipment. This paper proposes a concept of equipment automation that utilizes new information and communication technologies and more flexibly adds equipment intelligence. The feasibility and applicability of the concept is demonstrated via case studies of a grinding circuit and a pulp drying process.

Keywords: Networked equipment automation, intelligent process equipment.

1. Introduction

The globalization process is tightening competition and setting, higher efficiency requirements in the process industries. In addition, factories now have to be considered as a part of sustainable development. Energy savings and pollution prevention have become priorities. Safety and maintenance are also listed as the main equipment design parameters. At the present time, however, the intelligence level of the process equipment and their networks is still low. In order to enhance the efficiency of the production chain, new functionalities and information networking must be incorporated into the production equipment. This can be further enhanced by applying embedded automation middleware to manage complex networked machines and distributed resources.

The goal of automation middleware is to increase the intelligence level of the production machines by networking and integrating the embedded systems of the machines into larger automation system entities. The functionalities embedded in the middleware cover the management of the whole process system at the life-cycle scale. The methods can utilize connections to higher-level systems and remote resources, as well as to the information-refining tools provided by the middleware platform. The middleware thus integrates lower level process resources, and further refines the collected information into a more convenient form that is subsequently presented to the users and/or to other information systems, such as MES or ERP.

The purpose of this paper is to propose a common way of incorporating the current communication technologies, and especially the communication protocol and network topology design, into process equipment in such a way that it facilitates the implementation of new methods and subsequently satisfies the efficiency requirements.

2. Survey of Industrial Automation Architecture and Communication Technologies

The evolution of communication technologies has had a strong influence on changes in the structure of industrial automation systems. Up until now, communication support in plant automation systems has been defined according to the Computer Integrated Manufacturing (CIM) concept. In this hierarchical structure, levels of functionality are identified in such a way that each device is designed for a specific task, and specific networks are used to interconnect devices at the same level, i.e. running the same task.

However, the devices have recently started to include more than one function, or module, which increases the intelligence level of the equipment automation. Devices like sensors that have traditionally been used for measurement now have to support e.g. maintenance or monitoring tasks. This means that the traditional hierarchical structure is increasingly being replaced by a distributed communication architecture. Nevertheless, the hierarchical structure still exists - and this is also advisable - in most of the process control strategies and plant automation systems, as illustrated in Fig. 1 and 2. This phase in the evolution of automation systems has been called FCS (Field Control System).

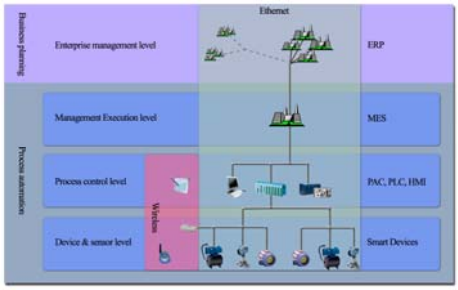


Fig. 1. Hierarchy of the plant automation.

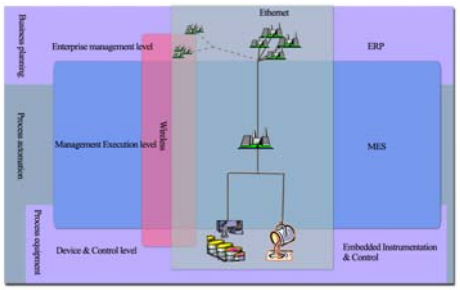


Fig. 2. Future scenario of the plant automation

2.1 Low-layer communication protocols

Currently the most widely available industrial networks can be classified into three main categories: traditional fieldbusses, Ethernet-based networks and wireless networks. A comparison of the key properties of the currently, most widely available networks in each of the three main categories is given in Table 1.

2.2 High-layer data specifications

The information standards for process operation and maintenance are driven by OpenO&M Initiative joint working groups, which mainly represent three industrial organizations: MIMOSA, the OPC Foundation, and ISA's SP95. One of the most strongly established standards, the OPC, also enables the use of state-of-the-art technologies such as web services. On the other hand, the traditional fieldbusses (like Profibus or Foundation Fieldbus) have defined concepts for the manufacturer-independent integration of field devices such as FDT/DTM.

3. Proposed Architecture for the Networked Equipment Automation

3.1 Components of the automation concept

The concept of equipment automation must be applicable over a wide range of different industrial processes. The main starting point is to bring the automation, with intelligent software-based functionalities, near to the process equipment while, on the other hand,

designing a communication network that robustly provides the services that are already incorporated. Therefore, the automation devices for one specific piece of equipment or sub-process are networked into one process system node, thereby providing connectivity to other parts of the process. The equipment 'intelligence' is embedded as a part of the automation devices, using software-based modules (Fig.3.)

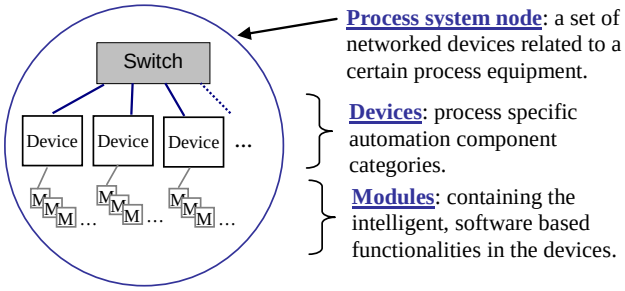


Fig. 3. Structure of the process system node.

The process system nodes are further connected to the supervisory automation level with a high bandwidth network (Fig. 4). In addition, the nodes can also provide a wireless connection link. Finally, in order to implement the equipment's advanced operating algorithms and procedures, a data processing unit (PLC or soft-PLC) is needed as a part of the process system node. This leads to a scheme in which the equipment automation provides the intelligent operating functionalities related to a specific task, which typically means one unit operation in the process chain. Compared to the traditional DCS hierarchy, the addition of intelligence offers a more dynamic platform for autonomous systems. Intelligent devices can act more independently and the improved communications enable interaction between the distributed assets. In fact, the FCS hierarchy provides a new platform for the implementation of agent technologies in the process industry.

3.1. Communication network topology

The architecture of the concept relies on a *backbone* that is capable of providing the main connectivity between the different devices. Moreover, it includes the connection point to remote resources via the Internet. The process system nodes, which are linked to the backbone, form sub-networks. These provide supplementary networks for establishing the instrumentation power supply and wireless links. It also manages the mutual interconnections between local devices. Redundancy is also provided for the backbone. The network system topology is presented in Fig. 5.

The communication protocol applied in the backbone should be able to link with office and remote resources and, as a result, have the same low layer technology as used in Internet networks, to provide a large bandwidth to integrate all the sub-networks, and also to support several high layer technologies, such as web and multimedia applications. The use of *Ethernet* networks is therefore proposed. The Ethernet technology is applied in the process system node network level. The connection to the devices is shared by using *switch* in order to improve the time determinism and isolate the local traffic. The Ethernet also has the advantage of facile connectivity to traditional fieldbusses. Additionally, Ethernet ensures device interoperability by supporting open standards, such as OPC.

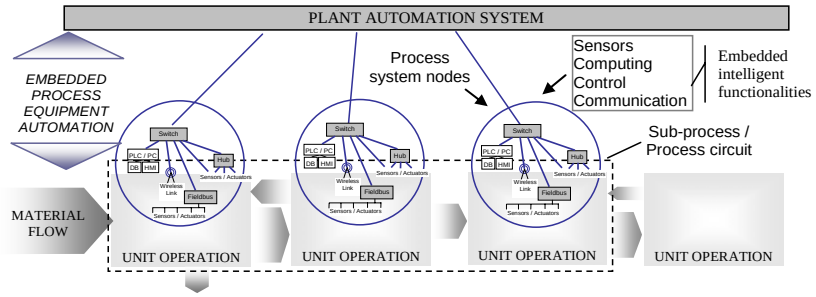


Fig. 4. Concept of the equipment automation level between the physical process and the supervisory plant automation system.

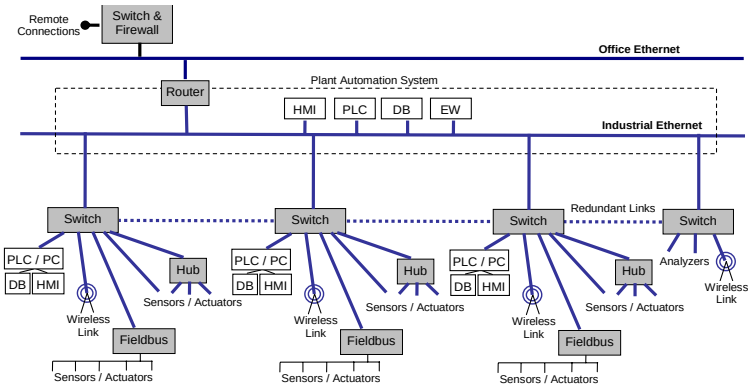


Fig. 5. Topology of the equipment automation network architecture.

Table 1. Comparison of the currently, most widely available industrial networks

Type	Protocol	Manufacturer	Throughput	Cycle Time	Range (Length)	Number of Devices	Energy Efficiency
Traditional Fieldbus	CAN	Bosch	10kb/s - 1Mb/s		40m (1Mb/s), 5km (10kb/s)	max. 32	
	WorldFIP(1158-2)	Schneider	31,25 kb/s, 1Mb/s, 2,5Mb/s (5Mb/s optical fiber)	2 ms, 5 ms	1km (1Mb/s), max 4km	max. 256, 64/seg	Power over network solution
	Profibus DP	Siemens	9,6kb/s - 12Mb/s	depending on configuration <2ms	100m (12Mb/s), 1,2km (9,6kb/s)	max. 126	
	Profibus PA	Siemens	31,25 kb/s	depending on configuration <2ms	1900m	max. 32/seg	Power over network solution
	DeviceNet			2,0 ms, 4,2 ms, 10 ms			
	ControlNet	Rockwell Automation	5Mb/s	<0.5 ms	5km (5Mb/s), 30km+ optical fiber)	max. 99	
	Interbus	Phoenix Contact	500kb/s	1,8 ms, 7,4 ms, 140 ms	~13km	max. 512	
	AS-i	Siemens	167kb/s	4,7 ms	100m/seg, 300m	max. 64	Power over network solution
	Foundation Fieldbus H1	Fieldbus Foundation	31.25 kb/s	36 ms,100 ms, 600 ms	max 1900m	max. 32	Power over network solution
Type	Protocol	Manufacturer	Throughput	Cycle Time	Range (Length)	Number of Devices	Energy Efficiency
Ethernet Based	Ethernet/IP	Rockwell Automation	10Mb/s, 100Mb/s, 1Gb/s	>10 ms			
	Foundation Fieldbus HSE	Fieldbus Foundation	100Mb/s		100m/seg		
	EtherCat	Beckhoff	100Mb/s	<50 µs	100m/seg	~65535	Power over network solution
	Powerlink	B&R	100Mb/s	<500 µs	100m/seg		
	ProfiNet IO	Siemens	100Mb/s	1 ms, 10ms	100m/seg		
	Sercon III	Bosch Rexroth	100Mb/s	31,5 µs	100m/seg	max. 254	
	Modbus TCP	Schneider	10Mb/s, 100Mb/s, 1Gb/s	>10 ms			
Type	Protocol	Manufacturer	Throughput	Cycle Time	Range (Length)	Number of Devices	Power Consumption
Wireless	IEEE 802.11 (b/g/n)		54Mb/s, 100Mb/s		30-100m		Medium
	Bluetooth, IEEE 802.15.1		1Mb/s		10m	7	Low
	IEEE 802.15.4 /ZigBee		20kb/s, 40kb/s, 250kb/s		10m	~65k	Very Low
	WirelessHart	Hart Communication Foundation	250 kb/s		10m		

4. Add–On Service Modules of the Intelligent Process Equipment

Intelligent process equipment in modern automation systems can also be thought of as forming a functional hierarchy. The equipment provides *services* designed to meet the higher level *objectives*. There are production management services consisting of daily operations, abnormal event management, operator support and training, product data and lifecycle management services. All these consist of lower level *activities*, such as process control, performance monitoring and field management. The implementation of the activities consists of *tools and routines* as presented in Fig. 6.

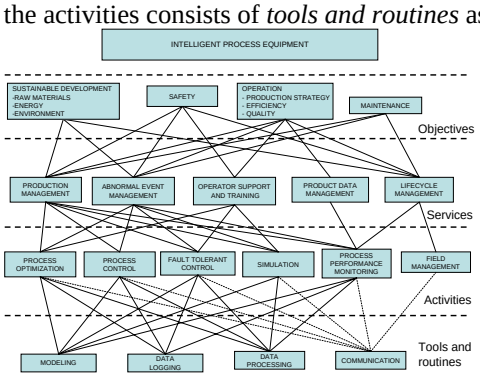


Fig. 6. Functional hierarchy of the intelligent process equipment

5. Case Examples

5.1. Case example: applying the concept of equipment automation to a grinding circuit

The proposed equipment automation concept is applied to an example of a mineral grinding process. It is proposed that the instrumentation of the process is grouped into four categories, typical for the process. These are: basic sensors, machine vision, process analyzers, and data processing units.

The process system nodes are divided according to the data processing units for the operation and monitoring of the mill, pump and hydrocyclone. The fourth node is reserved for an elemental analyzer, which is located physically further from the grinding circuit. One of the most important nodes, which includes the 'mill automation' data processing unit, has a wireless link enabling connection to the particle size and elemental assay analyzers. Other wireless sensors located in problematic points, such as the rotating mill shell, can also be added to the system. The industrial Ethernet connection is favored for all the instrumentation connections, as illustrated with the 'machine vision' device in the figure. The network topology of the equipment automation in this case is illustrated in Fig. 7.

5.2. Case example: applying the concept to monitor a pulp drying process

The concept was also implemented and deployed in the drying section of a Finnish pulp mill. The application is primarily intended to be used for production management services through process optimization and performance monitoring tools. It is also useful for abnormal event management, lifecycle management and operator support. The tools available for process optimization and performance monitoring include clustering, conditional histogram and control loop monitoring. An extensive performance measurement system has also been developed. The measurement system is hierarchical representing the whole system, subprocesses, control loops and the

equipment. On the equipment level features of intelligent field devices were exploited while on the control loop level a control loop monitoring tool called LoopBrowser was used. Examples of subprocess measurements are relative energy consumption and relative variations of controlled variables. The whole drying process was represented by an OEE indicator.

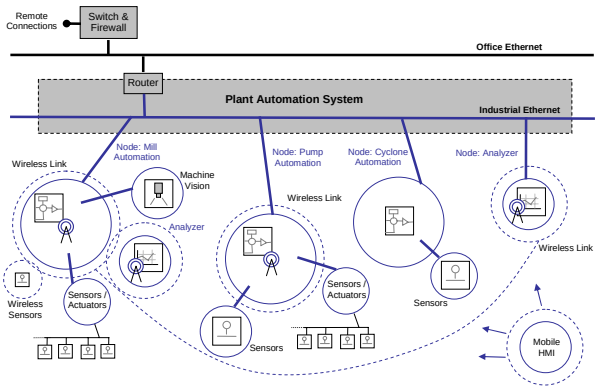


Fig. 7. Simplified abstraction of the network topology, with device connections for the process.

The clustering method is based on a modified K-means algorithm for identifying operating points and for detecting abnormal situations. The algorithm was chosen based on its computational simplicity enabling easy online implementation. The information is also used for diagnostic purposes because the application provides a contribution plot for showing the deviations of the input variables from their typical values. The conditional histogram method is used for constructing a database representing normal performance in each operating point. The method also includes routines for detecting whether the current measured performance deviates from the normal performance. The control loop monitoring system produces data about the performance of the monitored control loops. Four performance indices are used, which basically measure the accuracy and stability of the control loops.

The refined information serves the lifecycle module by providing monitoring performance throughout the equipment lifecycle. Intelligent field devices can be incorporated to deepen the scope of the application through embedded device level monitoring, as well as diagnostics which provide valuable information about the performance and ambient conditions of the field devices. For example, valve positioners typically provide control error and stiction measurements that can be used for determining whether a detected abnormal situation originates from the monitored valve-actuator-positioner package or from an outside source. Efficient exploitation of field device level information requires the use of device integration technologies, such as FDT/DTM. The overall performance measurement system was designed to give a comprehensive picture of the overall state of the whole process through the use of sub-process performance indicators. These measure e.g. relative variations in the controlled variable and efficiency measures.

6. Conclusions

In order to meet the tightening performance requirements in the process industries, the efficiency of the process equipment must be improved. Therefore, an equipment automation concept was formulated and proposed The concept outlines the procedure for efficient utilization of the latest information and communication technologies.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Multivariate Statistical Control of Emulsion and Nanoparticle Slurry Processes Based on Process Tomography, Dynamic Light Scattering, and Acoustic Sensor Data

Rui F. Li,[a] Lande Liu,[a] Xue Z. Wang,[a] Richard Tweedie,[b] Ken Primrose,[b] Jason Corbett,[c] Fraser McNeil-Watson[c]

[a] *Institute of Particle Science and Engineering, University of Leeds, Leeds LS2 9JT,UK*
[b] *Industrial Tomography Systems Ltd, Speakers House, 39 Deansgate, Manchester M3 2BA, UK*
[c] *Malvern Instruments Ltd, Enigma Business Park, Malvern, Worcestershire WR14 1XZ,UK*

Abstract

This paper describes the use of multiple on-line sensors including electrical resistance tomography (ERT), dynamic light scattering (DLS) and ultrasound spectroscopy (USS) for real-time characterization of process operations processing emulsions and nanoparticle slurries. The focus is on making novel use of the spectroscopic data to develop multivariate statistical process control (MSPC) strategies. The ERT data at different normal operating conditions was processed using principal component analysis and used to derive two MSPC statistics, T^2 and SPE (squared prediction error) for detecting abnormal changes in mixing conditions. The corresponding particle size distribution was monitored using DLS and USS. Two case studies, a sunflower oil–water emulsion system and a silica suspension system, were examined.

Keywords: dynamic light scattering, acoustic spectroscopy, process tomography, multivariate statistical process control.

1. Introduction

An experimental system for studying on-line sensing techniques for nanoparticle slurries and emulsions has been built at the University of Leeds. The rig is equipped with dynamic light scattering (DLS) and ultrasound spectroscopy (USS) instruments for particle size and z-potential measurement, as well as electrical resistance tomography (ERT) for characterising the mixing conditions and solid concentration. In addition, flowrate, pH and temperature are also measured. This paper describes a novel method for making effective use of the spectroscopic to develop multivariate statistical process control (MSPC) charts.

2. Materials, Instruments and Experiments

2.1. Materials

Two materials were investigated in this study, a sunflower oil – water emulsion and a silica suspension. The emulsion was manufactured by a membrane emulsification rig in our lab. The oil volume concentration is 33%, and the mean size of droplets is around

700 nm. The silica suspension, Nissan SnowTex ZL, of 24.0% vol solid concentration and particle mean size between 70 nm – 100 nm was purchased from Nissan America Incorporation.

2.2. Experimental Rig and Instruments

Figure 1 shows the flow diagram of the experimental rig. It consists of mainly two parts, a 10 litre jacket reactor and a circulation loop. The 10 litre reactor is used as a buffer tank to store the multi-phase fluid. It is equipped with sensors of temperature, pH and conductivity. A Julabo water bath (FP50) is used to control the temperature inside the tank. Mixing in the tank was achieved using a pitched-blade stirrer. A mono pump with flow control is used to feed fluid to the circulation loop. A flow metre and a glass window were also fitted into the loop for visual observation of the flow. The ERT and USS systems were directly connected to the circulation pipe. A by-pass pipe with flow control and indication was designed for connection of the DLS sensor.

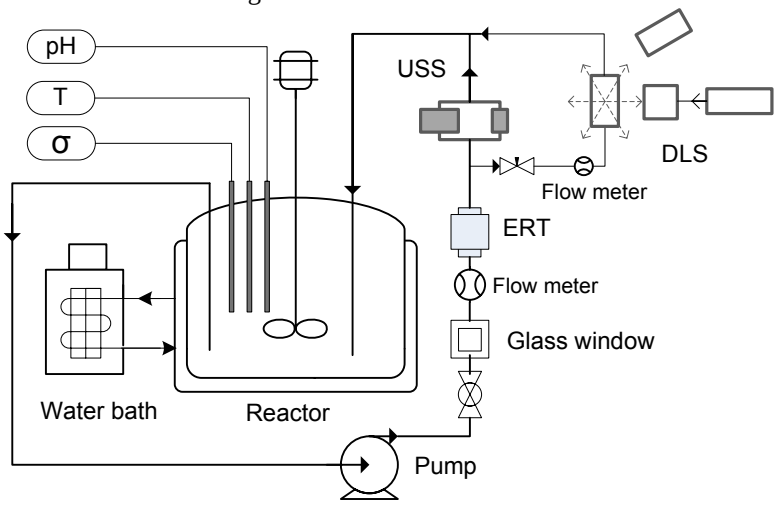


Figure 1. The experimental rig

The ERT instrument was provided by Industrial Tomography Systems Ltd [1]. It has 16 electrodes and can make 104 voltage measurements in approximately 25 ms. The tomography images are reconstructed from the 104 voltage measurements using a reconstruction algorithm. Both USS and DLS are supplied by Malvern Instruments Ltd [2].

2.3. Experiments

The purpose of the experiments is to investigate the mixing performance and measure particle size for two particulate processes at various concentrations and flowrates. During the experiments, in order to obtain different concentrations, the original emulsion, or suspension was gradually diluted to lower concentrations by adding sodium dodecyl sulfate surfactant solution or water into the concentrated emulsion or suspension. Under each concentration of the emulsion or suspension, different flowrates to the circulation loop were used to study the effect of flowrate on the mixing performance and particle size. During the experiments, tomography measurement was recorded at a rate of 60 frames per minute. At each concentration and each flowrate, one acoustic measurement was recorded by USS. In addition three size and three zeta-potential measurements were taken by the on-line DLS at every concentration and every

flowrate. Table 1 summarises the concentrations, flowrates and the numbers of each type of measurement obtained during the experiments for the two materials.

Table 1. Summary of measurements

Material	Volume concentration	Flowrate (l/s)	Number of DLS particle size measurement	Number of DLS zeta potential measurement	Number of USS particle size measurement
SunflowerOil – Water Emulsion	33%, 30%, 25%, 20%, 15%, 10%, 5%, 3%, 2%, 1%	0.190, 0.166, 0.133, 0.108, 0.083, 0.055, 0.041	3×10×7	3×10×7	10×7
SnowTex-ZL Suspensions	24%, 20%, 16%, 12%, 8%, 4%, 2.3%, 1.6%, 0.8%.	0.190, 0.166, 0.133, 0.108, 0.083, 0.055, 0.041	3×9×7	3×9×7	9×7

3. Multivariate Statistical Process Control Based on ERT Measurements

A significant development in multivariate statistical process control (MSPC) in recent years has been due to the introduction of principal component analysis (PCA) and Independent component analysis (ICA) for compression of data prior to deriving the MSPC's two statistics[4,5], Hotelling's T^2 and SPE. If PCA is used, the latent variables, i.e., the first few PCs are used rather than the original variables. The Hotelling's T^2 can then be calculated by

$$T^2 = \sum_{i=1}^{k} \frac{t_i^2}{S_{ti}^2} \tag{1}$$

where S_{ti}^2 is the estimated variance of t_i, k the number of PCs used, and t_i the ith score. The second statistic SPE (squared prediction error) model is:

$$SPE = \sum_{i=1}^{m} (y_{new,i} - \hat{y}_{new,i})^2 \tag{2}$$

where $\hat{y}_{new,i}$ and $y_{new,i}$ are the reconstructed value of the ith process variable from the reference PCA model and the measured value of this variable respectively. SPE is also refereed to as Q statistic or distance to the model. It represents the squared perpendicular distance of a new multivariate observation from the projection space. The procedure for developing the tomography based MSPC is described below.

The first step is the collection of normal operational data. The data is then pre-processed using two scaling methods. Since the tomography measurements (in this case the measurements are a set of voltages between different electrodes) are dependent on the conductivity and concentrations of fluids in the pipeline, it is necessary to remove the effect of these parameters on the measurements so that the developed MSPC model can be generalised and is independent of the concentrations and conductivity of the fluids.The first scaling method, called scaling over a frame, is applied to each frame of the collected data:

$$V_{new}(i,j) = \frac{V_{raw}(i,j) - V_{\min}(i)}{V_{\max}(i) - V_{\min}(i)} \quad (\text{j} = 1,2,3,\ldots,104) \tag{3}$$

where $V_{new}(i, j)$ and $V_{raw}(i, j)$ are the j[th] pre-processed and raw voltage measurement on the i[th] frame in the collected data, $V_{\min}(i)$ and $V_{\max}(i)$ the minimum and maximum raw measurement values on the i[th] frame, respectively. The second method for pre-processing the data is auto scaling to make each measurement have the same variance.

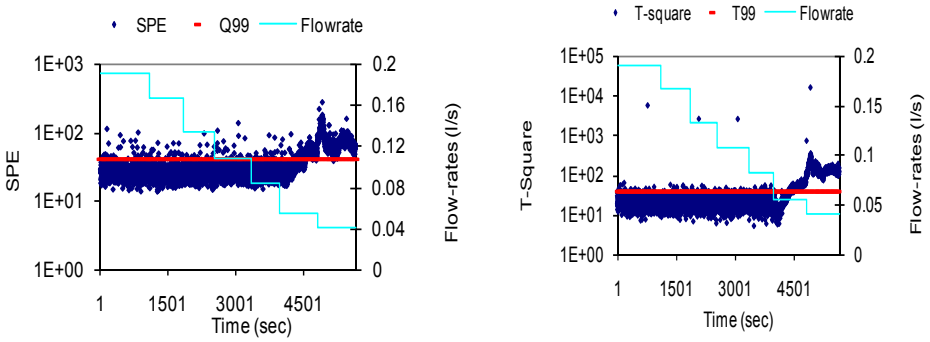
In the third step, the pre-processed data is analysed using PCA. A smaller number of latent variables, i.e. principal components (PCs) are derived. This transformation reflects the correlations between the tomography measurements at normal mixing states. The T^2 and SPE control limits at different confidence levels are derived.

Finally, the MSPC model is applied to new tomography data for mixing condition monitoring. It needs to point out that all new measurements were also pre-processed using the two scaling techniques described above. If the SPE is over the control limits, it could be an indication that the current mixing state had gone out of or was very different from the predefined normal mixing states. If the SPE is in the control limit and T^2 out of the limit, this could be an indication of bad mixing. After poor mixing is detected, contribution plots can be applied to analyse the key measurements contributed to the abnormal mixing.

4. Results and Discussions

4.1. Sunflower-oil – water emulsion

The ERT voltage measurements of sunflower oil – water emulsion, at flowrates 0.166, 0.133 and 0.108 l/s of concentrations from 33% to 5%, were used to develop the PCA based MSPC model. There are 23768 frames in total. The voltage data were firstly pre-processed using the scaling methods described in the previous section. Then PCA was applied to the pre-processed data to develop the MSPC models. 20 PCs were selected which explained about 73% variance of the data. Based on the selected PCs, control limits of two statistics SPE and T^2 at different confidence levels were calculated. The MSPC models were applied to monitor the mixing performance at different operating conditions. Figure 2 shows an example of emulsion at 20% concentration and various flowrates. It can be seen that significant increases of both SPE and T^2 occurred after the flowrate was reduced to 0.055 and 0.041 l/s. This indicates that the mixing became poor at these flowrates. The reconstructed conductivity tomography during the flowrate change, shown in the Figure 3, confirmed the deteriorating mixing at reduced flowrates.



(a) (b)

Figure 2. SPE and T^2 Monitoring charts for the sunflower oil – water emulsion system. The concentration is fixed at 20%, while the flowrate is decreased. Q99 – control limit is 99%.

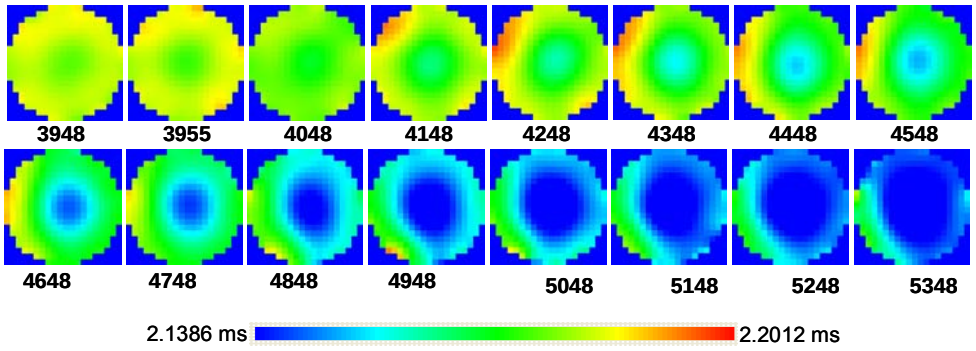


Figure 3. Reconstructed conductivity tomography images for the sunflower oil – water emulsion system. Concentration is fixed at 20%, but flowrate changes.

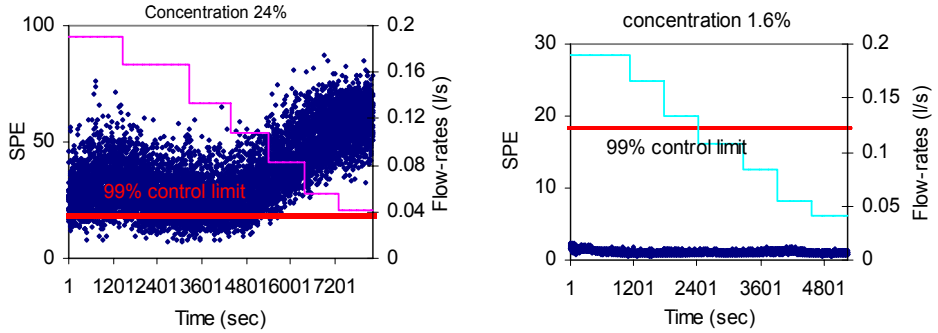


Figure 4. SPE charts for silica suspension at solid volume concentrations of 24% and 1.6%.

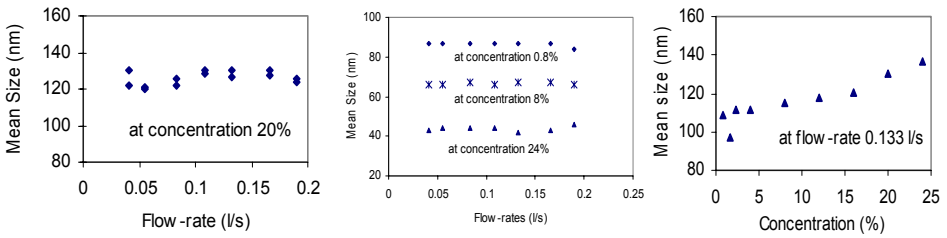
4.2. Mixing performance of SnowTex-ZL silica suspension

MSPC models were developed based on the tomography measurements at flowrates of 0.166, 0.133 and 0.108 l/s, for all concentrations studied. Twenty principal components were chosen, capturing about 88% of the variance.

The MSPC models were applied to monitor the suspension at different concentrations and flowrates. Figure 4 shows the SPE control charts at concentrations 24% and 1.6%, which indicates that the higher the concentration, the poorer the mixing.

4.3. Particle size measurements by on-line Zetasizer and Ultrasizer

Particle size distribution was measured using DLS and USS for both materials at different concentrations and flowrates. For the silica suspension, both DLS and USS gave consistent results when the flowrate was changed while concentration remained constant, as indicated by figures 5(a) and 5(b). However with the increase of solid concentration, the mean size measured by the DLS is increased, while the mean size measured by the USS is decreased, as shown in figure 5 (b) and (c), reflecting the effects of solid concentration on the measurement results.



(a) (b) (c)

Figure 5. Particle mean size measurements. (a) DLS result at a fixed concentration; (b) USS result at different flowrates and concentrations, and (c) DLS result at a fixed flowrate but varied solid concentrations.

5. Final Remarks

Nanoparticle processing in industry faces major challenges in process scale-up and in maintaining consistency and reproducibility in product quality. Online characterization of product quality and process conditions can help industry effectively address the challenges, because based on real-time measurements, process control and quality assurance strategies can be developed. The work has investigated the use of dynamic light scattering (DLS), ultrasound spectroscopy (USS) and electrical resistance tomography (ERT) for online characterization of processes of nanoparticle suspensions and emulsions. The system provides real-time information of particle size distribution (PSD), zeta potential, conductivity and solid concentration and visualizes the mixing behaviour. In addition, ERT provides solid concentration data that is needed by DLS and USS for PSD characterization. This study proposed the use of multivariate statistical process control, or MSPC for making use of the data for process monitoring and control. Two cases studies proved the feasibility of the proposed approach. The combined sensors and the MSPC system together provide a platform for on-line characterization of emulsions and nanoparticle slurries.

6. Acknowledgement

The authors would like to thank the financial support form the Technology Strategy Board (TP/SC/6/I/10097) and UK Engineering and Physical Sciences Research Council (Grant reference: EP/E040624/1).

References

1. Industrial Tomography Systems. http://www.itoms.com. Accessed February 2009
2. Malvern Instruments Ltd. http://www.malverns.com. Accessed Februatry 2009.
3. MacGregor J. F., T. Kourti, Statistical process control of multivariate processes. Control Engineering Practice, 1995, **3:** 403-414.
4. Albazzaz H., X. Z. Wang, Statistical process control charts for batch operations based on independent component analysis. Industrial & Engineering Chemistry Research, 2004, **43**: 6731-6741.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Control of Fed-Batch Yeast Cultivation Using a Capacitance Sensor

Giann B. Reis, Antonio C. L. Horta, Teresa C. Zangirolami, Roberto C. Giordano and Antonio J. G. Cruz

UFSCar - Department of Chemical Engineering - Federal University of São Carlos.

LaDABio – Laboratory for Development and Automation of Bioprocesses.

Rod. Washington Luis (SP 310), Km 235, CEP 13565-905, São Carlos - SP, Brazil.

Abstract

Saccharomyces cerevisiae biomass is still one of the most important biotechnological products in the world, mainly in the form of baker's yeast. Costs of substrates have an important position in the overall economics of baker's yeast production, which is usually run in fed-batch bioreactors. In this context, the maximization of biomass yields, aiming at the highest achievable volumetric productivity is a driving force for improving the control of the process. Classical model-based control strategies for this process, however, face difficulties due to the inherent variability of this system: microorganisms have a complex growth dynamics, lumped in very simplified growth models; raw materials are variable and may be not traceable; the system response depends on non-controlled previous process stages, such as strain selection and inoculum preparation. Hence, the possibility of using feedback information from *in situ* sensors for re-tuning control parameters is an important issue to ensure sub-optimal performances, at least. A capacitance sensor is a device that can monitor cell concentration on-line. In fed-batch cultivations, the feed rate and correlated inputs can be controlled by systems coupled with softsensors that infer the state of the system from on-line measurements of primary variables. Several estimation techniques have been proposed in the literature, and among them the gas balance technique is widely used. The specific respiration rate (qO2), the specific carbon dioxide production rate (qCO2), and the respiratory quotient (RQ) are the main variables determined from the gas balance. Values of RQ for complete oxidation of some carbon sources to carbon dioxide and water are found in the literature. For the baker's yeast production process growing in glucose, a value of RQ close to unity indicates the preponderance of the aerobic route. The capacitance sensor, after calibration, may provide information about cellular growth and viability: the capacitance of the medium is linearly proportional to viable cell concentration. In this work, the signal of capacitance sensor and the RQ value are coupled to a fuzzy algorithm in order to control the glucose feed rate during baker's yeast aerobic cultivation.

Keywords: biomass, process control, *Saccharomyces cerevisiae*, soft-sensor.

1. Introduction

One of the most important variables in bioprocess control is the viable biomass concentration, which can be monitored by capacitance probe (Kiviharju et al. 2007; Xiong, et al., 2008). Until today, the feed rate in fed-batch fermentation was controlled based only in gas balance technique and in simulations results. Some works, as the one

of Xiong et al. (2008), demonstrate that the capacitance probe is a powerful tool for monitoring *Saccharomyces cerevisiae* fermentation, providing a more reliable process control when it is employed.

The capacitance probe functioning principles are shortly described in the followings. At frequencies between 0.1 and 10 MHz, the undamaged cell membrane, *i.e.* a viable cell, acts as an electric capacitor. The dielectric permittivity (ε in pF/cm) is proportional to the viable biomass concentration (Neves et al., 2000). The permittivity and the conductivity (σ in mS/cm) values can be calculated by the Equation 1 and 2 respectively. K (cm^{-1}) is the geometric constant of the electrodes (Markx and Davey, 1999).

$$\varepsilon = C \cdot K \quad (1)$$

$$\sigma = G \cdot K \quad (2)$$

At low frequency, the oscillation of the electric field is slow enough to completely polarize the cells and the medium. When the frequency of the electric field is slow enough, cell polarization reaches a plateau which corresponds to the maximal polarization. At high frequency, the field oscillates quickly and the intracellular charges (ions) do not have time to reach the plasma membrane before the change in the field polarity. Therefore the higher the frequency is, the lower the cell polarization will be. At high frequency, only the rotation of charged molecules induces a permittivity signal. Thus the permittivity value at 10 MHz gives a measurement of the medium's influence only. The difference of the permittivity measurements at low and high frequency is proportional to viable cellular concentration, as illustrated at Figure 1.

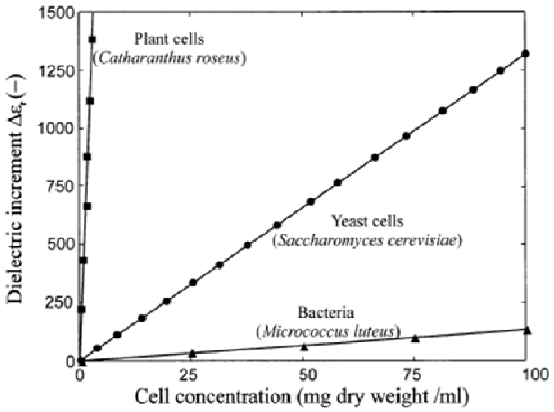


Figure 1. Permittivity (dielectric) increment of the β-dispersion as a function of biomass concentration, for a number of cell types (Markx and Davey, 1999).

2. Materials and Methods

2.1. Experiment

Experiments were carried out in a 6 L stirred tank bioreactor (5 liters working volume), filled with 4.0 liters of medium solution (glucose, 1.2 g/l; KH_2PO_4, 5.0 g/l; $MgSO_4 \cdot 7H_2O$, 0.5 g/l; $(NH_4)_2SO_4$, 4.5 g/l, yeast extract, 3.0 g/l), in fed-batch mode. The inoculum was inserted at concentration of 1 g/l and the cultivation process takes place. Stirrer speed was set to 350 rpm. The level of dissolved oxygen (DO) was maintained

over 45% of saturation (Kasperski and Miskiewicz, 2008). A schematic representation of the experimental set-up used in this study is shown in Figure 2.

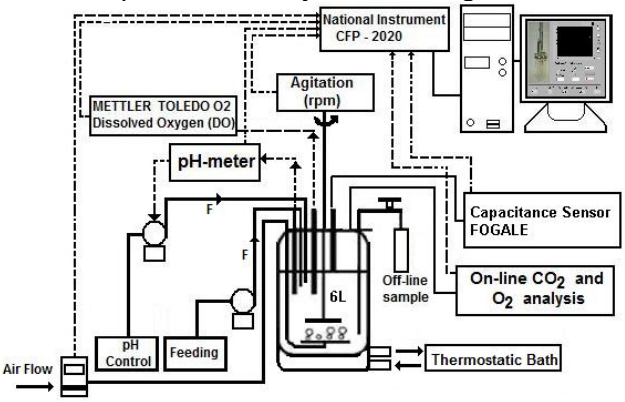


Figure 2- Schematic representation of the experimental system used in this study.

2.2. Off line variables

Dry cell weight (DCW, g/l) was measured by gravimetric method (samples were centrifuged at 8000×g for 15 min followed by drying at 100°C for 24 h). Cell viability was assessed through the Methylene Blue staining method, followed by counting of stained and non-stained cells in Neubauer chamber (Antonini, 2004).
Glucose and ethanol concentrations were measured by HPLC.

2.3. On-line variables

Cell concentration (viable cells) was measured by Biomass System – Viable Cell Density Monitoring (Fogale Nanotech). Equation 3 presents a calibration curve relating viable cell concentration (C_x) and permittivity (ε).

$$C_x(t) = 0.029 \cdot \varepsilon(t) - 3.23 \quad (3)$$

pH was monitored by pHmeter (GLI PRO P3, Hach, USA). The exhaust gas (oxygen and carbon dioxide molar fraction) was measured by the gas analyzer (SICK-MAIHAK S710, German). The dissolved oxygen concentration was monitored (O_2 4050, Mettler Toledo, Switzerland).

2.3.1. The control algorithm

In this study, a Mamdani configuration for the fuzzy algorithm was used. A basic fuzzy logic control structure consists of a *Fuzzifier* to transform measured data into suitable fuzzy sets; an *Inference engine-fuzzy rule base* to evaluate assigned fuzzy rules and has the capability of simulating human decision making by performing approximate reasoning; a *Defuzzifier* to convert fuzzy sets into a single number.
There are several methods for extracting a crisp value from a dispersion of experimental data to build up a fuzzy set. The center-of-gravity method was adopted in this work for the *defuzzification* step. It should be stressed that the decision on which *defuzzification* method should be applied relies mostly on heuristics and on the experience of the programmer. More details about fuzzy logic can be found in Zadeh (1965), Hisbullah et al. (2003), Tahera (2008) and Karakuzu et al. (2006).
The first step in designing a fuzzy control is to identify fuzzy input and output monitoring variables. During *S. cerevisiae* cell growth, a strong relationship was observed between experimental permittivity and carbon dioxide concentration in the exhaust gas. The fuzzy

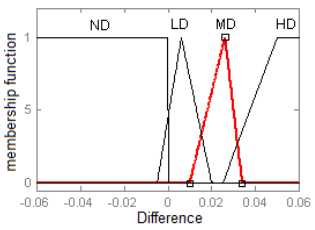
logic algorithm presented here has 16 rules (Table 1). It was written in MatLab® and implemented in a LabVIEW® environment. Data strings were exchanged in real time between the MatLab program and the data acquisition system (programmed in LabVIEW).
The variables chosen as input to the fuzzy algorithm were the Respiratory Quotient (RQ) and the error (e), a variable defined as the difference between the reference cellular growth coefficient (μ_{REF}) and its on-line estimated value (μ). The reference value had to be modified during cultivation because the growth behaved differently as the experiment progressed. Its value ranged from 0.06 to 0.1. The output variable was glucose feed flow rate powered by peristaltic pump (Ismatec, Switzerland). The rules of fuzzy logic algorithm included the two input variables, each one associated to four fuzzy sets: a) antecedent: input RQ – Low (RQL), Good (RQG), High (RQH) and Very High (RQVH); input error (e) – Negative Difference (ND), Low Difference (LD), Medium Difference (MD) and High Difference (HD); b) consequent: glucose feed flow rate – Null Flow (NF), Low Flow (LF), Medium Flow (MF) and High Flow (Table 1).

Table 1 – Fuzzy set rules of input and output variables.

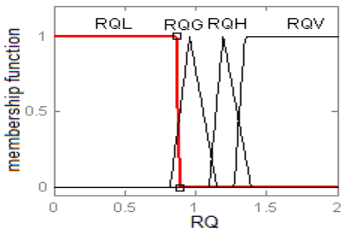
Rule	IF (antecedent)	THEN (consequent)
1	RQL **AND** ND	LF
2	RQG **AND** ND	LF
3	RQH **AND** ND	LF
4	RQV **AND** ND	NF
5	RQL **AND** LD	MF
6	RQL **AND** MD	MF
7	RQL **AND** HD	HF
8	RQG **AND** LD	LF
9	RQG **AND** MD	LF
10	RQG **AND** HD	MF
11	RQH **AND** LD	NF
12	RQH **AND** MD	NF
13	RQH **AND** HD	NF
14	RQV **AND** MD	NF
15	RQV **AND** ND	NF
16	RQV **AND** HD	NF

3. Results and Discussion

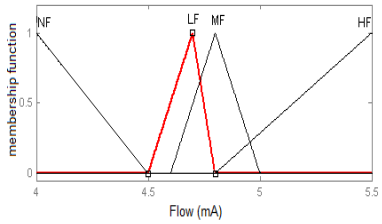
The correct choice of membership function plays an essential role in achieving a successful design of fuzzy logic controller. Unfortunately, this is not deterministic and has no unique solution. Furthermore, this choice is based on subjective decision criteria and relies heavily on time consuming trial and error. Different membership functions were tested during the fuzzy parameters tuning (*Gaussian*, *Bell*, *Triangular* and *Trapezoidal*) and the best result was obtained using triangular and trapezoidal membership functions, as illustrated in Figure 3. Figure 4 shows the cell, glucose and ethanol concentrations during the experiment.
The exponential growth phase started shortly after the inoculation. Glucose was exhausted within 1 hour of cultivation. Figure 5 shows the behavior of the manipulated variable (glucose feed flow rate) and controlled variable as well (RQ).
At the start of the feed (between 1 and 4.5h), the parameters of the fuzzy controller were changed by the operator, leading to sharp fluctuations of the feeding pump. After the right set of parameters were adjusted, the behavior of the glucose feed flow rate stabilized, as can be seen in Figure 5.



(a) Input variable – error (μ_{REF}- μ).



(b) Input variable- RQ.



(c) Output variable - Glucose feed flow rate.

Figure 3- Final membership functions of antecedent after parameters tuning: (a) input variable (μ_{REF} - μ); (b) input variable RQ; (c) consequent part – Glucose feed flow rate.

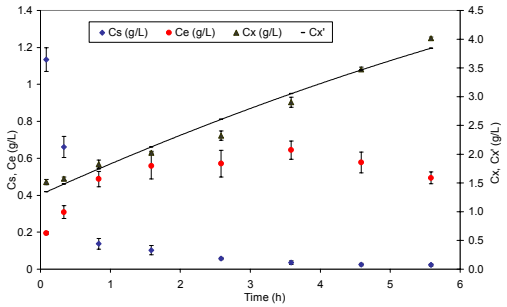


Figure 4 - Glucose concentration (Cs), ethanol concentration (Ce), biomass concentration measured by dry weight (Cx) mass, viable biomass concentration provided by capacitance probe (Cx').

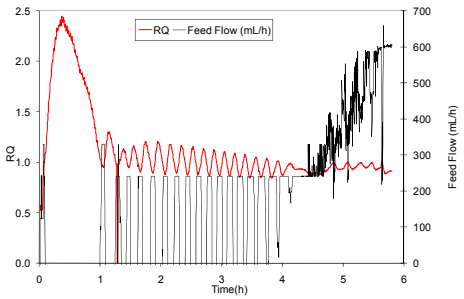


Figure 5 - Behavior of the manipulated variable (glucose feed flow rate) and controlled variable (RQ).

4. Conclusion

The fuzzy control algorithm using the permittivity and the RQ values as input variables was successfully implemented. The glucose concentration into the medium was kept at low values (close to zero) during the whole assay. As a consequence, ethanol concentration was kept low. Cellular growth rate was kept constant until the end of the experiment. The cell viability remained at 98% throughout the experiment. This information allows to correlate directly the probe signal to the cell concentration. The cell productivity ($\Delta Cx/\Delta t \cong 0.45 g/l/h$) reached was considered high if compared to others reported at the litterature, as $\Delta Cx/\Delta t \cong 0.41 g/l/h$ obtained by Oliveira (2006). The results demonstrate that the use of the capacitance sensor to reach the optimun control of system is viable, and the fuzzy control was able to control glucose feed flow rate efficiently using this probe signal.

Acknowledgement

FAPESP, CNPq and CAPES for the financial support.

References

Antonini, S.R.C., 2004, Métodos de análises e monitoramento microbiológico em laboratórios de destilaria. Dept. Tecnologia industrial e Sócio-Economia Rural Centro de Ciências Agrárias – Campus Araras.

Hisbullah, M.A.H., Ramachandran, K.B., 2003, Design of a fuzzy logic controller for regulating substrate feed to fed-batch fermentation. *Institution of Chemical Engineers*, 81,138-146.

Karakuzu, C., Turker, M., Özurks., 2006, Modelling, on-line state estimation and fuzzy control of production scale fed-batch baker's yeast fermentation. *Control Engineering Practice*, 14, 959-974.

Kiviharju, K.; Salonen, K.; Moilanen, U.; Meskanen, E.; Leisola, M.; Erikäinen, T., 2007, On-line biomass measurements in bioreactor cultivations: comparison study of two on-line probes. J Ind Microbiol Biotechnol, 34:561–566.

Kasperski, A., Miskiewicz, T., 2008, Optimization of pulsed feeding in a Baker's yeast process with dissolved oxygen concentration as a control parameter. *Biochemical Engineering Journal*. (*In press).*

Markx, G.H.; Davey, C.L., 1999, The dielectric properties of biological cells at radiofrequencies: Applications in biotechnology. Enzyme and Microbial Technology, 25: 161–171.

Neves, A.A., Pereira, D.A., Vieira, L.M., Menezes, J.C., 2000, Real time monitoring biomass concentration in *Streptomyces clavuligerus* with industrial media using a capacitance probe. Journal of Biotechnology, 84: 45-52.

Oliveira, J. A., 2006, Modelagem Matemática do Processo de Leveduras de Panificação – Um Estudo de Caso, dissertation of masters degree, *Federal University of São Carlos – SP- Brazil.*

Tahera, K., Ibrain, R.N., Lochert, P.B., 2008, A fuzzy logic approach for dealing with qualitative quality characteristics of a process. Expert System with Applications, 34, 2630-2638.

Xiong, Z.Q.; Guo, M.J.; Guo, Y.X.; Chu, J.; Zhuang, Y.P.; Zhang, S.L., 2008, Real-Time Viable-Cell Mass Monitoring in High-Cell-Density Fed-Batch Glutathione Fermentation by *Saccharomyces cerevisiae* T65 in Industrial Complex Medium. Journal of Bioscience and bioengineering, Vol. 105, No. 4, 409–413.

Zadeh, L. A., 1965, Fuzzy sets. *Information and Control*, 8, 338–353.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Monitoring of Vinyl Chloride Suspension Polymerization Using NIRS. 2. Proposition of a Scheme to Control Morphological Properties of PVC

João Miguel de Faria Jr.,[a,b] Fabricio Machado,[b,‡] Enrique Luis Lima,[b] José Carlos Pinto[b]

[a]*Braskem S.A., Rua Hidrogênio, 3342, Pólo Petroquímico, CEP: 42810-000, Camaçari, BA, Brazil*
[b]*Programa de Engenharia Química / COPPE, Universidade Federal do Rio de Janeiro, Cidade Universitária, CP 68502, Rio de Janeiro, 21945-970, RJ, Brazil*

Abstract

This article introduces the use of in-line and *in situ* monitoring of the dynamic evolution of PVC morphological properties in suspension polymerization reactions based on near infrared spectroscopy (NIRS) intended for control applications. It is shown for the first time that it is possible to follow the dynamic evolution of morphological properties (such as the bulk density, BD, and the cold plasticizer absorption, CPA) in real time. The obtained results indicate that closed-loop control schemes can be properly implemented to control end-use resin properties, which depend upon the morphology of PVC polymer particles. Suspension polymerization of vinyl chloride performed in bench and pilot scale showed that process variables, such as the agitation speed and the amount of suspending agents) can be successfully manipulated in order to control the morphological parameters of PVC resin. Based on this information, a control scheme for these parameters is proposed.

Keywords: Near Infrared Spectroscopy (NIRS), MVC Suspension Polymerization, Dynamic Evolution of Morphological Properties, Control of Polymer Particle Morphology.

1. Introduction

It is well known that profits and plant operation quality may be significantly increased through the implementation of plant automation and the adoption of advanced control strategies. This is particularly true in the polymer industry, as polymer materials are characterized as “product-by-process” materials, which means that the history of the reaction is of great importance to define the final product properties and quality. In the polymer industry it is generally very difficult to minimize specifications drifts caused by process disturbances and uncertainties through blending of different batch products and/or additional separation steps, as usually performed in other fields (Vieira et al., 2002). Specifically, in the case of poly(vinyl chloride) (PVC) resin, a sophisticated arrangement of particles morphology appears as a special feature of PVC resins

‡ Current Address: Instituto de Química, Universidade de Brasília, *Campus* Universitário Darcy Ribeiro, CP 04478, 70910-900, Brasília, DF, Brazil, E-mail: fmachado@unb.br

obtained from suspension polymerization process. Very poor solubility of PVC in the vinyl chloride monomer (VCM) causes precipitation of PVC in the VCM droplet at early stages of polymerization, proceeding with the generation of complicated particle structures. This PVC particle structure is very important for understanding of process capability and quantities of produced product (Saeki and Emura, 2002).

The production of polymers with pre-specified properties has placed great emphasis on the development of accurate and robust instruments for the on-line monitoring of polymerization reactions. The complex nature of polymerization systems is one of the main contributing factors to the difficulty of on-line measurements. The presence of long measurement delays and poor reliability are two common problems associated with the on-line characterization of the polymer quality. Even the off-line measurement of many polymer properties is a difficult task, requiring sophisticated and time-consuming analytical techniques. In addition, the development of on-line polymer sensors requires a multidisciplinary effort, including mathematical modeling and data processing, improved knowledge and understanding of the process, reactor design, and modern advanced control and instrumentation engineering (Machado and Bolzan, 1998; Kammona et al., 1999).

An important problem is that, generally, in batch processes, the polymer quality can be assessed only a certain period of time after the end of the reaction. Therefore, many times, the quality of raw material changes (e.g. the quality of monomer) during this period and probably these changes, that could probably affect the resin quality parameters, may not be detected immediately. Consequently, off-specification products can be produced. In addition, as the process will be in a different state, adjustments in the formulation performed based on this delayed information would not assure the solution of polymer quality problems. With the methodology proposed here, it will be possible to track the dynamic evolution of PVC morphological parameters in real time, which means that the long delay times involved in the suspension polymerization of PVC will be properly eliminated, making possible the use of control actions related to the real state of the process.

2. Experimental Procedure

A standard recipe consisting of poly(vinyl alcohol), initiator, water, monomer and additives was fed into the reactor to perform the polymerization reactions. The detailed recipe cannot be presented for proprietary reasons. Detailed description of reactor apparatus is presented by de Faria Jr. et al. (2009), and the reader is referred to this publication for a more detailed description of the protocol.

The studies were divided in three groups. The first one was performed to obtain the dynamic evolution of bulk density (BD) and cold plasticizer absorption (CPA). The second one was performed to verify if changes in the agitation speed would lead to modifications of the BD, CPA and average diameter of final PVC particles. The tests were performed in a bench stainless-steel reactor with a volume of 0.01 m^3 and maximum working pressure and temperature of 20 bar and 250ºC, respectively. A thermocryostatic bath was used to control the reaction temperature. The agitation speed could be varied in the range between 0 and 2400 rpm and a Pt 100 thermocouple was used to monitor and control the reactor temperature.

After 20 minutes of reaction, the normalized agitation speed was changed in the range from 0.40 to 1.00. In all batches the normalized agitation speed at the beginning of the reaction was equal to 0.6. The runs were replicated to analyze the process

reproducibility. For each run, BD, CPA and PSD of the final PVC powder were obtained.

The third group of studies was performed by dosing suspending agent into the reaction medium in order to verify if changes of the suspending agent concentration would lead to modifications of the BD, CPA and average diameter of final PVC particles, as in the previous. Tests were performed in a pilot stainless-steel reactor with volume of 0.1 m^3. A cascade control mode (PID-PI) combined with a Distributed Control System was used to control the flow rate of hot water in the jacket. After 20 minutes of reaction, the normalized amount of suspending agent was dosed into the reactor (0.58, 0.75, 0.83, 1.00). Additions were performed in different batches, using a solution of dispersant agent in water. The solution was kept inside a process vessel that was isolated from the reaction medium by a handle valve. After 20 minutes of reaction, the valve was opened, allowing for the solution to flow into the reactor.

3. Results and Discussion

The correlations obtained to predict BD and CPA were excellent and able to represent qualitatively the dynamic evolution of both variables during the polymerization reactions, as shown in Figure 1. The quality of prediction for both BD and CPA, as obtained with the calibration model, can also be regarded as very good. It is very important to emphasize that calibration was performed only with samples obtained at the end of the batch, because quality parameters can only be measured when the polymerization is finished. In spite of that, the dynamic evolution of BD and CPA are predictable and follow very well defined trajectories, showing the robustness of the proposed monitoring scheme (in Figure 1, only the last points of the BD and CPA dynamic trajectories are real experimental points).

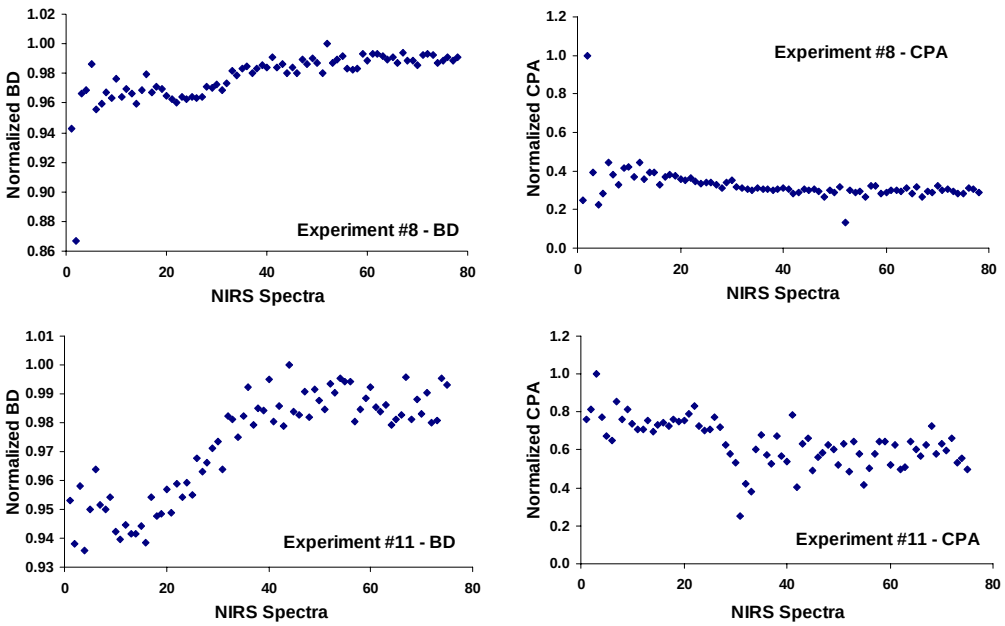


Figure 1 - Dynamic Evolution of BD and CPA during the VCM Polymerization, as predicted by the PLS model.

Figure 1 shows that *in situ* and on-line monitoring of morphological PVC properties using NIRS can be employed successfully in MVC suspension polymerization process, as dynamic trajectories are smooth and well-behaved, indicating that model predictions can be useful along the whole reaction course and used for monitoring and control purposes.

The average diameter of PVC particles obtained at the end of the polymerization reactions was analyzed by screening and light scattering, showing good agreement, as shown in Figure 2. This figure illustrates the effect of agitation speed on the average particle size of PVC. Based on these results, one can conclude that it is possible to vary the agitation speed in order to control the particle size distribution and CPA of PVC particles. As shown, the sudden increase of the agitation speed during the polymerization leads to the diminishment of the average droplet diameter and consequently of the final PVC particle (Dp). After attainment of a minimum Dp value, the overall superficial particle area increases so much that PVC particles agglomerate to reduce the interfacial energy of the system. If stirring continues to increase, the break-up rates due to the high shear forces predominates and Dp decreases again (see in Figure 2 the value of normalized agitation speed equal to 1).

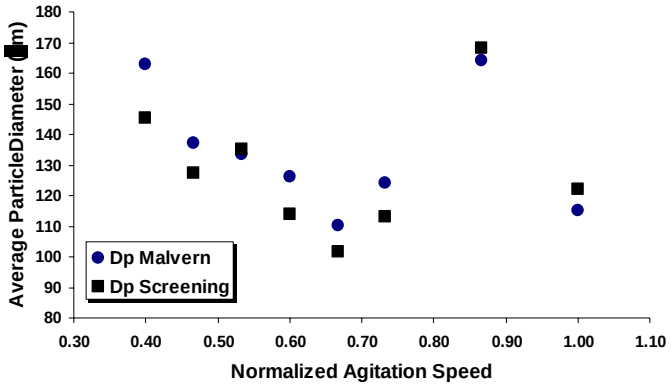


Figure 2 - PVC Particle Diameter Obtained by Screening and Light Scattering.

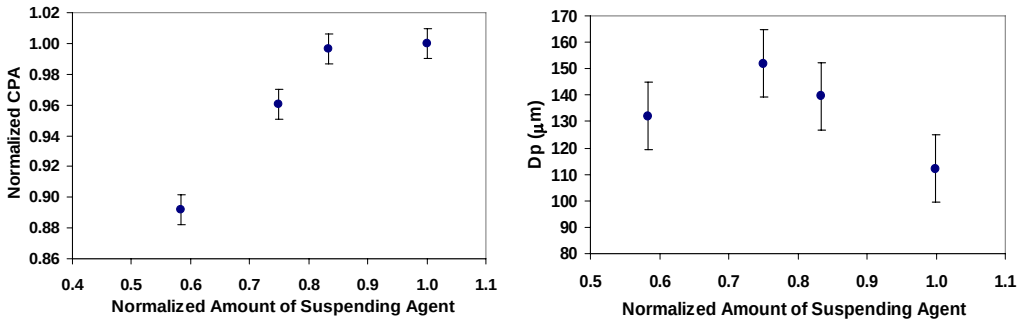


Figure 3 - Effect of Dosing of Suspending Agent on the Bulk Density and Cold Plasticizer Absorption.

According to the results obtained from the third group of experiments, one can conclude that it is possible to dose suspending agent during the reaction to control the particle size distribution, BD and CPA of the final PVC resins, since polymer particles with different diameters, BD and CPA can be obtained when the dispersant agent is dosed

into the reaction medium during the run, as shown in Figures 3. It is important to observe that these experiments were performed in a pilot reactor, where the reactor volume was bigger than the bench reactor volume where the experiments of group two were performed. These observations are in very good agreement with independent observations in the industrial process (although not shown here).
Based on the results presented in Figures 1 to 3, a control scheme is proposed for monitoring and control of important morphological properties of PVC obtained in suspension polymerization process. The scheme takes into account the acquisition of signal available at polymerization site plus the signal from de NIR probe. Among the signals of practical interest available in the polymerization system, one may cite the agitation speed (that provides information about the evolution of the medium viscosity), the reactor pressure (that provides information about the monomer conversion), the reactor temperature (that provides information about the reaction rate, the medium viscosity and the relative volatility of VCM), the inlet and outlet jacket temperature (that allows for implementation of calorimetric schemes for evaluation of monomer conversion, heat transfer coefficients and reaction rates). Figure 4 shows the proposed scheme for the monitoring and control of PVC morphology and the reaction state in real time.

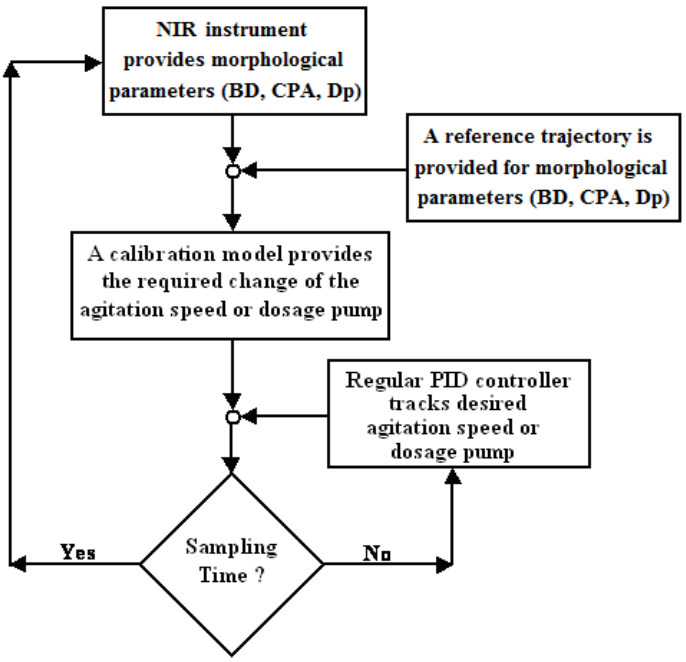


Figure 4 - Scheme for Monitoring and Real Time Control of the Morphological Properties of PVC Particle.

The implementation of the proposed in-line monitoring and control scheme will permit to follow, in real time, the evolution of conversion, of morphological features of PVC particles and the quality of the heat exchange. According to this scenario, the measured trajectories can be compared with ideal trajectories designed off line for control purposes. These ideal trajectories can result from optimization of any variable of interest (for example, minimum reaction time) or from a standard run that can be regarded as a reference one. In the second case, the NIRS signal obtained along a

standard run can be transformed into average size, CPA and/or BD (obtained from calibration model) and used as setpoints for others batch runs. In this case, different setpoints can be used to produce different PVC grades. According to this scheme, the process model can also be used to estimate in real time the monomer conversion and the quality of heat exchange (with the help of a simple balance model).
Reactor temperature should not be manipulated for control of morphological characteristics of PVC particles because it affects the K-value of the resin very significantly. When it is not possible to manipulate the initiator concentration along the polymerization batch, the reaction time can be manipulated through selection of initiator cocktails in the initial charge. Finally, the morphological variables (BD, CPA and average particle size, predicted by NIRS signal) can be controlled through proper manipulation of agitation speed and\or dosing of suspending agent. Thus, whenever a deviation in the variable of interest is detected (for example, Dp), based on the obtained calibration model, a control action must be sent to the agitation speed and/or dosing pump, allowing for satisfaction of the desired trajectories (Dpset). For example, if Dp is bigger than the Dpset, the stirring can be increased and/or the dosing pump can feed an additional amount of suspending agent into the reactor vessel. On the other hand, if Dp is lower than Dpset, the system agitation can be reduced.

4. Conclusions

It was shown for the first time that it is possible to follow the dynamic evolution of morphological properties (such as the bulk density, BD, and the cold plasticizer absorption, CPA) in real time during the PVC suspension polymerization. The obtained results indicated that closed-loop control schemes can be properly implemented to control end-use resin properties, which depend upon the morphology of PVC polymer particles. Suspension polymerization of vinyl chloride performed in bench and pilot scale showed that process variables, such as the agitation speed and the amount of suspending agents) can be successfully manipulated in order to control the morphological parameters of PVC resin. Based on this information, a control scheme for these parameters is proposed.

Acknowledgements

The authors thank Braskem S.A. and FINEP for supplying technical resources required for this work.

References

J. M. de Faria Jr., F. Machado, E. L. Lima, J. C. Pinto, 2009, Monitoring of Vinyl Chloride Suspension Polymerization Using NIRS. 1. Morphological Properties Prediction, Computer-Aided Chemical Engineering (Accepted).

R. A. F. Machado, A. Bolzan, 1998, Control of batch suspension polymerization reactor, Chemical Engineering Science 70, 1-8.

R. A. M. Vieira, C. Sayer, E. L. Lima, J. C. Pinto, 2002, Closed-Loop Composition and Molecular Weight Control of a Copolymer Latex Using Near-Infrared Spectroscopy, Industrial & Engineering Chemistry Research, 41, 12, 2915-2930.

O. Kammona, E. G. Chatzi, C. Kiparissides, 1999, Recent Developments in Hardware Sensors for the On-Line Monitoring of Polymerization Reactions, Journal of Macromolecular Science - Reviews in Macromolecular Chemistry and and Physics, C39, 1, 57-134.

Y. Saeki, T. Emura, 2002, Technical Progresses for PVC Production, Progress in Polymer Science, 27, 2055-2131.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modified Unscented Recursive Nonlinear Dynamic Data Reconciliation for Constrained State Estimation

Sachin C. Kadu,[ab] Mani Bhushan,[a•] R.D. Gudi,[a] Kallol Roy[c]

[a]*Department of Chemical Engineering, IIT Bombay, Mumbai 400076, India*
[b]*Reactor Projects Division, BARC, Mumbai 400085, India*
[c]*Research Reactor Maintenance Division, BARC, Mumbai 400085, India*

Abstract

In state estimation problems, often, the true states satisfy certain constraints that need to be incorporated during the estimation procedure. Amongst various constrained nonlinear state estimation algorithms proposed in literature, the unscented recursive nonlinear dynamic data reconciliation (URNDDR) proposed by Vachhani et al. (2006) seems to be promising since it is able to incorporate constraints while maintaining the recursive nature of estimation. In this article, we propose a modified URNDDR algorithm that gives superior performance compared to basic URNDDR. The improvements are obtained via better constraint handling and are demonstrated via a representative case study. **Keywords**: URNDDR, Kalman Filter.

1. Introduction

Kalman filter (KF) based state estimation techniques are widely employed to produce accurate estimates of the true states in the presence of noisy sensor information and an uncertain process model. For most of the real world problems of interest, the system dynamics and observation equations are nonlinear (NL) in nature. Most of the NL estimators proposed in literature are extensions of KF. The most widely used estimator for nonlinear systems is extended KF (EKF). The EKF involves simple linearization of the nonlinear terms in the model so that the traditional linear KF equations can be applied. However, the implementation of EKF requires the Jacobian matrices which may be expensive to compute and are prone to numerical errors, especially for higher order systems. Further, the linear approximation of the system may introduce significant errors in the state estimates, which may even lead these estimates to diverge over time. To overcome these difficulties, Julier et al. (1995) proposed the unscented KF (UKF). It yields performance equivalent to the KF for linear systems, and can be used for nonlinear systems without the linearization steps, which are required by the EKF. The fundamental component of UKF is the unscented transformation (UT), which uses a set of deterministically chosen weighted points (called as sigma points) to parameterize the mean and covariance of a probability distribution. However, the UKF suffers from two drawbacks: (i) the updated state estimates are obtained using a linear estimator (Vachhani et al., 2006), (ii) similar to EKF, UKF also does not incorporate constraints on states during estimation. This can lead to meaningless estimates (such as negative concentrations, etc.) in several applications. The recursive nonlinear dynamic data

• Corresponding Author, Email: mbhushan@iitb.ac.in; Tel:+91-22-25767214; Fax:+91-22-25726895

reconciliation (RNDDR) method developed by Vachhani et al. (2005) to incorporate constraints on state estimates, also does not account for the effect of nonlinearity in state and measurement equations as well as the effect of constraints while calculating covariance matrices of the error in the predicted and estimated states. The subsequent work by Vachhani et al. (2006) overcame this limitation by combining merits of RNDDR and UKF to propose the unscented recursive nonlinear dynamic data reconciliation (URNDDR). A rigorous analytic method of incorporating state equality constraints in the KF by projecting the unconstrained KF estimates onto the boundary of the feasible region at each time step was developed by Simon and Chia (2002). But this method is applicable only for equality constraints and requires Jacobian calculation. Recently, Kandepu et al. (2008) proposed a simple procedure of projecting transformed infeasible sigma points onto the boundary of feasible region in order to include state constraints in the UKF algorithm. However, this projection idea was not applicable to the case of nonlinear equality constraints. Due to various limitations of alternative constraint handling algorithms as discussed above, the URNDDR proposed by Vachhani et al. (2006) seems to be promising. The key idea in URNDDR is to replace the update step in UKF by a set of optimization problems that are solved at each sigma point to satisfy the constraints on the states. The updated (filtered) state is then obtained as a weighted combination of the individual updated sigma points. Recently, Teixeira et al. (2008) have proposed various approaches based on UKF and URNDDR combinations to solve constrained, nonlinear state estimation problems. However, URNDDR fails to yield feasible state estimates if the constraints are non-convex. Hence, URNDDR can lead to state estimates which do not satisfy state constraints. In this article we propose a modified URNDDR (MURNDDR) formulation that ensures that the filtered state estimate satisfies the specified constraints irrespective of the nature of the constraints (linear/nonlinear, convex/nonconvex, etc.). The improvements due to the modified approach over the basic URNDDR algorithm are demonstrated on a representative case study. Section 2 presents a brief summary of the URNDDR and MURNDDR approach. The utility of the modifications is demonstrated in Section 3 by applying it to a gas phase reaction simulation case study.

2. Estimation Algorithms

2.1. Unscented Recursive Nonlinear Dynamic Data Reconciliation (Vachhani et al., 2006; Narasimhan and Rengaswamy, 2008)

Consider the following state-space description of the process:

$$x_{k+1} = f(x_k, u_k) + w_k \tag{1}$$

$$y_{k+1} = h(x_{k+1}, u_{k+1}) + v_{k+1} \tag{2}$$

$$\text{s.t. } x_L \le x \le x_U; \quad d(x) \le 0; \quad g(x) = 0 \tag{3}$$

In the above model $x_k \in \Re^n$, $u_k \in \Re^m$ and $y_k \in \Re^r$ are respectively the states, manipulated inputs and the measurements at k^{th} time instant. Process noise $w_k \in \Re^n$ and measurement noise $v_k \in \Re^r$ are assumed to be Gaussian with zero mean and covariance matrices Q_k, R_k and uncorrelated with each other. Functions $f()$ and $h()$ represent the process and the measurement model respectively. x_L and x_U are lower and upper bounds on the state vector and $d(x)$ and $g(x)$ represent inequality and equality constraints. For given estimated state $\hat{x}_{k(+)}$ and covariance $P_{k(+)}$, URNDDR can be summarized as:

A) Prediction Step: The UT is used to determine mean and covariance of the propagated states. A set of $(2(n+r)+1)$ sigma points and associated weights (z_i) are calculated from following augmented matrices as:

$$\hat{x}_k^a = \begin{bmatrix} \hat{x}_k(+) \\ y_k \end{bmatrix} \text{ and } P_k^a(+) = \begin{bmatrix} P_k(+) & [0]_{nxr} \\ [0]_{rxn} & R_k \end{bmatrix}$$

$$\chi_{i,k}^a(+) = \begin{cases} \left[\hat{x}_k^a(+)\right] & i = 0 \\ \left[\hat{x}_k^a(+) + \theta_{i,k} S_{i,k}\right] & i = 1, \ldots, 2(n+r) \end{cases} \tag{4}$$

where, $S_{i,k} = \left[\sqrt{P_k^a(+)}\right]_i, i = 1, \ldots, n+r$ and $S_{i,k} = -\left[\sqrt{P_k^a(+)}\right]_{i-(n+r)}, i = n+r+1, \ldots, 2(n+r)$ where $\left[\sqrt{P_k^a(+)}\right]_i$ is the i^{th} column of square root of $P_k^a(+)$ matrix and indicates the direction along which the sigma points are selected. In URNDDR, for better approximation of covariance information in presence of bounds, the sigma points used in propagation step are located asymmetrically and the step size $\theta_{i,k}$ for i^{th} sigma point is chosen as:

$$\begin{aligned} \theta_{i,k} &= min(\sqrt{n+\kappa}, \theta_{1,k}, \theta_{2,k}) \\ \theta_{1,k} &= \min_{j:S_{i,k}^j > 0} (\infty, (x_U^j - \hat{x}_k^j(+))/S_{i,k}^j); \quad \theta_{2,k} = \min_{j:S_{i,k}^j < 0} (\infty, (x_L^j - \hat{x}_k^j(+))/S_{i,k}^j) \end{aligned} \tag{5}$$

In the above, $S_{i,k}^j$ is j^{th} entry of i^{th} column at k^{th} time instant. Since sigma points are not symmetric the weights corresponding to each sigma points are selected as:

$$z_i = a\theta_{i,k} + b \tag{6}$$

where, $a = \frac{2\kappa - 1}{2(n+r+\kappa)(S_\theta - (2(n+r)+1)(\sqrt{n+r+\kappa}))}, b = \frac{1}{2(n+r+\kappa)} - \frac{2\kappa - 1}{2\sqrt{n+r+\kappa}(S_\theta - (2(n+r)+1)(\sqrt{n+r+\kappa}))}, S_\theta = \sum_{i=1}^{2(n+r)} \theta_i$

where κ is a tuning parameter. For any symmetric prior distribution with kurtosis *Ku,* choosing κ such that $0 < n + r + \kappa \le Ku$, leads to mean and covariance predictions that are more accurate than those of EKF (Julier et al., 1995) . The transformed set of sigma points are obtained by propagating each sigma point through process model as:

$$\chi_{i,k+1}(-) = f(\chi_{i,k}(+), u_k) \tag{7}$$

The predicted mean and state prediction error covariance matrix (SPECM) are then computed as:

$$\hat{x}_{k+1}(-) = \sum_{i=0}^{2(n+r)} z_i \chi_{i,k+1}(-) \tag{8}$$

$$P_{k+1}(-) = \sum_{i=0}^{2(n+r)} z_i [\chi_{i,k+1}(-) - \hat{x}_{k+1}(-)][\chi_{i,k+1}(-) - \hat{x}_{k+1}(-)]^T + Q_{k+1} \tag{9}$$

In order to account for the effect of Q_{k+1}, sigma points are regenerated from computed $\hat{x}_{k+1}(-)$ and $P_{k+1}(-)$. The new predicted sigma points and modified weights are computed in the same manner as in Equations (4)-(6). These new predicted sigma points and weights are used for update step.

B) Update Step: The optimization problems are solved to obtain the updated sigma points set. The optimization problem corresponding to the i[th] sigma point is given by:

$$\begin{aligned} \min_{\chi_{i,k+1}(+)} \quad & \left(\chi_{i,k+1}(+) - \chi_{i,k+1}(-)\right)^T (P_{k+1}(-))^{-1} \left(\chi_{i,k+1}(+) - \chi_{i,k+1}(-)\right) + \\ & \left(y_{i,k+1} - h(\chi_{i,k+1}(+), u_{k+1})\right)^T (R_{k+1})^{-1} \left(y_{i,k+1} - h(\chi_{i,k+1}(+), u_{k+1})\right) \end{aligned} \tag{10}$$

subject to the following constraints

$$x_L \le \chi_{i,k+1}(+) \le x_U; \quad d(\chi_{i,k+1}(+)) \le 0; \quad g(\chi_{i,k+1}(+)) = 0$$

The updated sigma points $\chi_{i,k+1}(+) \quad i = 0,1,...,2(n+r)$, are obtained from the solution of above optimization problems, one for each i[th] point. The updated state and state estimation error covariance matrix (SEECM) are then calculated as:

$$\hat{x}_{k+1}(+) = \sum_{i=0}^{2(n+r)} z_i \chi_{i,k+1}(+); \quad P_{k+1}(+) = \sum_{i=0}^{2(n+r)} z_i [\chi_{i,k+1}(+) - \hat{x}_{k+1}(+)][\chi_{i,k+1}(+) - \hat{x}_{k+1}(+)]^T \tag{11}$$

It can be seen from the formulation (10) that while the individual updated sigma points will satisfy the state constraints (3), the same cannot be said for the updated state since it is a convex combination (11) of the individual updated sigma points. If the feasible region defined by the constraints is nonconvex, then there is no guarantee that a convex combination of the feasible points will yield another feasible point.

2.2. Modified Unscented Recursive Nonlinear Dynamic Data Reconciliation

In order to account for constraints on individual sigma points as well as on the updated state (which is a convex combination of the individual sigma points), in this work we propose to solve the following single optimization problem for the update step:

$$\begin{aligned} \underset{\substack{\chi_{i,k+1}(+) \\ i=0,...,2(n+r)}}{min} \quad & \sum_{i=0}^{2(n+r)} \left[\begin{array}{l} \left(\chi_{i,k+1}(+) - \chi_{i,k+1}(-)\right)^T (P_{k+1}(-))^{-1} \left(\chi_{i,k+1}(+) - \chi_{i,k+1}(-)\right) + \\ \left(y_{i,k+1} - h(\chi_{i,k+1}(+), u_{k+1})\right)^T (R_{k+1})^{-1} \left(y_{i,k+1} - h(\chi_{i,k+1}(+), u_{k+1})\right) \end{array} \right] \\ \text{s.t.} \quad & x_L \le \chi_{i,k+1}(+) \le x_U; \quad d(\chi_{i,k+1}(+)) \le 0; \quad g(\chi_{i,k+1}(+)) = 0; \quad i = 0,...,2(n+r) \\ & x_L \le \sum_{i=0}^{2(n+r)} z_i \chi_{i,k+1}(+) \le x_U; \quad d(\sum_{i=0}^{2(n+r)} z_i \chi_{i,k+1}(+)) \le 0; \quad g(\sum_{i=0}^{2(n+r)} z_i \chi_{i,k+1}(+)) = 0 \end{aligned} \tag{12}$$

In the above formulation, constraints are imposed on individual updated sigma points as well as on the updated state. The decision variables for this optimization problem are the *2(n+r)+1* updated sigma point vectors. It is also to be noted that in absence of constraints for linear system models, the updated state as computed by the modified optimization problem will be the same as that given by linear KF (this feature is similar to URNDDR). The advantage of (12) over URNDDR is that we are directly minimizing the error in state estimates subject to the constraints on individual sigma point vectors as well as convex combination of the individual updated sigma points. Hence, modified URNDDR approach is guaranteed to satisfy arbitrary nonlinear and nonconvex constraints on the states. In the proposed approach, at the k^{th} time instant, a single optimization problem with *(2(n+r)+1)n* decision variables is solved instead of solving *2(n+r)+1* separate optimization problems each with *n* decision variable as was the case in URNDDR. Rest of the procedure is the same as discussed in section 2.1. The proposed modification ensures that individual sigma points as well as the updated state estimate satisfy the given constraints irrespective of the nature of the constraints. The utility of the proposed algorithm is now demonstrated by applying it to a simulation gas phase reaction case study.

3. Results and Discussion

The URNDDR methodology as proposed by Vachhani et al. (2006) and modified approach as proposed in this article are applied on the gas phase reaction simulation case study (Haseltine and Rawlings, 2003). It involves an irreversible gas phase reaction: $2A \rightarrow B$, with reaction rate $r = k_e P_A^2$. The partial pressure (P.P.) of species A (P_A) and B (P_B) are the state variables and the measurement is of total pressure (i.e. $P=P_A+ P_B$). The discretized nonlinear model for the system is (Haseltine and Rawlings, 2003) given in equation (13) where $x=[P_A \quad P_B]^T$. The initial state ($\hat{x}_0$) and associated covariance matrix (P_0) are taken as $[0.001,0.045]^T$ and $3600*I_{2X2}$ (where I is identity matrix) respectively, and other parameters are taken from Haseltine and Rawlings (2003).

$$x_{k+1} = \begin{bmatrix} \frac{x_{k,1}}{2k_e\Delta t x_{k,1}+1} \\ x_{k,2} + \frac{k_e\Delta t x_{k,1}^2}{2k_e\Delta t x_{k,1}+1} \end{bmatrix} + w_k; \quad y_k = [1 \quad 1]x_k + v_k \tag{13}$$

The number of time samples considered is 150. The states as estimated by URNDDR and MURNDDR approach are presented in Figures 1 and 2 respectively. It can be seen from the figures that both approaches accurately track the true states while satisfying the upper and lower bounds (100 and 0 respectively). To demonstrate the ability of the MURNDDR approach to handle arbitrary nonlinear constraints which cannot be incorporated in basic URNDDR, we introduce an additional nonlinear constraint on state as:

$$193 \le D \le 259 \tag{14}$$

where at k^{th} time point, $D_k = \frac{(\hat{x}_k^1(+)-15)^2}{1} + \frac{(\hat{x}_k^2(+)-15)^2}{4}$. The values of upper and lower bounds in (14) are specified such that the true state satisfies this constraint and hence this constraint should improve the quality of estimation. In basic URNDDR, at the k^{th} time point, 7 (number of sigma points) optimization problems, each involving 2 decision variables (number of states) are solved. For the i^{th} optimization problem, constraint (14) is imposed as:

$$193 \le D_{k,i} \le 259; \text{ where } \quad D_{k,i} = \frac{(\chi_{k,i}^1(+)-15)^2}{1} + \frac{(\chi_{k,i}^2(+)-15)^2}{4} \tag{15}$$

On the other hand, in the MURNDDR approach, at the k^{th} time point, a single optimization problem is solved that involves 14 (i.e. 7x2) decision variables and the constraints (14) are given as:

$$193 \le D_{k,i} \le 259 \ \text{for } i = 1, 2, ..., 7; \text{ and } \quad 193 \le D_k \le 259 \tag{16}$$

The corresponding performances of URNDDR and MURNDDR at first time point are shown in Figures 3 and 4 respectively. It can be seen from these figures that the state (weighted sigma point) estimated by URNDDR violates the nonlinear constraint whereas the state estimated by the proposed MURNDDR satisfies the specified nonlinear constraint. Note that the URNDDR state estimate violates this constraint even though individual sigma points satisfy it since it was incorporated in the optimization formulation corresponding to the i$^{\text{th}}$ sigma point. This is due to the inability of URNDDR to guarantee satisfaction of nonconvex constraints by the updated state. However, MURNDDR suggested in this article is able to handle the specified constraint. The performance of URNDDR and MURNDDR with nonlinear constraints is compared by calculating average mean square error (MSE) in estimated states and the average time required for computation for a time point**.** All the computations were performed in Matlab version 7 running in Windows XP environment on a P4, 2GB RAM machine. The results are shown in Table 1. It can be seen from the Table that the quality of estimates is better for MURNDDR due to its ability to handle arbitrary nonlinear constraints. To summarize, this formulation enables incorporation of any arbitrary nonlinear, non-convex constraints on the states while retaining the recursive feature of URNDDR for nonlinear state estimation. These features, as illustrated by the presented case study, make the modified URNDDR a promising tool for nonlinear constrained state estimation. Implementation of the proposed approach on larger problems is currently under investigation.

Table 1 Comparison of performance of URNDDR and MURNDDR

Description	Average MSE	Time (sec)/iteration
URNDDR	0.0301	0.5
MURNDDR	0.0186	2.5

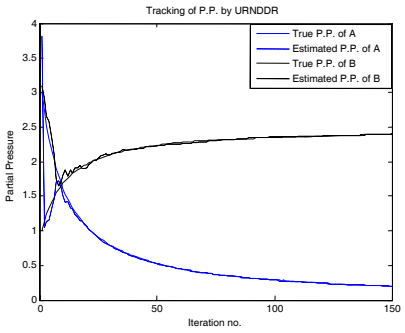


Figure 1: P.P. tracking by URNDDR

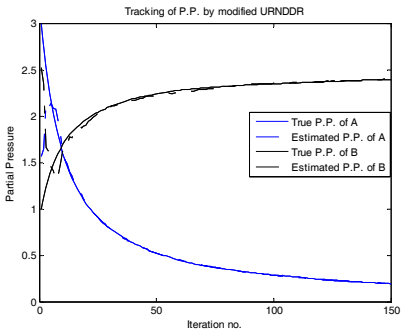


Figure 2: P.P. tracking by MURNDDR

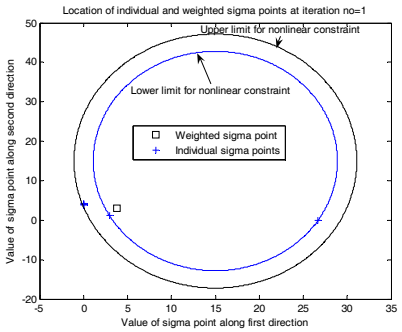


Figure 3: Location of points at first time instance for URNDDR

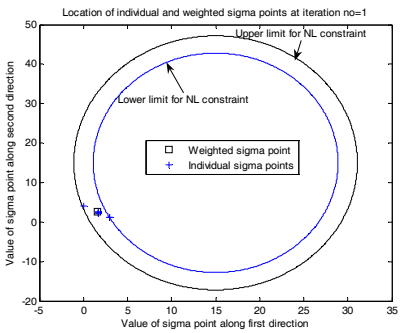


Figure 4: Location of points at first time instance for MURNDDR

References

P. Vachhani, S. Narasimhan, and R. Rengaswamy, 2006, Robust and reliable estimation via unscented recursive nonlinear dynamic data reconciliation, Journal of Process Control, Vol. 16, pp. 1075–1086.

S. J. Julier, J. K. Uhlmann, and H. C. Durrant-Whyte, 1995, A new approach for filtering nonlinear systems, in Pro. of the American Control Conf., USA, pp. 1628–1632.

P. Vachhani, R. Rengaswamy, V. Gangwal, and S. Narasimhan, 2005, Recursive estimation in constrained nonlinear dynamical systems, AIChE, Vol. 51, No. 3, pp. 946–959.

D. Simon, and T. L. Chia, 2002, Kalman filtering with state equality constraints, IEEE Trans. on Aerospace and Electronic Systems, Vol. 38, No. 1, pp. 128–136.

R. Kandepu, B. Foss, and L. Imsland, 2008, Applying the unscented Kalman filter for nonlinear state estimation, Journal of Process Control, Vol. 18, pp. 753–768.

B. O. S. Teixeira, L.A.B. Torres, L. A. Aguirre, and D. S. Bernstein, 2008, Unscented filtering for interval-constrained nonlinear systems, in Pro. of the 47th IEEE Conf. on decision and control, Cancun, Mexico, pp. 5116–5121.

S. Narasimhan, and R. Rengaswamy, 2008, Reply to Comments on robust and reliable estimation via unscented recursive nonlinear dynamic data reconciliation, Journal of Process Control, in press doi:10.1016/j.jprocont.2008.08.002.

E. Haseltine, and J. Rawlings, 2003, A critical evaluation of extended Kalman filtering and moving horizon estimation, Tech. report, TWMCC, University of Wisconsin, Madison, WI.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A performance study of dynamic RTO applied to a large-scale industrial continuous pulping process

Pablo A Rolandi[a], José A Romagnoli[b]

[a]*Process Systems Enterprise Ltd, London, UK*
[b]*Louisiana State University, Luisiana, USA*

Abstract

A first-principle model-centric approach for advanced control-and-optimisation of industrial manufacturing processes is presented in this work. The elements of this advanced framework are described along with performance studies conducted on industrial control problems using an optimiser based on linear and nonlinear models. The optimal operation of an industrial continuous pulping system is used as a case-study. The software implementation is executed in this dynamic real-time-optimisation (DRTO) environment.

1. Introduction

We envision flexibility and interoperability as the key technological breakthrough of the next generation of model-based APC systems. For example, such APC engine would allow embedding linear models as easily as nonlinear models. Similarly, the APC engine would support empirical/semi-empirical models derived from identification- or reduction-based techniques, as well as fundamental mechanistic models derived from first principles. Concurrently, the APC system would allow unconstrained, quadratic cost problem formulations or general constrained control problems. Finally, this next-generation APC engine would support discrete- and continuous-time formulations interchangeably, and would be integrated into the MPC multivariable control and/or RTO economic optimisation layers of the APC hierarchy. Of course, the form of the optimal control problem does not depend on the characteristics of the APC application and, therefore, a set of mechanisms to formulate (and subsequently interpret) this control problem must be provided. In summary, we argue that the next generation of APC systems should be founded on a domain framework and software platform that allows the interchange of models, solutions methods and control/optimisation settings and strategies seamlessly. Ultimately, such a framework/platform would enable a transparent comparison of controller design and control architecture alternatives.

In this paper, a model-centric platform for advanced process control-and-optimization of industrial processing systems is presented. It provides the appropriate framework through which the aforementioned research issues can be investigated and addressed in a thorough and systematic way. In this work, we present the architecture of the DRTO platform, and we discuss the formulation of industrial optimisation-and-control problems and its subsequent interpretation into a mathematical formalism. The analysis of the controller/optimiser performance using linear and nonlinear models is studied.

2. A Framework for Dynamic Real-Time Optimisation

2.1. Control hierarchy

In this work we devote our attention to the performance of online real-time dynamic optimisation (DRTO) and, therefore, we adopt a control-hierarchy architecture where DRTO directly provides the setpoints of PID controllers (i.e. the regulatory control system), as opposed to the conventional industrial approach where (steady-state) RTO provides the optimum targets to the multivariable MPC controllers.

2.2. Component architecture

The dynamic RTO engine (DRTOE), a software application developed in this work, comprises several software entities: the process-data server (PDS) (which interfaces to OPC servers), the modelling-and-solution engine (MSE), the problem-definition manager (PDM) and the solution-feasibility supervisor (SFS). These are described next.

2.3. Modelling and Solution Engine

The MSE encapsulates all modeling-and-solution services needed by the DRTOE kernel. The advanced process modelling package gPROMS, from Process Systems Enterprise Ltd, has been chosen as the MSE. gPROMS is a state-of-the-art, equation-oriented modelling system that supports hybrid continuous/discrete (HCD) integro-partial-differential-algebraic systems (IPDAEs) of arbitrary complexity. gPROMS meets the requirements of a model server as well as a robust, efficient, general-purpose solution engine.

2.4. Optimiser and virtual plant models

By choosing gPROMS as the MSE, no restrictions are imposed on the type of models that can be embedded in DRTOE. Implicit differential-algebraic systems (DEA) are the most common class of models arising from first-principles' modeling, as well as ordinary differential systems (ODE). In general, these models are highly nonlinear due to natural phenomena such as kinetics, transport and equilibrium (thermodynamics) relationships.

Linear models are a common class of models used in linear constrained/unconstrained APC. Although these models rarely arise from first principles' modelling, they result from linearising nonlinear models around one or more nominal operating points. In the former case, this result in a linear time-invariant (LTI) model; in the latter case, known as successive linearisation, the model is linear time-variant (LTV). gPROMS has built-in linearisation capabilities based on a hybrid structural and numerical algorithm which presents the model in a linear state-space form. LTI models are fully supported by the DRTO engine and LTV models will be included in future work.

A large-scale first-principles' model is available from previous work. This model is used as the virtual plant for all closed-loop runs. In order to avoid plant/model mismatch, this model is also embedded in the control/optimisation engine; naturally plant/model mismatch does exist when this model is subject to linearisation.

2.5. Problem Definition Manager

The industrial operation of continuous and batch processes demand control-and-optimisation problems that change constantly as time progresses. This is due to changes in process conditions and operation specifications. The translation of the control-and-optimisation problem into a high-level dynamic optimization problem is executed by the Problem Definition Supervisor component (PDM). This component interprets the control events posted by users/operators and implements the resulting dynamic-optimisation problem formulation as requested by the MSE. This could be either input

arguments to low-level numerical routines or high-level constructs of declarative modeling languages.

In practice, multivariable, constrained control-and-optimisation problems seeking to improve process performance are difficult to pose and may even result in incorrect or infeasible formulations. Previous studies show that approximately a dozen different types of control events are needed to define a typical industrial control/optimitsation problem. These events have been presented in previous work (Rolandi & Romagnoli, 2008). The so-called "elementary events" have a direct interpretation into the parameterisation of dynamic optimisation problems. For instance, in the case of gPROMS, this means that these events have a direct representation into the gPROMS optimisation entity language. On the other hand, "composite events" are high-level nontrivial instructions that must be reinterpreted into elementary events. Path and zone constraints are two of the most common composite events, and they are supported in the current version of the software. In principle, this concept can be extended to encompass whole transition planning and scheduling scenarios. Overall, the introduction of these composite events greatly simplifies the definition of industrial control-and-optimisation problems by encapsulating operation recipes.

Scaling of decision variables is performed automatically by gPROMS (and most optimisation routines). Scaling of constraints, which is not normally built-in in optimisation modules, is performed by the PDM using the event mechanism described above. This transformation, which is derived from understanding of the optimality criterion, allows fine control of the enforcement of constraints which is needed to improve the robustness and efficiency of industrial control problems.

2.6. Solution and Feasibility Supervisor

The outcome of the PDM is, in general, a constrained dynamic optimisation problem whose solution demands not only optimality but also feasibility considerations. Indeed, the existence of a feasible solution is a key issue to be addressed by any control application. As a consequence, the proposed framework/platform has different build-in strategies to recover from infeasibilities: the Solution-Feasibility Supervisor (SFS) is a dedicated component of the DRTO engine for monitoring the solution and handling of infeasibilities. The SFS is composed by a Solution Interpreter (SI) and a Constraint Manager (CM).

Ideally, the SI would continuously monitor the progress of the open-loop optimisation computation to detect slow convergence between successive iterations due to constraint infeasibility/inconsistency. Due to technical restrictions, the current implementation of the SI logs the magnitude of control variables and constraints violations and overall computation statistics (e.g. number of NLP iterations, line searches and corresponding times) upon termination of the inner optimisation loop, where failing to converge is not considered fatal.

The implementation of the Constraint Manager, a concept that had already been presented in previous publications by these authors (Rolandi & Romagnoli, 2008), is a novelty of this work. The CM is built-in with two independent infeasibility-recovery strategies: i) relaxation vs elimination and ii) ranking vs identification. These individual strategies can be combined to create four different constraint-management policies.

The most common recovery mechanism in commercial MPC packages is constraint ranking and elimination (Qin & Badgwell, 2003). Here, one or more constraints are associated to a ranking level and removed from the control problem formulation upon infeasibility when their priority is below a cut-off level. We have found that this mechanism is sub-optimal and, furthermore, does not guarantee immediate recovery,

leading to inefficient computations. Constraint identification, on the contrary, uses information on the constraints' bounds and values (or their Lagrange multipliers, if available) to detect the subset that is active at any given iteration. Constraints that are violated are either eliminated or relaxed. The relaxation takes place by applying a factor which is provided by the user/operator via a console event mechanism (by default, it is set to 1% of the nominal variable value). Identification is coupled with ranking information because only constraints below a cut-off level can be relaxed or eliminated.

3. Case-study and analysis of results

3.1. Industrial wood pulping

The case study presented in this section is based on an industrial continuous pulping digester (and its auxiliary units) of a state-of-the-art pulp and paper mill. Papermaking is a highly competitive industry where meeting quality control parameters and the demand of highly dynamic markets is paramount to the economics of the business; also, new plants with more flexible and efficient productions schemes often replace old production systems. The key unit in this process is a vertical tubular reactor with several fluid-phase extraction and introduction points along the length of the reactor. The reactor is a solid-liquid heterogeneous system with inter- and intra-particular transport phenomena, where the column of wood chips moves continuously down and the fluid phase moves in co-current or counter-current flow at different heights. Besides the tubular reactor, there is auxiliary heating and heat-recovery equipment consisting of nine unit operations (condensers, heaters, pre-heaters and a kettle-reboiler), and seven transportation and handling devices (the so-called feed line). There are approximately 50 PID control loops, switched to manual or automatic operating mode.

3.2. 3.2 First principles model

Large-scale mechanistic models (that is, models of order ~1E5 variables or so) have seldom been used in advanced model-based control systems, with only a few examples resulting from academic studies rather than industrial applications (e.g. Wisnewski & Doyle, 2001). In this work, a mechanistic model of the continuous pulping system described above has been implemented in gPROMS, resulting in a system of approximately 14,000 algebraic, 1,000 differential equations, and 300 degrees-of-freedom.

3.3. Other implementation aspects

In this work, real-time dynamic optimisation provides the setpoints of key loops of the regulatory control system, condensing the two intermediate optimisation/control layers of a conventional hierarchical control structure (i.e. MPC and RTO) into a single, consistent model-centric application (DRTO).

Depending on the configuration, either this full nonlinear model is embedded in the optimizer or a reduced-order linear model. The latter is obtained by exact linearisation and minimal realisation of the former (the first-principles' nonlinear model) at the nominal steady-state starting point (in both case studies this is at 600.0 ad.ton/day); this is computed by the MSE (gPROMS). Linearisation requires full knowledge of the operating point, that is, degrees-of-freedom and initial conditions of the system. In order to satisfy this requirement as well as the need for state feedback during the movement of the receding horizon, a perfect observer is implemented. This is achieved by accessing the values of states and inputs of the dynamic model used as virtual plant, which is implemented via gPROMS' FOI/FPI communication mechanisms.

While a perfect observer is not feasible in an industrial setting, this choice allows us to focus on algorithmic and performance aspects and avoid any issues introduced by the ubiquitous plant/model mismatch that would characterise any real industrial application.

3.4. Case-study 1

In this case study the main target is to control a production-rate transition from 600.0 ad.ton/day to 650.0 ad.ton/day. From an operations-and-control perspective, the goal is to maximise the pulp yield (Y) at a given production target (P) while maintaining the deviation of pulp selectivity from its quality-control target below a given threshold. Selectivity is measured in terms of the "kappa number" and it is an indication of the proportion of non-cellulosic components present in the wood pulp. The process is initially at steady-state and control actions start taking place three hours before the scheduled production-rate change. The manipulated variables (MVs) optimized by DRTO are the setpoints of three key PID controllers: the chip meter (CM) speed (feed rate of wood chips), and the temperature of the lower (LH) and wash (WH) circulation heaters (which perform indirect heating of the chip column). Two interior-point and two end-point constraints are imposed on the trajectories and final values of selectivity and pulp production rate, and one path constraint is introduced for the deviation of selectivity with respect to its control target. The prediction and control horizons are set to 7 and 5 hr, respectively, and the control window is 1 hr; these settings are representative of industrial practices.

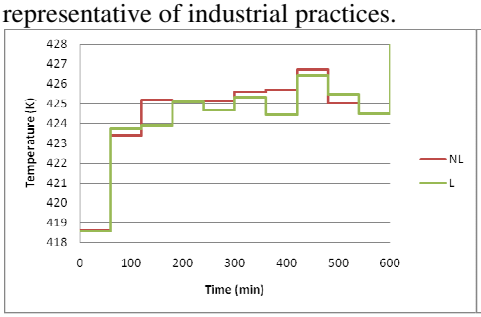


Figure 1: Wash heater (WH), MV.

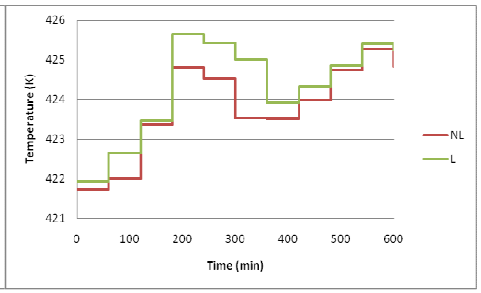


Figure 2: Lower heater (LH), MV.

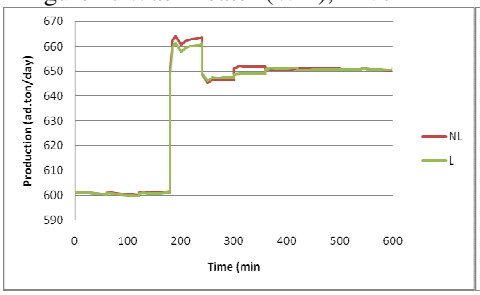


Figure 3: Production (P), CV.

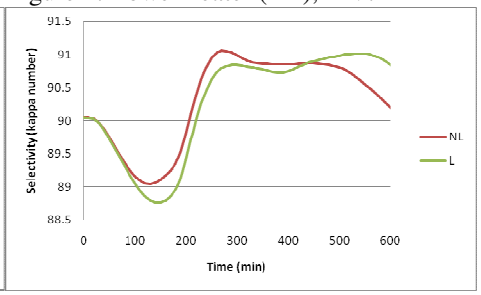


Figure 4: Selectivity (S), CV.

Figures 1 to 4 show the trajectories of key process variables for this case-study, for the linear (L) and nonlinear (NL) embedded model. While the progression of control moves (MVs) is similar for both controllers, the L-controller imposes a higher temperature setpoint for the LH PID loop and a lower setpoint for the WH PID than the NL-controller does. In general, the selectivity trajectories are within the control limits (between kappa number 89 and 91), although the L-controller slightly violates this constraint between 100min and 200min approximately. Overall, the NL-controller and the L-controller exhibit similar control performance. The good performance of the L-

controller is due to fact that the plant is reasonably linear in this limited range of operation and an accurate model of the plant has been obtained by exact linearisartion of the first principles' nonlinear model. Also, the prefect observer periodically corrects the effects of plant/model mismatch by resetting the linear model to the true operating point at every control window. While both controllers execute roughly the same number of NLP iterations, the L-controller is approximately 5 times faster than the NL-controller. Finally, the L-controller hits one infeasibility and recovers, while the NL-controller always remains feasible.

3.5. Case-study 2

This case study is a variation of case study 1. Again, a production-rate transition from 600.0 ad.ton/day to 650.0 ad.ton/day is scheduled, this time, two hours in advance. However, two hours after the transition a failure in downstream processing equipment triggers a production slow-down to the original rate of 600.0 ad.ton/day, which is enforced two hours later. This is a common yet challenging operating scenario in integrated pulp and paper mills which have low inventories by design.

See figures 5 and 6. This time, the NL-controller outperforms the L-controller, since it is able to maintain the constraint on selectivity at all times. The NL-controller also gives a higher mean yield (objective function) throughout the operation horizon, which is an indication that optimality has been achieved (at least locally).

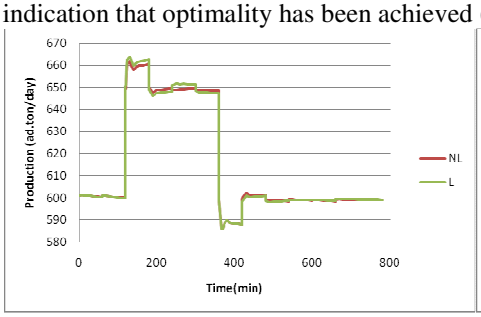


Figure 5: Production (P), CV.

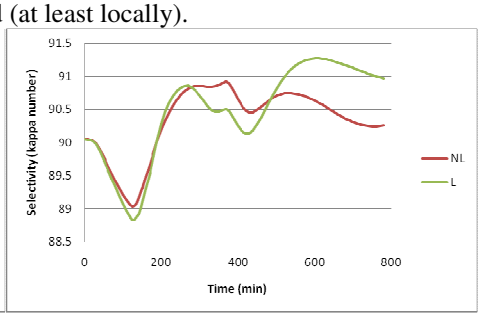


Figure 6: Selectivity (S), CV.

4. Conclusions and Future Work

In this manuscript we presented a framework and software platform that enables the adoption of different model types, solutions methods, and control-and-optimisation strategies transparently and effectively. We examined the performance of a novel DRTO engine through two industrial case studies using linear and nonlinear models. Both first principles model-based controllers performed well and were able to drive the plant through production transitions satisfactorily; while the nonlinear model handled constraints better, the linear model resulted in much faster computational times. The constraint identification and relaxation mechanism for recovery of infeasibilities performed well. Future work will involve refinements to this framework as well as the development of observers for realistic closed-loop feedback with plant/model mismatch.

5. Acknowledgements

The authors wish to acknowledge the contribution of Dr Joseph Zeaiter to this work, who enthusiastically shared his vast experience in and knowledge of industrial APC.

References

Wisnewski, P.A. and Doyle, F.J. (2001). IEEE Trans Control Sys Tech, 9, 435-444.
Rolandi P.A., Romagnoli, J.A. (2008). ESCAPE-18, Lyon, France.
Qin, S.J., Badgwell, T.A. (2003). Control Engineering Practice, 11, 733–764.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Control Structure Selection Based Upon Rigorous Dynamic Process Models

Le Chi Pham[a], Sebastian Engell[a]

[a]Process Dynamics and Operations Group, Department of Biochemical and Chemical Engineering Technische Universität Dortmund
Emil-Figge-Str. 70, 44221 Dortmund, Germany
E-mail address: le-chi.pham | s.engell@bci.tu-dortmund.de

Abstract

In this contribution, the selection of control structures based upon rigorous stationary process models is extended to include the dynamic performance in a consistent manner. A key problem in the comparison of control structures with respect to dynamic performance is that the performance depends on the type and the parameterization of the controllers that are used. In our approach, linear MPC-controllers are assumed and the weights of these controllers are optimized for each structure to yield an optimal economic performance for the disturbance scenarios considered. Thus the result is a good approximation of the best performance that is attainable for each structure and the structures are compared on equal grounds. The approach is demonstrated for a distillation problem.

Keywords: Control structure selection, economic performance, distillation control.

1. Introduction

Control structure selection deals with the choice of manipulated and measured variables used in feedback control. The result of it often has a stronger influence on the performance of the resulting closed-loop system than the choice of control algorithms. Most of the published work from the control community on control structure selection focuses on the dynamic tracking and regulation performance rather than the resulting economic performance of plant, which however should be the main goal of control design for chemical processing plants (see e.g. Engell, 2007). Morari et al. (1980) discussed several important aspects of the control of chemical processes from the point of view of plant performance. This was followed by Skogestad (2000) with the idea of "self-optimizing" control which means that a close-to-optimal steady state performance of the process in the presence of disturbances is established by a well chosen linear control structure. The focus is on the resulting performance at the steady state, and a stepwise approach is proposed to determine the most promising structures. Engell et al. (2005) extended this idea by some refinements replacing the informed judgement by objective criteria and optimization procedures, in particular considering the effect of measurement errors. After the static analysis, the remaining structures were screened by using dynamic performance indicators from linear control theory (Engell et al, 2004).
In the work mentioned above, the disturbances are assumed to be slowly time-varying or constant, so the steady-state behavior is dominant in the economic performance. An incorrect judgment of the performance of the controlled plant may result because the reaction to fast disturbances is not included. In this paper, we extend the procedure

proposed in (Engell et al, 2005, Scharf, 2007) by an evaluation of the dynamic closed-loop performance based upon full rigorous process models.

2. Plant Performance and Control Structure Selection

From a process engineering point of view, the role of automatic feedback control is to keep the process operation as close as possible to the economic optimum in the presence of disturbances and differences between the plant model used for the design of the plant and the real plant. The goal is to regulate the controlled variables to their setpoints in the presence of disturbances in such a way that the optimal manipulated variables u_{opt} are approximated. The effect of feedback control on the profit function J can be expressed as (Engell et al., 2005, Engell, 2007):

$$\Delta J = \left(J(\underline{u}_{nom}, 0) - J(\underline{u}_{nom}, \underline{d}_i) \right) + \left(J(\underline{u}_{nom}, \underline{d}_i) - J(\underline{u}_{opt}, \underline{d}_i) \right) + \left(J(\underline{u}_{opt}, \underline{d}_i) - J(\underline{u}_{con}, \underline{d}_i) \right)$$

The first term is the loss if the manipulated variables are fixed at their nominal values when disturbances occur, the second term is the effect of the optimal adaptation of the manipulated variables in the presence of disturbances, and the third term is the difference between the optimal adaptation and the one which is realized by the chosen feedback control structure. If the first term is large compared with the other terms then changing the manipulated variables will bring little profit, neither online optimizing nor feedback control is useful here. On the other side if the third term is large then online optimizing control (Engell, 2007) should be used instead of feedback control. The overall static performance of a control structure can be quantified by the expected loss of profit (Engell et al., 2005):

$$\Delta J = \int_{-d_{1,\max}}^{d_{1,\max}} \dots \int_{-d_{n,\max}}^{d_{n,\max}} w(\underline{d})(J(\underline{u}_{nom}, \underline{d}_i) - J(\underline{u}_{con}, \underline{d}_i))dd_1 \dots dd_n$$

where $w(\underline{d})$ is the probability of the occurrence of the disturbance $\underline{d}$. Practically this expression is approximated by the weighted sum over a set of disturbance scenarios.

There are two kinds of disturbances which affect the process: slow and fast varying disturbances. The effect of the former ones can be taken by into account by using a static optimization which analyzes the worst case performance of the regulatory control keeping the controlled variables within a range around the set-points defined by the measurement errors (Engell et al., 2005, Scharf, 2007). How to include the effect of the latter ones is discussed in this paper.

It is proposed to include the dynamic performance in a consistent manner by extending the above analysis to incorporate a full dynamic process model, dynamic controllers, and fast disturbances. In order to avoid the assumption of simple linear control structures and to be reasonably close to the best possible performance that a control structure can yield without excessive numerical effort, linear MPC controllers are assumed As the focus is on the economic performance, the weights used in the cost functions of the MPC-controllers are optimized for each structure with respect to the optimal economic performance for all disturbance scenarios considered.

3. Control Structure Selection Procedure

The proposed control structure selection procedure consists of 6 steps:

3.1. Define the optimization problem:

Analyze the available degrees of freedom of the process during plant operation and choose the manipulated variables. Formulate a profit function *J* to be maximized and specify the constraints that have to be satisfied during the process operation.

$$\max_{u} J(\underline{x}, \underline{u}, \underline{d}_i)$$

s.t system model

$$\underline{h}(\underline{x}, \underline{u}) \le 0$$

The system model can be either dynamic or static, $\underline{h}(\underline{x}, \underline{u})$ are the process constraints, d_i are the disturbances. The output mapping is given by: $\underline{y} = \underline{m}(\underline{x})$, in steady state $\underline{y} = \underline{M}(\underline{u}, \underline{d}_i)$.

3.2. Choose the disturbances:

It is distinguished between two types of disturbances: measurement errors and external disturbances. The former ones can be obtained from the instrument data-sheets. The latter ones may be caused by errors in the assumed model, disturbances, etc.

3.3. Pre-selection of the control structures

The number of possible control structures is given by:

$$C_n^k = \binom{n}{k} = \frac{n!}{(n-k)!\,k!}$$

with *n*: the number of available measurements and *k*: the number of controlled variables. This number increases rapidly with the number of measurements. Pre-screening of the unpromising structures is useful here especially for large problems. Several indices can be used, e.g. sensitivity to measurement errors (Engell et al., 2004) or linear performance indicators (Engell et al., 2004).

3.4. Selection of the set-points for regulatory control

The set-points are determined by optimization over the set of disturbance scenarios to fully utilize the potential of feedback control. The optimal set-points are found by solving:

$$\max_{\underline{y}_{set}} \sum_{i=1}^{n} J(\underline{x}, \underline{u}_i, \underline{d}_i)$$

$$s.t.: \forall \underline{d}_i :$$

$$\dot{\underline{x}} = f(\underline{x}, \underline{u}_i, \underline{d}_i) = 0$$

$$\underline{h}(\underline{x}, \underline{u}) \le 0, \underline{y}_{set} = \underline{m}(\underline{x})$$

$$u_{\min} \le u_i \le u_{\max}, x_{\min} \le x \le x_{\max}$$

If the above optimization problem is infeasible, it means that there is no common set-point which can be attained for all disturbances and the given constraints.

3.5. Quantitative evaluation of the benefits of the control structures with constant disturbances

For all scenarios of disturbances $\underline{d}_i$, the worst case control performance for regulation of the controlled variables to values in the range around the nominal set-point $\underline{y}_{set}$ defined by the measurement error $\underline{e}_{sensor}$ is obtained by solving the following optimization problem:

$$\min J(\underline{x}, \underline{u}_i, \underline{d}_i)$$

$$s.t.: \dot{\underline{x}} = f(\underline{x}, \underline{u}_i, \underline{d}_i) = 0$$

$$\underline{h}(\underline{x}, \underline{u}) \le 0, \underline{y} = \underline{m}(\underline{x})$$

$$\underline{y}_{set} - \underline{e}_{sensor} \le \underline{y} \le \underline{y}_{set} + \underline{e}_{sensor}$$

A comparatively large value of the maximum loss means that the corresponding control structure is not able to avoid a poor stationary performance in the presence of the measurement errors and should be excluded.

3.6. Quantitative evaluation of the benefits of the control structures with dynamic disturbances

After the controlled variables that are regulated to their set-points have been determined, the maximum attainable dynamic performance of the controlled system is compared. In order to accomplish this, a simulation of a linear MPC controller which is based upon a linearized model of the plant is employed. The objective function of the MPC controller is defined as

$$P(t_k) = \min_{\mathbf{u}} \left(\int_{t=t_k}^{t_k+H_P} \left(\|\Delta y(t)\|_P + \|\Delta u(t)\|_Q \right) dt \right)$$

$$s.t.: \quad \Delta \dot{x}(t) = A\Delta x(t) + B\Delta u(t) + C\Delta d(t) \qquad (P1)$$

$$\Delta y = D\Delta x(t)$$

where $\|.\|_X$ denotes the norm which is defined by: $\|u\|_X = u^T Xu$, X is a positive semi-definite matrix. A, B, C, D are matrices which are obtained by linearizing the system at the operating points found in step 4. P and Q are degrees of freedom and should be chosen such that the economic profit function J is maximized. The prediction horizon H_p is chosen long enough to capture all effects of the disturbances. An upper layer optimization is used in order to compute P and Q over all disturbances:

$$\max_{P,Q} \int_{t=0}^{t=t_{end}} \sum_{i=1}^{n} J(\underline{x}, \underline{u}_i, \underline{d}_i) dt$$

$$s.t.: P, Q \succ 0$$

$$(P1)$$

Thus the result is a good approximation of the attainable performance for each structure. The structure which yields the best performance using the same criterion as step 5 will be chosen.

4. Case study

The methodology described above is applied to a continuous distillation column shown in Figure 1. The distillation column is used for the separation of a binary mixture of Methanol and n-Propanol. The model of the process (Diehl, 2001) is based on the following assumptions: total condenser, negligible vapour holdup, variable liquid holdup, liquid outflow determined by Francis weir formula, constant pressure losses, perfect mixing, the mixture is at equilibrium temperature, Murphree efficiency is applied for each tray. The system model for 40 trays is large and stiff with 82 differential and 122 algebraic variables. There are two operating degrees of freedom after closing the inventory loops: the reflux rate and the heat supply. A composition measurement is assumed to be too expensive so the controlled variables are two tray temperatures. There are: $C_{40}^2 = 780$ possible control structures. The profit function is:

$$J = c_{Methanol} \tanh(x_{Methanol} - desiredvalue)\dot{n}_{Methanol} - c_{heat}\, heatinput$$

tanh implies that the purity of distillate product should satisfy the specification which is 0.99 in our case. The disturbances are chosen to be: a step change in the feed flow rate

by ±4 l/h, a step change in the feed concentration by ±0.1, a step change in the feed temperature by ±5 K for both the steady state and the dynamic analysis. For the steady state case, additional disturbances are considered: a change in the condenser temperature by ±5 K, a change in the heat loss by ±0.2 kW. In the dynamic scenario, the disturbances are assumed to be uniformly distributed within the lower and upper bound. The sensor error is assumed to be ±0.33 K. The heat input and the reflux flow rate are restricted to 5 kW and 7 l/h.

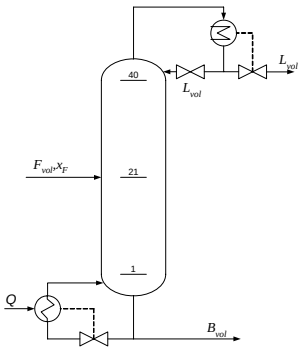


Fig. 1. Distillation Column

First a pre-screening is performed to take unpromising structures out of the large number of options. The generalized criterion of non square RGA (Skogestad and Postlethwaite, 1996) is used. 429 structures with a sum of the elements in row of the RGA less than 0.1 are discarded. Next, a measurement error sensitivity analysis is performed and 78 structures are left for the next step. In step 4, all the structures have a common set-point so 78 structures are considered further in step 5. Only those structures which lead to a profit larger than 85% of the nominal operating profit of 114.893 in the worst case are retained in the final step. 12 structures that satisfy this criterion are compared in the final step with respect to their dynamic performance using the linear MPC simulation with the nonlinear rigorous process model. For the optimization of the weights of the MPC controllers, a gradient-free optimization procedure was used (Berghen and Bersini, 2005). The performance indices of several structures are given in Table 1. Structure (10,31) yields the best performance.

Structure	Profit
(6,35)	107.67
(7,33)	106.36
(10,31)	111.02
(11,32)	109.85

Table 1. Performance indices of several structures resulting from a simulation of optimized linear MPC controllers

The results of the MPC simulation of this structure and of the structure (6,35) with disturbances of the feed flow and of the feed temperature are shown in Fig. 2 and Fig. 3. It can be seen that in the presence of strong disturbances in the feed flow by 30% of the nominal flow, the controller with the structure (10,31) manages to keep the controlled variables at their set-points and the top product purity still satisfies the requirement. The structure (6,35) could not keep the purity at the specified value. In the case of a disturbance in the flow rate, the purity drops down close to 0.985. This result verifies the heuristics that the best positions for controlling the composition of a distillation column using two tray temperatures are somewhere in the middle of the rectifying and stripping sections.

5. Conclusion

A methodology for control structure selection was presented with the aim of optimizing the plant performance taking into account the presence of both steady-state and dynamic disturbances. The method was applied successfully to the example of a distillation column. Because of the large computation times for NMPC simulations and in view of the industrial practice to use linear MPC controllers, the performance evaluation in the dynamic case is based upon linear MPC controllers with weights that are optimized for economic performance in an outer loop. Steps 1 – 6 of the control structure selection procedure yield a set of control structures that lead to good performance with respect to

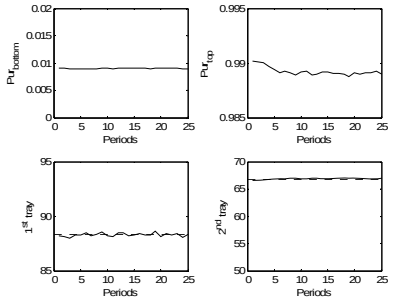


Fig. 2.aSimulation of linear MPC for structure (10,31) with disturbance in the feed flow rate

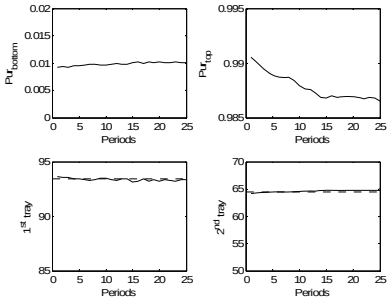


Fig. 2.b Simulation of linear MPC for structure (6,35) with disturbance in the feed flow rate

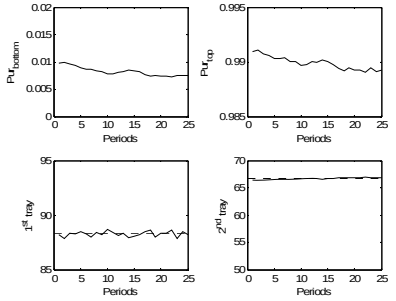


Fig. 3.a Simulation of linear MPC for structure (10,31) with disturbance in the feed temperature

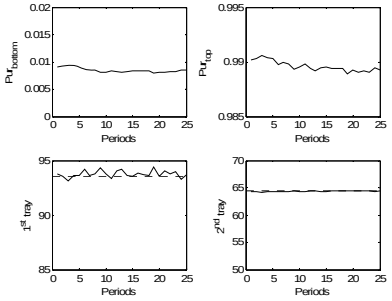


Fig. 3.b Simulation of linear MPC for structure (6,35) with disturbance in the feed temperature

the economic performance of the plant both for stationary or slowly-varying and for fast-varying disturbances. However, it may happen that some of these structures are not suitable for dynamic operation in the presence plant-model mismatch (lack of robustness). To check this, a linear dynamic controllability analysis should also be performed, cf. Engell at al. (2004).

References

F.V. Berghen, H. Bersini, 2005, CONDOR, a new parallel, constrained extension of Powell's UOBYQA algorithm. Experimental results and comparison with the DFO algorithm. Journal of Computational and Applied Mathematics, 157-175.

M. Diehl, 2001, Real-time Optimization for Large Scale Nonlinear Process, Dr. rer. nat. thesis, Universität Heidelberg.

M. Morari, Y. Arkun, G. Stephanopoulos, 1980, Studies in the Synthesis of Control Structures for Chemical Processes, AIChE Journal 26, 220-260.

S. Engell, J.O. Trierweiler, S. Pegel, M. Völker, 2004, Tools and indices for dynamic controllability assessment and control structure selection. In: Integration of Design and Control (P.Seferlis, M.C. Georgiadis, Eds.), 430-464, Elsevier.

S. Engell, T. Scharf, M. Völker, 2005, A methodology for control structure selection based on a rigorous process model. Proc. 16th IFAC World Congress, Prague, paper Tu-E14-TO/6.

S. Engell, 2007, Feedback control for optimal process operation. Journal of Process Control 17, 203-219.

S. Skogestad and I. Postlethwaite, 1996, Multivariable Feedback Control: Analysis and Design, John Wiley and Sons.

S. Skogestad, 2000, Plantwide control: the search for the self-optimizing control structure. Journal of Process Control 10, 487-507.

T. Scharf, 2006, Prozessbezogene optimierungsbasierte Regelungsstrukturauswahl mit Anwendung auf die Reaktiverektifikation, Dr-Ing thesis, Fachbereich BCI, Universität Dortmund (Shaker Verlag, Aachen, 2007).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Model-Based Fault-Tolerant Control of Particulate Processes: Handling Uncertainty, Constraints and Measurement Limitations

Arthi Giridhar[a], Sathyendra Ghantasala[a], Nael H. El-Farra[a]

[a]*Department of Chemical Engineering and Materials Science, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA*

Abstract

This paper presents a methodology for the integrated synthesis of a fault detection and isolation (FDI) and fault-tolerant control (FTC) system for particulate processes described by population balance models. The approach is based on a suitable low-order model of the process and accounts explicitly for model uncertainty, control constraints and measurement errors. The main idea is to shape the fault-free closed-loop dynamics via robust control in a way that enables the derivation of FDI and reconfiguration criteria that are less sensitive to the uncertainties and measurement errors. The results are illustrated through an application to a continuous crystallizer with a fines trap.

Keywords: Robust control, Process monitoring, FDI, Reconfiguration, Crystallizer.

1. Introduction

Particulate processes characterized by the co-presence of continuous and particulate phases are encountered in a large number of processing industries including agricultural, chemical, food, minerals, and pharmaceuticals. Despite the significant and growing body of research work on particulate process control (e.g., see Semino and Ray (1995), Christofides (2002), Zhang and Rohani (2003), Park et al. (2004), Larson et al. (2006) for some results and references in this area), the problem of designing model-based fault diagnosis and fault-tolerant control systems for particulate processes has received limited attention. This is an important problem given the fact that the erosion of controller authority caused by control system malfunctions can directly impact the end product quality and lead to substantial production losses (El-Farra and Giridhar, 2008). It is also a challenging problem due to the infinite-dimensional nature of particulate process models, which precludes their direct usage for control and monitoring purposes, the presence of model uncertainty, control constraints and the inherent limitations on the number and accuracy of process measurements, which can seriously erode the diagnostic and fault-tolerance capabilities of the control system if not explicitly accounted for at the controller design stage. In this paper, we present a robust fault diagnosis and fault-tolerant control (FTC) structure for particulate processes with control actuator faults. Unlike our previous work in El-Farra and Giridhar (2008), the structure presented here addresses explicitly the problems of time-varying model uncertainty and measurement errors. Initially, a finite-dimensional system that captures the dominant process dynamics is derived and decomposed into a number of interconnected subsystems that are each excited by a single manipulated input. A bounded Lyapunov-based controller that satisfies the control constraints is then

designed for each subsystem, leading to an explicit characterization of (1) the fault-free behavior of each subsystem in terms of a time-varying bound on the dissipation rate of the corresponding Lyapunov function, and (2) the robust stability region where constraint satisfaction and robustness with respect to the uncertainty and measurement errors are guaranteed. Using the fault-free bounds as alarm thresholds for fault detection and isolation (FDI), faults in a given actuator are diagnosed by monitoring the evolution of the process within the stability region and declaring a fault when the corresponding threshold is breached. The threshold depends on the achievable degree of asymptotic uncertainty attenuation and thus can be adjusted by properly tuning the controller. To ensure robustness of the FDI scheme to measurement errors, the monitoring region is confined within a subset of the stability region, and the FDI thresholds are enlarged appropriately. The robust FDI scheme is integrated with a controller reconfiguration strategy that orchestrates stabilizing transitions from the faulty actuators to a well-functioning fallback configuration following FDI. Practical implementation issues, such as robustness to model reduction errors when the model-based FTC architecture is implemented on the particulate process, as well as the problem of incomplete state measurements, are also discussed. Finally, the proposed methodology is applied to robustly stabilize a continuous crystallizer with a fines trap under control actuator faults**.**

2. Particulate processes with uncertainty, constraints and actuator faults

2.1 Process modeling using population, mass and energy balances

We focus on spatially homogeneous particulate processes modeled by the following general nonlinear system of partial integro-differential equations:

$$\begin{aligned}
&\frac{\partial n}{\partial t} = -\frac{\partial\left(G(x,r)n\right)}{\partial r} + w\left(n,x,r,\theta(t)\right) + g_1\left(n,x,r\right)\left[u_1^k + f_{a1}^k\right], n(0,t) = b(x(t)) \\
&\dot{x} = f(x) + g_2(x)\left[u_2^k + f_{a2}^k\right] + \chi\left(x,\theta(t),\int_0^{r_{\max}} q\left(n,x,r\right)dr\right) \\
&\left\|u_i^k(t)\right\| \le u_{i,\max}^k, i = 1,2; \left\|\theta(t)\right\| \le \theta_b, k(t) \in \{1,2,...,L\}, L < \infty
\end{aligned} \tag{1}$$

where $n(r,t)$ is the particle size distribution function, which is assumed to be continuous and sufficiently smooth in its arguments, $r \in [0, r_{max})$ is the particle size, t is the time, $x \in R^n$ is the vector of state variables that describe properties of the continuous phase (e.g., solute concentration, temperature and pH in a crystallizer), $u^k_1 \in R$ is the manipulated input associated with the particulate phase (e.g., fines destruction rate), $u^k_2 \in R$ is the manipulated input associated with the continuous phase (e.g., solute feed concentration in a crystallizer), $u^k_{i,max}$ is a real number that captures the size of constraints on the magnitude of the *i-th* manipulated input, $f^k_{a,i} \in R$ denotes a fault in the *i-th* control actuator, $k(t)$ is a discrete variable that denotes which control configuration is active at time t, $\theta \in R^q$ denotes the vector of uncertain process variables (e.g., unknown process parameters and time-varying external disturbances), θ_b is a known bound on the size of the uncertainty. In the population balance of Eq.1, $G(x,r)$ is the growth rate, $w(n,x,r,\theta) := w_0(n,x,r) + w_u(n,x,r)\theta$ is the net rate of introduction of new particles into the system, and $f(\cdot), q(\cdot), g_1(\cdot) g_2(\cdot), \chi(\cdot)$ are smooth nonlinear functions. The term containing the integral accounts for mass and heat transfer from the continuous phase to all the particles in the population and has the form $\chi(\cdot,\cdot,\theta) = \chi_0(\cdot,\cdot) + \chi_u(\cdot,\cdot)\theta$.

2.2 Model order reduction using the method of weighted residuals

To derive a finite-dimensional system that captures the dominant dynamics of the distributed parameter system of Eq.1, we first expand the solution in an infinite series in

terms of an orthogonal and complete set of basis functions, $\{\phi_k(r)\colon r\in [0,r_{max})\}$, as $n(r,t)=\Sigma_{k=1}^{\infty} a_k(t)\phi_k(r)$, where $a_k(t)$ are time-varying coefficients. Substituting the series expansion into Eq.1, multiplying the population balance with appropriate weighting functions, integrating over the entire particle size spectrum, and finally truncating the series expansion up to order N (keeping the first N equations), we obtain the following approximate finite-dimensional system:

$$\dot{z}_i = h_i(z_1,z_2)+m_i(z_1,z_2)\left[u_i^k + f_{ai}^k\right]+d_i(z_1,z_2)\theta, i=1,2 \quad (2)$$

where $z_1=a^N=[a_1^N \ \dots \ a_N^N]^T$, $z_2=x_N$, x_N and a_k^N are the approximations of x and a_k, respectively, obtained by an *N-th* order truncation of the infinite series, h_i, m_i, d_i are nonlinear functions whose explicit forms are omitted for brevity. The asymptotic validity of the approximation can be justified using standard results from perturbation theory (Christofides, 2002). In the next section, we present the design methodology for the FDI-FTC structure on the basis of the reduced-order system of Eq.2, and then discuss in Section 4 its implementation on the infinite-dimensional system of Eq.1.

3. Methodology for integrated robust FDI and fault-tolerant control

3.1 Bounded Lyapunov-based controller synthesis

The objectives of this step are to: (a) synthesize, for each control actuator configuration, a family of feedback controllers that enforce (in the absence of faults) constraint satisfaction and robust stability with an arbitrary degree of asymptotic attenuation of the effect of uncertainty on the closed-loop system, and (b) explicitly characterize the robust stability region associated with each configuration in terms of the constraints and the size of the uncertainty. To this end, we first design the controllers assuming the availability of accurate full-state measurements and then analyze the effects of measurement errors to derive sufficient conditions for closed-loop stability. Specifically, let V_i be a robust control Lyapunov function (Freeman and Kokotovic, 1996) for the *i-th* subsystem in Eq.2, and consider the feedback control law given by:

$$u_i^k(z)=-\left(\frac{\alpha_i(z)+\sqrt{\alpha_i^2(z)+\left(u_{i,\max}^k\beta_i^k(z)\right)^4}}{\left(\beta_i^k(z)\right)^2\left[1+\sqrt{1+\left(u_{i,\max}^k\beta_i^k(z)\right)^2}\right]}\right)L_{gi}V_i, i=1.2, k(t)\in\{1,2,...,L\} \quad (3)$$

where $\alpha_i = L_{hi}V_i + (\rho_i|z_i|+\chi_i\theta_b|L_{di}V_i|)(|z_i|/(|z_i|+\phi_i^k))$, $\beta_i^k = |L_{mi}V_i|$, $L_{hi}V_i$, $L_{mi}V_i$ and $L_{di}V_i$ are Lie derivatives of V_i, $z=[z_1\ z_2]^T$ and $\rho_i>0$, $\chi_i>1$, $\phi_i^k>0$ are adjustable parameters. It can be verified that each controller satisfies the control constraints within the state space region described by the set $I\Gamma_i^k:=\{z\in R^{n+N}\colon \alpha_i(z,\rho_i\chi_i\phi_i^k,\theta_b)\le u_{i,\max}^k\beta_i^k(z)\}$. To account for measurement errors, let $z_m(t)=z(t)+e(t)$ be the actual state measurement used to implement the controllers, where $e(t)$ is a bounded time-varying function ($|e(t)|\le\mu$) that captures the measurement error. From the continuity of the control laws, it follows that the error in the implemented control action is also bounded, i.e., $|u_i^k(z_m)-u_i^k(z)|\le M_i(\mu)$ for some class $\mathcal{K}$ function M_i (a function is said to be of class $\mathcal{K}$ if it is non-decreasing and vanishes at zero; see Khalil (2002)). To ensure that the control constraints continue to be satisfied, evolution of the closed-loop states must be restricted within the sub-region defined by $\Lambda_i^k:=\{z\in I\Gamma_i^k\colon |u_i^k(z)|\le u_{i,\max}^k - M_i(\mu)\}$. Specifically, it can be shown using a Lyapunov argument that if $z(t)\in\Lambda_i^k$ and $f_{ai}^k(t)=0$, for some t and i, then for sufficiently small ϕ_i (and/or sufficiently large χ_i), the time-derivative of V_i along the trajectories of the *i-th* fault-free closed-loop subsystem of Eqs.2-3 satisfies:

$$\dot{V}_i(z(t)) \le -\gamma_i V_i(z_i(t)) + \sigma_i(\phi_i) + \Xi_i(\mu) := B_i(z_{m,i}(t), \gamma_i, \phi_i, \mu) \quad (4)$$

for some constant $\gamma_i>0$ and some class $\mathcal{K}$ functions σ_i and Ξ_i. Furthermore, if $f^k_{a,i}=0$ for all i and the closed-loop system is initialized within any invariant subset of $\Lambda^k := \Lambda^k_1 \cap \Lambda^k_2$, e.g., $z(0) \in \Omega^k_s(\theta_b, u^k_{\max}) := \{z \in \Lambda^k : V_1(z)+V_2(z) \le \delta_s\}$ for some $\delta_s>0$, then each controller in the *k-th* control configuration (implemented with measurement errors) satisfies the control constraints and drives the trajectory of its corresponding subsystem in finite-time into a small neighborhood around the desired steady-state where it remains confined for all future times. The size of this terminal set can be made small by proper tuning of the controller parameters, but is ultimately constrained by the size of the measurement error μ. The set $\Omega^k_s(\theta_b, u^k_{\max})$ therefore represents an estimate of the robust stability region starting from where practical closed-loop stability is guaranteed.

3.2 Robust fault detection and isolation scheme

Since the inequality of Eq.4 provides an explicit characterization of the expected fault-free behavior of each subsystem, it can be used to derive rules for FDI. Specifically, any breach of the given bound while $z(t) \in \Omega^k_s(\theta_b, u^k_{\max})$ is an indicator of a fault in the corresponding actuator. However, since verifying this requires monitoring of the state, which is known with only limited accuracy, we introduce two key modifications to safeguard against possible false alarms due to measurement errors. The modifications include (1) limiting the monitoring region to an appropriate subset of the stability region, i.e., $\Omega^k_c \subset \Omega^k_s$ where $z_m(t) \in \Omega^k_c \Rightarrow z(t) \in \Omega^k_s$ (this is needed to be able to infer the location of z in the state-space from the available measurement, z_m) and (2) enlarging the bounds on the dissipation of the Lyapunov functions. Specifically, by exploiting the continuity of both V_i and its time-derivative in their arguments, one can show that given any desired subset $\Omega^k_c \subset \Omega^k_s$, there always exists a class $\mathcal{K}$ function $X_i > \Xi_i$ such that:

$$\dot{V}_i(z_m(t)) \le -\gamma_i V_i(z_{m,i}(t)) + \sigma_i(\phi_i) + \mathrm{X}_i(\mu) := B^*_i(z_{m,i}(t), \gamma_i, \phi_i, \mu) \quad (5)$$

Note that any breach of Eq.5 while $z_m(t) \in \Omega^k_c$ implies that Eq.4 is breached while $z(t) \in \Omega^k_s$. Therefore, by using the time-derivative of V_i as a residual, and the right hand side of Eq.5 as an alarm threshold, a fault in the *i-th* actuator can be declared at time T_d if $z_m(T_d) \in \Omega^k_c$ and $\dot{V}_i(T_d) > B^*_i(z_{m,i}(T_d), \gamma_i, \phi_i, \mu)$. Comparing Eqs.4 and 5, wee see that the alarm threshold is essentially increased to ensure that any discrepancy between the actual and fault-free residuals exceeds what can be accounted for by measurement errors alone, and thus is due solely to faults.

3.3 Control system reconfiguration

Following FDI, the supervisor needs to determine based on the available (inaccurate) measurements which backup configuration can be activated to maintain robust closed-loop stability. To this end, consider the closed-loop system of Eqs.2-3 with $k(0)=j$, $z(0) \in \Omega^k_s$ and let T_d be the earliest time that a fault is detected in the *i-th* actuator. Then choosing the configuration $k(t)=\nu$ for $t \ge T_d$ where $z_m(T_d) \in \Omega^\nu_c(\theta_b, u^\nu_{\max})$ ensures that the origin of the closed-loop system remains practically stable. Recall from the definition of Ω^k_c in Section 3.2 that $z_m(t) \in \Omega^k_c \Rightarrow z(t) \in \Omega^k_s$; therefore, the switching logic ensures that the fallback actuator configuration that is activated and implemented following FDI is one whose stability region contains the state at the time of switching. In the event that multiple configurations satisfy this stability condition, additional performance criteria can be introduced to aid in choosing the appropriate configuration.

4. Implementation issues

While the FDI-FTC structure presented above is designed on the basis of an approximate reduced-order model of the particulate process, it can be shown using regular perturbation techniques (e.g., see El-Farra and Giridhar (2008)) that it continues to enforce the desired stability and fault-tolerance properties in the infinite-dimensional system of Eq.1 provided that the reduced-order model is of a sufficiently high-order and the FDI rules are appropriately modified to account for the approximation errors. Specifically, the closeness of solutions between the approximate and infinite-dimensional systems can be exploited to obtain new alarm thresholds that can be chosen as close as desired to the bounds obtained for the approximate system for sufficiently large N. Furthermore, the problem when only measurements of the principal moments of the particle size distribution and the continuous-phase variables are available for on-line implementation can be addressed by incorporating a suitable nonlinear state observer to generate estimates of z from the measured outputs, which are then used to implement the controllers, FDI and reconfiguration laws. The observer must be designed to ensure sufficiently fast convergence of the state estimates. This property is necessary to ensure that the same diagnostic and fault-tolerance properties obtained under full-state feedback control can be practically recovered under output feedback control by proper choice of the observer tuning parameters (see El-Farra and Giridhar (2008) for an example observer design that satisfies this requirement).

5. Case study: Application to a continuous crystallizer with a fines trap

To demonstrate the implementation of the proposed FDI-FTC methodology, we consider in this section an example involving an isothermal continuous crystallizer with a fines trap. The trap is used to remove small crystals and increase the mean crystal size. The process model and parameters can be found in Christofides (2002). The control objective is to regulate the crystal and solute concentrations at desired set points by manipulating the flow rate of suspension through the fines trap and the inlet solute concentration in the presence of (1) actuator constraints and faults, (2) time-varying errors in the crystal and solute concentration measurements, and (3) time-varying uncertainties in the nucleation rate and the density of crystals. To ensure actuator fault-tolerance, a backup actuator is assumed to be available for each manipulated input for use in the event that the primary actuator fails. Using the method of moments and approximating the size distribution function using a Laguerre polynomial expansion to close the set of moment equations (Lei et al., 1971), an approximate fifth-order nonlinear ODE system describing the evolution of the first four moments of the crystal size distribution and the solute concentration is obtained and used for the synthesis of the FDI-FTC structure which is then implemented on a sufficiently high-order discretization of the process model obtained using a finite-difference scheme. The model derivation and controller synthesis details are omitted due to space limitations.

To detect and isolate faults during crystallizer operation, we consider the following two residual signals. The first is dedicated to faults in the first manipulated input (flow rate through fines trap) and captures the discrepancy between the actual and fault-free evolutions of z_1 in Eq.2 (the fault-free behavior is estimated from Eq.5). The second residual is dedicated to faults in the second manipulated input (solute feed concentration), and captures the discrepancy between the actual and fault-free evolutions of z_2 in Eq.2. To account for the effects of uncertainties, state estimation, measurement and model reduction errors in the absence of faults, the alarm thresholds for FDI are set at $\delta_1=\delta_2=0.012$ by properly tuning the controllers and observers. To

demonstrate how the integrated FDI-FTC scheme works, the process is initialized using the healthy controllers which, as can be seen from the solid profiles in the top plots in Fig.1, successfully steer the crystal (left) and solute (right) concentrations near the desired set points very quickly. At *t=15 hr*, a fault is introduced in the actuator manipulating the suspension flow rate through the fines trap (valve stuck in the closed position; see dashed line in the left middle plot of Fig.1). The bottom plots in Fig.1 depict how this failure is detected and isolated since it causes the corresponding residual, r_1, to cross the specified threshold at *t=15.5 hr*, while practically not influencing r_2. Following FDI in the first controller, the supervisor switches to a backup actuator that maintains the closed-loop outputs near their desired set-points, as can be seen by the solid profiles in the top plots in Fig.1. After recovering from the first fault, a second failure in the actuator manipulating the solute feed concentration occurs at *t=25 hr* as shown by the dashed line in the right middle plot of Fig.1. By examining the residual profiles, it can be seen that this failure is detected and isolated in a timely manner since it causes r_2 to cross the threshold at *t=25.07 hr*, while r_1 remains within its threshold limit. Following the detection and isolation of the second actuator failure, the supervisor switches to the available backup actuator to preserve robust closed-loop stability. The dashed profiles in the top two plots of Fig.1 show the destabilizing effect of the two consecutive faults when no corrective action is taken.

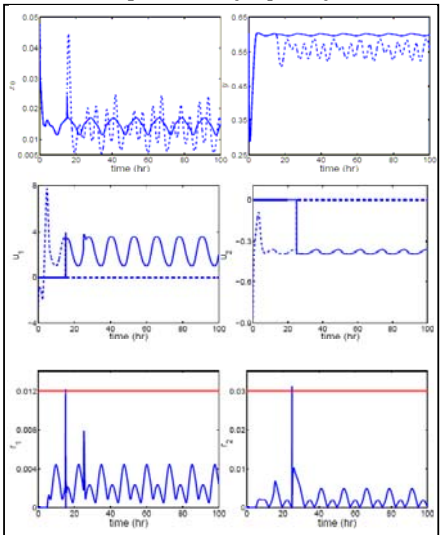


Fig.1. Evolution of the crystallizer's controlled outputs (top), manipulated inputs (middle) and FDI residuals (bottom) under two consecutive failures in the control actuators and subsequent reconfiguration.

References

P.D. Christofides, 2002, *Model-Based Control of Particulate Processes*, Netherlands: Kluwer Academic Publishers.

N.H. El-Farra and A. Giridhar, 2008, Detection and management of actuator faults in controlled particulate processes using population balance models, *Chem. Eng. Sci.*, 63, 1185-1204.

R.A. Freeman and P.V. Kokotovic, 1996, *Robust Nonlinear Control Design: State-Space and Lyapunov Techniques*, Boston: Birkhauser.

H.K. Khalil, 2002, *Nonlinear Systems*, Upper Saddle River, NJ: Prentice Hall.

P. Larsen, D. Patience, and J.B. Rawlings, 2006, Industrial crystallization process control, *IEEE Contr. Syst. Mag.*, 26, 70-80.

S.J. Lei, R. Shinnar, and S. Katz, 1971, The stability and dynamic behavior of a continuous crystallizer with a fines trap, *AIChE J.*, 17, 1459-1470.

M.J. Park, M.T. Dokucu, and F.J. Doyle, 2004, Regulation of the emulsion particle size distribution to an optimal trajectory using partial least squares model-based predictive control, *Ind. Eng. Chem. Res.*, 43, 7227-7237.

D. Semino and W.H. Ray, 1995, Control of systems described by population balance equations-II: emulsion polymerization with constrained control action, *Chem. Eng. Sci.*, 50, 1825-1839.

G. Zhang and S. Rohani, 2003, On-line optimal control of a seeded batch cooling crystallizer, *Chem. Eng. Sci.*, 58, 1887-1896.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Iterative Feedback Tuning of State Space Control Loops with Observers given Model Uncertainty.

Jakob K. Huusom,[a] Niels K. Poulsen,[b] Sten Bay Jørgensen[a]

[a]*CAPEC, Department of Chemical and Biochemical Engineering, Technical University of Denmark, Building 229, DK-2800 Lyngby, Denmark, jkh@kt.dtu.dk;sbj@kt.dtu.dk*
[b]*Department of Informatics and Mathmatical Modelling, Technical University of Denmark, Buliding 321, DK-2800 Lyngby, Denmark, nkp@imm.dtu.dk*

Abstract

Direct tuning is investigated as an alternative to iterative model estimation and control design for state space systems in case of unsatisfactory loop performance. Direct tuning of the model parameters in the feedback control and the observer design by Iterative Feedback Tuning, optimize loop performance when the initial designs are based on an uncertain model estimate. The certainty equivalence design is kept throughout the iterations when the model parameters is updated. This methodology constitutes a promising algorithm when developing tools for online tuning of state space control systems.

Keywords: LQG control, Iterative Feedback Tuning.

1. Introduction

The need for optimal process operation has rendered methods for optimization of control loop parameters an active research area. Much attention has been directed on performing control oriented system identification, which implies model estimation from closed loop data [1,2]. Optimizing the parameters in a control loop is an iterative procedure since the data from one experiment will depend on the current controller, and repeated iteration is necessary for the loop performance to converge to a minimum. An alternative would be a direct data driven approach to tuning without utilizing a model estimate. Data driven tuning methods have mainly been reported for systems given in transfer function form. The most established method is Iterative Feedback Tuning [3]. This method optimizes the closed loop performance by adjusting the control parameters through a gradient based scheme. The gradient of the cost function is replaced by an unbiased estimate evaluated from special closed loop experiments. Direct tuning is often computationally less demanding than identification and model based control design. Direct tuning methods can be used when insufficient knowledge of the model structure limits nominal performance, where the system is tuned based on the certainty equivalence principle.

This paper investigates the use of the direct tuning method, Iterative Feedback Tuning, for optimization of the parameters in the feedback loop and the state observer for a control loop based on a state space system description. Based on the certainty equivalence principle, analytical solutions for optimal values of the feedback and Kalman gain exist. This assumption renders the loop performance sensitive to model

errors and bias. Direct controller tuning of the uncertain model parameters may serve as an interesting alternative, when fine tuning a control loop is desired or when degrading loop performance is observed.

This paper is organized as follows. In the following section a short introduction to the system and control loop description is given. Section 3 discusses optimal control and tuning. In Section 4 an illustrative simulation example is given before final conclusions eventually are drawn.

2. The state space control loop

Given the following linear, discrete time, single input/single output, time-invariant system description:

$$\begin{aligned} \boldsymbol{x}_{t+1} &= \boldsymbol{A}\boldsymbol{x}_t + \boldsymbol{B}u_t + \boldsymbol{e}_t^P, & \boldsymbol{e}_t^p &\in N(0, \boldsymbol{P}_{e^P}) \\ y_t &= \boldsymbol{C}\boldsymbol{x}_t + e_t^m, & e_t^m &\in N(0, \sigma^2_{e^m}) \end{aligned} \tag{1}$$

Where $\boldsymbol{x}_t$ represents the system states, u_t is the manipulated variable and y_t is the measurement at discrete time instants. $\boldsymbol{e}^p{}_t$ represents process noise and $e^m{}_t$ is measurement noise. Both $\boldsymbol{e}^p{}_t$ and $e^m{}_t$ will be assumed white and the cross correlation between these types of noise will be assumed zero. It is desired to control this system using the feedback law:

$$u_t = -\boldsymbol{L}\boldsymbol{x}_t + Mr_t \tag{2}$$

Where $\boldsymbol{L}$ is a constant feedback gain matrix and M is a controller gain for the reference signal. Since the exact values of the states are not known, an observer is used to generate the state estimates used in the control law based on measurements of the process output and a process model. The observer has the form of the predictive Kalman filter with the constant gain matrix $\boldsymbol{K}$, assuming stationary conditions.

$$\hat{\boldsymbol{x}}_{t+1|t} = \hat{\boldsymbol{A}}\hat{\boldsymbol{x}}_{t|t-1} + \hat{\boldsymbol{B}}u_t + \boldsymbol{K}(y_t - \hat{\boldsymbol{C}}\hat{\boldsymbol{x}}_{t|t-1}) \tag{3}$$

The structure of the state space feedback loop consisting of equation (1) and (3) is shown in Figure 1.

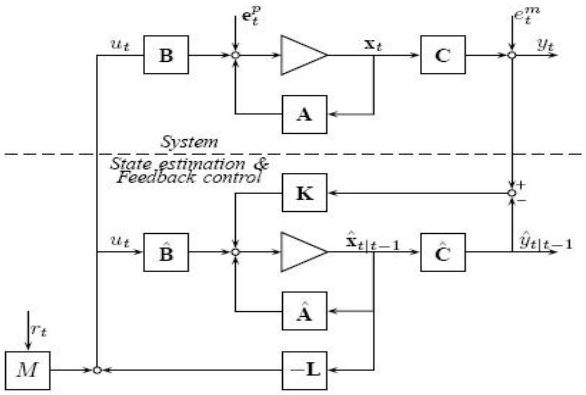


Figure 1. Block diagram of the closed loop state space system with state observer

In order to ensure a static gain from the reference to the estimated output of unity, the following requirements can be derived:

$$M = [\hat{\boldsymbol{C}}(\boldsymbol{I} - \hat{\boldsymbol{A}} + \hat{\boldsymbol{B}}\boldsymbol{L})^{-1}\hat{\boldsymbol{B}}]^{-1} \tag{4}$$

Introducing the state estimation error: $\tilde{\boldsymbol{x}}_t = \boldsymbol{x}_t - \hat{\boldsymbol{x}}_t$, provides a convenient description with a more clear distinction between feedback control and state estimation dynamics [4]. If the system (1) is stabilizable and detectable a set {***L,K***} exists which renders the system stable [5]. Hence if optimal values for the feedback and Kalman filter gains are used stability is guaranteed.

3. Optimal control and tuning

Optimal values for both the observer gain ***K*** and the feedback gain ***L*** exist and have known analytical solutions if the true system is known [6]. The optimal, stationary value for the gain matrix in the Kalman filter can be evaluated based on the process model and information of the noise intensities, from an algebraic Riccati equation. The optimal controller gain will also depend on the optimization criterion. In this paper the control design will minimize the value of a cost function for the loop performance for a single input/single output system:

$$F(y,u) = \frac{1}{2N}\sum_{t=1}^{N} y_t^2 + \lambda u_t^2 \tag{5}$$

Where λ determines the weighting between the penalty on the output and the control. For optimal tracking the output is replaced by the tracking error in the cost function. The optimal Linear Quadratic Gaussian controller (LQG) produces an optimal feedback gain which minimizes the quadratic cost function:

$$F_{LQG}(y,u) = \frac{1}{2N}\sum_{t=1}^{N} \hat{\boldsymbol{x}}_t^T \boldsymbol{Q}_R \hat{\boldsymbol{x}}_t + \lambda u_t^2 \tag{6}$$

using the linear system description in (3) and assuming that *N* approaches infinity. As for the observer gain, the solution is given by an algebraic Riccati equation. In case $\boldsymbol{Q}_R = \boldsymbol{C}^T\boldsymbol{C}$ the cost function (6) is equivalent to (5).

In absence of an accurate process model, direct tuning of the closed loop can provide the optimal control parameters. It has been shown in [7] that Iterative Feedback Tuning can be applied to a state space control system with state observer. The iterative procedure converges towards the optimal values for the two gains, when the observer is constructed from full model information. This result has more academic than practical interest since analytical solutions are known when a model is available. It is on the other hand interesting to investigate the potential for using direct tuning in case of model uncertainty. Parametric model uncertainty will give the following equations for the state and the estimation error:

$$\begin{aligned}
\boldsymbol{x}_{t+1} &= (\boldsymbol{A}-\boldsymbol{BL})\boldsymbol{x}_t + \boldsymbol{BL}\tilde{\boldsymbol{x}}_t + \boldsymbol{B}Mr_t + \boldsymbol{e}_t^P \\
\tilde{\boldsymbol{x}}_{t+1} &= (\hat{\boldsymbol{A}}-\boldsymbol{K}\hat{\boldsymbol{C}})\tilde{\boldsymbol{x}}_t - \boldsymbol{K}\boldsymbol{e}_t^m + \boldsymbol{e}_t^P + \boldsymbol{\Delta} \\
\boldsymbol{\Delta} &= [(\boldsymbol{A}-\hat{\boldsymbol{A}}) - \boldsymbol{K}(\boldsymbol{C}-\hat{\boldsymbol{C}})]\boldsymbol{x}_t + (\boldsymbol{B}-\hat{\boldsymbol{B}})[Mr_t - \boldsymbol{L}(\boldsymbol{x}_t - \tilde{\boldsymbol{x}}_t)]
\end{aligned} \tag{7}$$

Only if the true process and the model estimate are equivalent will the term $\boldsymbol{\Delta}$ be zero. In this case the optimal closed loop performance can be achieved by solving the Riccati equations for the two gains. In case of parametric uncertainty, a certainty equivalent control design will not produce optimal closed loop performance. Hence either model re-estimation with a subsequent update of the feedback and observer gain or direct tuning can be used for optimization. In [8] direct tuning of the two gains was performed using Iterative Feedback Tuning in order to investigate to which extent adjusting the gains could compensate for the erroneous model parameters in the state estimator. The loop performance was improved iteratively, and in some cases the performance degradation caused by an erroneous parameter could be completely compensated. This paper will move further and tune the model parameters in the state estimator by direct tuning. Given an updated set of model parameters, new optimal feedback and Kalman filter gains are recalculated from the Riccati equations.

4. Case study

In order to illustrate the potential of using direct tuning in form of the Iterative Feedback Tuning method the discrete time, scalar, state space system with state observer (8) is investigated. This system is too simple to have any industrial relevance but the objective of this example is to show the principles in the tuning method and it's ability to converge to optimality. The general method described in the paper concerns higher order linear models too.

$$\begin{aligned}
x_{t+1} &= 0.98x_t + 0.02u_t + e_t^P, \qquad e_t^P \in N(0,1^2) \\
y_t &= 1x_t + e_t^m, \qquad\qquad\quad e_t^m \in N(0,0.01^2)
\end{aligned} \tag{8}$$

In the optimization criterion λ= 0.001 is used. In the case when only the noise variance is unknown and all other parameters in the observer are correct, only the Kalman filter gain needs to be tuned. In case any of the parameters *a*, *b* or *c* in model estimate are erroneous, it will affect both the values of the Feedback and Kalman gains and the state estimate. In the following, three experiments are performed where one of the parameters *a*, *b* or *c* in the state estimate is wrong respectively. By direct tuning of the parameter in question and subsequent update of the Feedback and Kalman gains, to ensure optimality by certainty equivalence in each iteration, the closed loop performance is optimized. Fig. 2, 3 and 4 show the results from 15 iterations of the tuning, when the erroneous parameters is either too large or too small. All experiments are able to converge to the optimal solution in 5-10 iteration which is very good when tuning is conducted for the disturbance rejection case. The method also allows tuning of all the model parameters simultaneously and for multivariable systems, but these results have been omitted due to the spatial constraints of this contribution. It is however ongoing work.

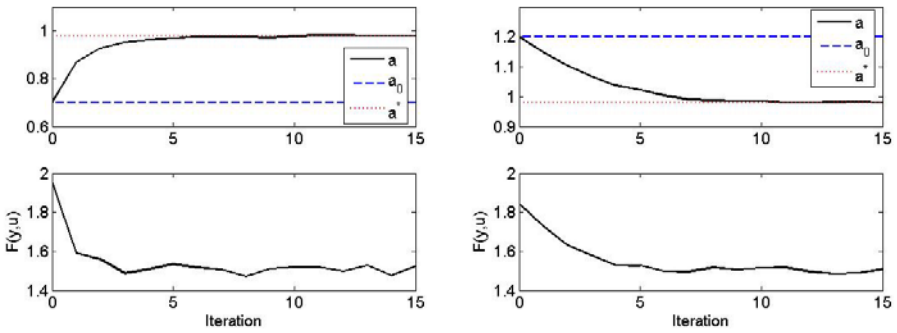


Figure 2. 15 iterations by the tuning method when the parameter *a* in the observer is erroneous. Both the iteration of the parameter value and the corresponding closed loop performance are given.

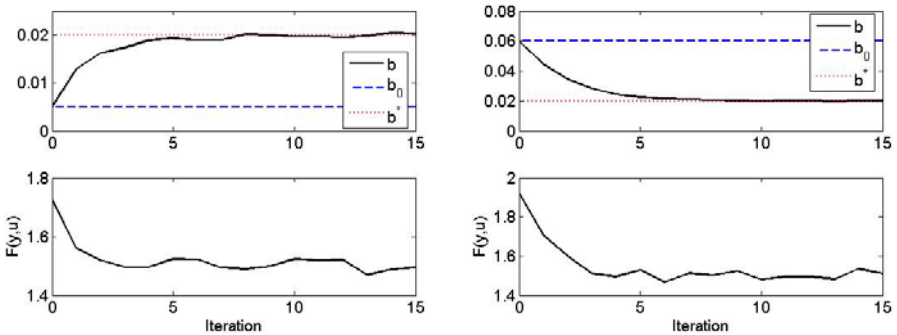


Figure 3. 15 iterations by the tuning method when the parameter *b* in the observer is erroneous. Both the iteration of the parameter value and the corresponding closed loop performance are given.

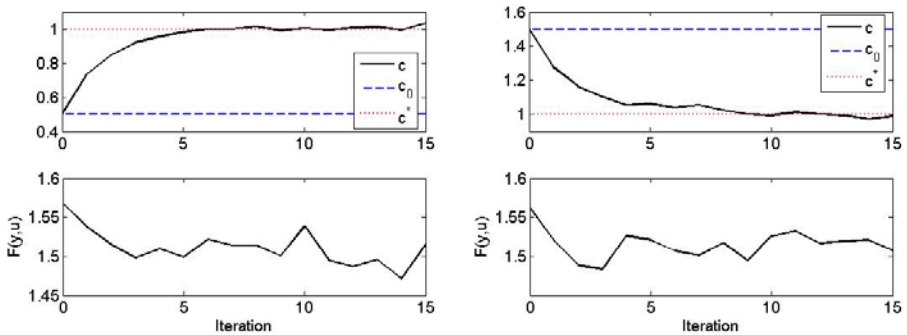


Figure 4. 15 iterations by the tuning method when the parameter *c* in the observer is erroneous. Both the iteration of the parameter value and the corresponding closed loop performance are given.

The cost function gradient estimate used in a Newton scheme by the Iterative Feedback Tuning method, utilizes a transfer function description of the state space control loop on Fig. 1 which is given in [7]. This estimate is constructed by filtering closed loop

input/output data through a filter which contains the gradient of the controller with respect to the tuning parameters. Since the feedback and the feed forward controller in the transfer function description of the state space control loop are functions of both the model estimate and the gain matrices, the partial derivatives of the optimal gains are needed with respect to the tuning parameters. These have been obtained by a first order forward difference approximation in the results presented here.

The results presented are by far superior to the results in [8] where gains were tuned while the erroneous parameters in the state estimator were kept constant. The increased complexity of tuning the system parameters and adapting the gains seems to be rewarded since convergence of the model estimate to the true system leads to optimality for the closed loop performance.

5. Conclusions

Direct tuning in form of Iterative Feedback Tuning has been used to adjust the system model parameters in the observer in a state space control loop in order to minimize the closed loop performance cost. When the model in the control and the observer design is iteratively updated so are the Feedback and Kalman gains in order to ensure optimality by certainty equivalence in the design. The tuning parameters converged to the true system hence the achieved closed loop performance converged to optimality.

Acknowledgements

This project is funded by The Danish Council for Technology and Innovation.

References

1. R. J. P. Schrama, Accurate identification for control: The necessity of an iterative scheme. IEEE Transactions on automatic control, vol. 37, (7), pp. 991-994, 1992.
2. M. Gevers, A decade of progress in iterative process control design: from theory to practice, Journal of process control, vol. 12,(4), pp. 519-531, 2002.
3. H. Hjalmarsson, M. Gevers, S. Gunnarsson, and O. Lequin, Iterative feedback tuning: Theory and applications, IEEE Control Systems Magazine, vol. 18, (4), pp. 26-41, 1998.
4. K. J. Åström, Introduction to Stochastic Control Theory, Academic Press, 1970.
5. H. Kwakernaak and R. Sivan, Linear Optimal Control Systems. John Wiley & Sons,1972.
6. B. D. O. Anderson and J. B. Moore, Optimal Control. Linear Quadratic Methods, Prentice-Hall Int. Ed., 1989.
7. J. K. Huusom, N. K. Poulsen and S. B. Jørgensen, Data Driven Tuning of State Space Controllers with Observers. Submited for ECC 2009
8. J. K. Huusom, N. K. Poulsen and S. B. Jørgensen, Data Driven Tuning of State Space Control loops with unknown state information and model uncertainty. Accepted for ESCAPE19 2009.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Real Time Optimization (RTO) with Model Predictive Control (MPC)

Glauce De Souza[a], Darci Odloak[b], Antônio C. Zanin[c]

[a]*Department of Chemical & Materials Engineering, The University of Auckland, 20, Symonds Street 8th floor, City Campus 1042, Auckland, New Zealand*
[b]*LSCP/CESQ - Department of Chemical Engineering - University of São Paulo, Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil. Brazil.*
[c]*PETROBRAS, Brazilian Oil Company*

Abstract:

This paper studies a simplified methodology to integrate the real time optimization of a continuous system into the model predictive controller in the one layer strategy. The gradient of the economic objective function is included in the cost function of the controller. One of the control objectives is to zero the reduced gradient of the economic objective while maintaining the system outputs inside their zones. Optimal conditions of the process at steady state are searched through the use of a rigorous nonlinear process model, while the trajectory to be followed is predicted with the use of a linear dynamic model that can be obtained through a plant step test. Moreover, the reduced gradient of the economic objective is computed taking advantage of the predicted input and output trajectories. The main advantage of the proposed strategy is that the resulting control/optimization problem can be solved with a quadratic programming routine at each sampling step. Simulation results show that the approach proposed here is comparable to the strategy that solves the full economic optimization problem inside the MPC controller where the resulting control problem becomes a nonlinear programming with a high computer load.

Keywords: Real Time Optimization, Model Predictive Control, Fluid Catalytic Converter, Non-linear programming

1. Introduction

Typically, the RTO application optimizes the process operating conditions and updates the set-points to local MPCs, which are based on linear dynamic models. A steady-state RTO may not be sufficient if, for instance, the set of active constraints with significant economic importance changes frequently. For this type of process, it is more suitable to use dynamic optimization with a non-linear model, which may be achieved using dynamic RTO (DRTO) or non-linear MPC with economic objective (Kadam et al., 2003; Tosukhowong et al., 2004). However, the dynamic optimization of a complex process is still not achievable with the available computing resources (Lu, 2003). Aske et al. (2008) propose the use of a coordinator MPC, which explores the degrees of freedom of the system in order to maximize throughputs at the bottlenecks of the system. Economic optimization studies for

oil refining and chemicals production systems have proven to be very beneficial (Georgiou et al., 1997: Rotava & Zanin, 2005). Similar algorithms to the ones of the petrochemical industry were applied by Mercangoz & Doyle III (2007) to the pulp and paper manufacturing system. To maximize the economic benefit while maintaining the desired product quality may be hard in a continuous production environment with frequent changes (Engell, 2007). Here, the focus is on the integration of real time optimization and control of continuous processes through the strategy called one-layer approach (Zanin et al. 2002) or full optimizing control (Rolandi & Romagnoli, 2005).

In process plants, the economically optimal operation is usually addressed by a two-layer structure (Marlin & Hrymak, 1996). In the upper layer, the operating point or steady state of the plant is optimized based upon a rigorous nonlinear stationary plant model. This two layer structure has some drawbacks, as optimization is only performed intermittently at a low sampling rate and inconsistencies may arise from the use of different process models at the two layers. These issues are partly addressed by schemes in which the economic optimization is integrated within a linear MPC controller. Typical sampling times for the RTO layer in a two layer approach are of the order of several hours (Engel, 2007) while the control layer provides tracking and disturbance rejection on shorter time-scales from seconds to minutes.

Zanin et al. (2002) included the economic objective computed at the predicted steady state as an additional term of the cost function of the controller. Here it is presented a simplified strategy to integrate the RTO problem into the linear model predictive controller in the one layer approach. An approximation of the gradient of the economic function is included in the MPC cost function with an appropriate weight. Simulation results obtained with the nonlinear model of Moro & Odloak (1995) are used to compare the proposed approach to the approach of Zanin et al. (2002), and show that the results of the two approaches may be comparable.

2. The one-layer RTO/MPC controller

Zanin et al. (2002) propose to integrate the real time optimization of continuous processes into the model predictive controller by solving at each sampling step k the optimization problem:

$$\min_{u(k+i);\, i=0,\ldots,m-1} \sum_{j=1}^{p} \left\| W_1 (y(k+j) - r) \right\|_2^2 + \sum_{j=0}^{m-1} \left\| W_2 \Delta u(k+j) \right\|_2^2 + W_3 f_{eco} \;\; u(k+m), y(k+\infty) \qquad (1)$$

subject to

$$h\;\; u(k+m), y(k+\infty) \;\; = 0 \qquad (2)$$

$$u_{\min} \le u(k+j) \le u_{\max} \quad j = 0,1,\ldots,m-1 \qquad (3)$$

$$\Delta u_{\max} \le \Delta u(k+j) \le \Delta u_{\max} \quad j = 0,1,\ldots,m-1 \qquad (4)$$

$$y(k+j) = g_j \;\; x(k), \Delta u_k \;\; , \quad j = 1,\ldots,p \qquad (5)$$

where

p and m are the output and input horizons, respectively; $u(k+j)$ is the control input computed at time k to be applied at time step $k+j$; $y(k+j)$ is the output prediction at time step $k+j$; r is the desired value of the output; $\Delta u(k+j) = u(k+j) - u(k+j-1)$; W_1, W_2 and

W_3 are positive weights; f_{eco} is the economic objective corresponding to the predicted steady state of the process; $y(k+\infty)$ is the process output steady state computed with Eq. (2).

Eq. (5) corresponds to the output prediction performed with the linear dynamic model of the system, and $\Delta u = \begin{bmatrix} \Delta u(k) & \Delta u(k+1) & \cdots & \Delta u(k+m-1) \end{bmatrix}^T$ is the control sequence.

In the RTO/MPC algorithm defined above, the outputs are controlled by zones, which means that the desired values *r* of the controlled outputs do not have fixed values.

Since in the general case, f_{eco} and *h* are nonlinear functions of the inputs and outputs of the system, the problem solved by the controller is a NLP while the problem solved in the conventional MPC is a QP. Although, the recent progress in the algorithms for the numerical solution of NLP problems, one can expect some difficulties in solving the RTO/MPC problem for large dimension systems. If the solver of the one layer RTO/MPC problem fails to produce a converged solution within the sampling period of the controller, no control action will be available to be implemented and the system will remain in open loop, which is highly undesirable. Thus, there is interest in developing an alternative solution to the one layer RTO/MPC problem that preserves some of the advantages of the one layer approach, but without delaying the computation of the optimizing control action.

3. The simplified one-layer RTO/MPC

Consider a multivariate system with *ny* outputs and *nu* inputs. At any time step *k*, suppose that the stationary prediction of the controlled output related to *u* is $\hat{y}$, and the economic function associated with this predicted steady state can be represented as follows:

$$F = f_{eco}(u, \hat{y}) \tag{6}$$

Assuming that the vector that represents the control action is changed to $u + \Delta\bar{u}$, the first order approximation to the gradient of the economic function can be represented as follows:

$$\varsigma_{u+\Delta u} = \left[\frac{\partial F}{\partial \hat{y}} K_p + \frac{\partial F}{\partial u} \right] + \left[K_p{}^T \frac{\partial^2 F}{(\partial \hat{y})^2} K_p + \frac{\partial^2 F}{\partial u \partial \hat{y}} K_p + \frac{\partial F}{\partial \hat{y}} \left(\frac{\partial^2 \hat{y}}{\partial u^2} \right) + K_p \frac{\partial^2 F}{\partial \hat{y} \partial u} + \frac{\partial F}{\partial u^2} \right] \Delta\bar{u}$$

where $K_p = \partial \hat{y} / \partial u$

The above equation can be represented as follows:

$$\varsigma_{u+\Delta u} = d + G\Delta\bar{u} \tag{7}$$

where $\Delta\bar{u} = u(k+m-1) - u(k-1)$ is the total move of the input vector.

Then, the simplified real time optimization integrated into the linear MPC controller (SRTO/MPC) proposed here has the following control objective:

$$\min_{u(k+i);\, i=0,\ldots,m-1} J = \sum_{j=1}^{p} \left\| W_1 (y(k+j) - r) \right\|_2^2 + \sum_{j=0}^{m-1} \left\| W_2 \Delta u(k+j) \right\|_2^2 + \left\| W_3 \varsigma_{u+\Delta u}^T \right\|_2^2 \tag{8}$$

subject to: (3) to (5) and (7)

The integration of the economic function into the controller can push one or more outputs towards their bounds. To consider this effect in the calculation of the gradient of the economic objective, the following algorithm is followed:

At each sampling time k, vector d and matrix G are computed using the predicted steady state corresponding to the control input implemented at time step k-1. Then, the problem defined in (8) is initially solved assuming that no controlled variables will be constrained. Next, the resulting new predicted values for the controlled variables are checked in order to confirm that there are no violations of the constraints. If the prediction values of the outputs do not violate any constraint, the solution obtained through the solution to the problem represented in (8) is implemented in the real system. If the prediction of one of the outputs is such that $\hat{y}_i \geq y_{\max,i}$ or $\hat{y}_i \leq y_{\min,i}$. Then, the problem defined in (8) is solved again with the following additional constraint.

$$\Delta y_i = \sum_{j=1}^{nu} K_{p_{i,j}} \Delta \bar{u}_j = 0 \tag{9}$$

The output predictions corresponding to the solution defined in the previous step should be checked to verify if any other output will break its bounds. If this happens additional constraints of the type represented in (9) should be included in the control problem. This procedure should be repeated until no other constraints are required to be included in the control problem.

It can be shown that the control cost function of the MPC with economic objective represented in (8) is a quadratic function of the control move. As the constraints represented in (3) to (5) and (7) are linear in the control, the optimization problem that produces the proposed MPC with economic objective is a QP, with a computer effort similar to the computer effort of the conventional MPC.

4. The FCC (Fluid Catalytic Cracking) unit

To compare the performance of the SRTO/MPC algorithm proposed in section 3 with the RTO/MPC of Zanin et al. (2002) summarized in section 2, it is considered here the same process system studied by Zanin et al. (2002). The system is the Fluid Catalytic Cracking Kellogg Model F, which is described in Moro and Odloak (1995).

In the simulations presented here, only a few variables of the industrial control system are considered. The main manipulated inputs are: u_1 (*ton/h*), the total air flow rate to the two-stage regenerator, u_2 (*%*), the valve opening of the regenerated catalyst and u_3 (m^3/h), the gasoil feed flow. Considering the approach of zone control the controlled outputs considered here are: y_1 (*C*), the temperature in the dilute phase of the regenerator and y_4 (*C*), the riser temperature. Here, it is assumed that the economic objective of the FCC unit is to maximize the production of liquefied petroleum gas (LPG), which depends on the feed properties and process operating conditions. In the simulations the tuning parameters adopted in the RTO/MPC of section 2 are the following: m=2, p=70, W_1=diag(0.2 0.1 0.1 1), W_2=diag(5 5 5 5) and W_3=-1. The sampling period is equal to 1 min.

The proposed algorithm SRTO/MPC has the same tuning parameters and constraints as the RTO/MPC except the weight of the economic component, which in this case is W_3= 0.3.

Fig. 1 shows the behavior of the LPG production in m^3/h, which is the economic objective for the two approaches when the controller is started from the steady state. It is also shown the controlled outputs and the manipulated inputs. One may observe that the simplified algorithm (SRTO/MPC) captures most of the economic benefit obtained in the one-layer (RTO/MPC) algorithm proposed by Zanin et al. (2002). The increase in the economic term

is faster in the RTO/MPC, but the final value of the economic benefit is not significantly different in the two approaches. Obviously, in both methods the speed of response of the economic benefit will depend on the tuning weight W_3. The larger the absolute value of this parameter, the faster the economic benefit will respond to a change in the operating conditions or disturbances. However, as the constraints on the system outputs are not explicitly included in the control/optimization problem, the result of a large economic weight W_3 may be undesirable. If the optimal unconstrained operating point of the controlled system lies outside the convex set defined by the constraints, the economic term in the control cost function tends to push the system outputs to outside the control zones. In the example simulated here, Fig. 1 shows that y_1 tends to stabilize slightly below the minimum bound, while y_4 tends to stabilize slightly above the maximum bound. Also, input u_3 is pushed by the controller to its maximum bound. This indicates that all the degrees of freedom of the system have been consumed by satisfying the system constraints and the optimal operating point lies at the intersection of the constraints. A straightforward way to reduce the distance from the point, in which the outputs stabilize to the control zones, is by increasing the corresponding elements of weight W_1. However, as the dynamic model used in the controller is a linear model, there is a limit on the maximum value of this weight before the closed loop system becomes unstable. For instance, keeping the remaining tuning parameters of SRTO/MPC as defined above, the system becomes unstable if the component of W_1 corresponding to y_1 is made equal to 1. Thus, the integration of real time optimization with linear MPC in a one layer strategy seems to pose a robustness problem related to the uncertainty of the linear model.

5. Conclusion

Here, it is proposed a simplified version of the one layer RTO/MPC controller presented in previous works that integrates the two functions, real time optimization and multivariable constrained control, in a single layer. The control objective is to zero the reduced gradient of the economic objective while maintaining the system outputs inside their zones. The main advantage of the proposed approach is that the integrated control/optimization problem becomes a QP, instead of a NLP as in the previous approach. Simulation results have shown that a substantial percentage of the possible economic benefit associated with real time optimization can be captured with the proposed approach.

References

Aske, E. M. B., Strand, S., & Skogestad S. (2008). Coordinator MPC for maximizing plant throughput. *Computers and Chemical Engineering*, 32 (1-2), pp.195-204.

Engell, S. (2007), Feedback control for optimal process operation, *Jour. Proc. Control*, 17, 203-219.

Georgiou, A., Taylor, P., Galloway, R., Casey, L., Sapre, A. (1997). Plantwide closed loop real time optimization and advanced control of ethylene plant (CLRTO) improves plant profitability and operability. In *Proceedings of NPRA Comp. Conf., Nat. Petr. Ass.*, CC 97-139, New Orleans..

Kadam, J., Marquardt, W., Schlegel, M., Backx, T., Bosgra, O., Brouwer, P. J. (2003), Towards integrated dynamic real-time optimization and control of industrial processes. In *Proceedings Foundations of Computer-Aided Process Operations (FOCAPO2003)*, pp. 593-596.

Lu, J. (2003). Challenging control problems and emerging technologies in enterprise optimization. *Control Engineering Practice, 11,* 847-858.

Marlin, T.E., Hrymak, A. N. (1996). Real-time operations optimization of continuous processes. *In Proceedings of CPC V, AIChE Symposium Series*, vol. 93, 1997, pp.156-164.

Mercangoz, M., Doyle III, F. J. (2008). Real-time optimization of the pulp mil benchmark problem. *Computers and Chemical Engineering*, 32 (4-5), pp. 789-804.

Moro, L.F.L., Odloak, D. (1995). Constrained multivariable control of fluid catalytic cracking converters. *Journal of Process Control, 5*, 1-11.

Rolandi, P. A., Romagnoli, J. A. (2005). An integrated environment for support of process operations, *AIChE Annual Meeting, Conference Proceedings*, pp. 6026-6027.

Rotava, O., Zanin, A. C. (2005). Multivariable control and real-time optimization-An industrial practical view. *Hydrocarbon Processing, 84*(6), 61-71.

Tosukhowong, T., Lee, J., Lee, J., Lu, J. (2004). An introduction to a dynamic plant-wide optimization strategy for an integrated plant. *Computers and Chemical Engineering*, 29, 199-208.

Zanin, A. C., Tvrzská de Gouvêa, M., Odloak, D. (2002). Integrating real-time optimization into the model predictive controller of the FCC system. *Contr. Engng. Pract., 10*, 919-831.

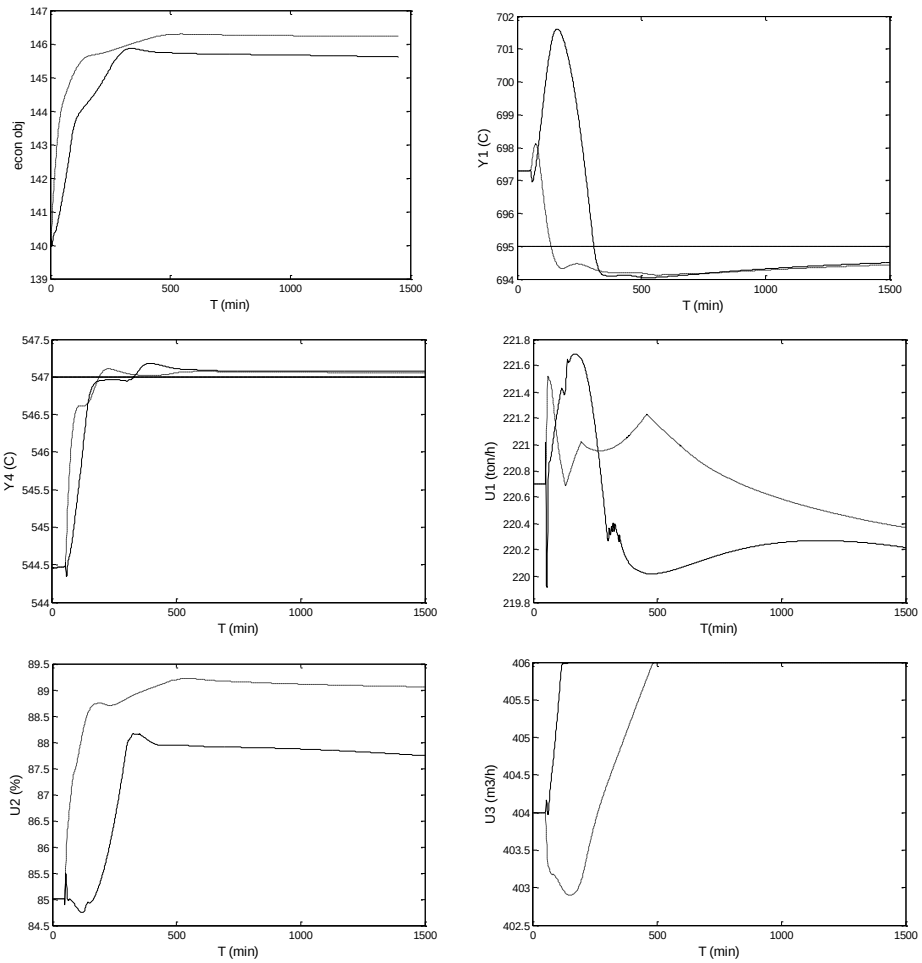


Fig. 1 FCC variables of SRTO/MPC (———) and RTO/MPC(− − −)

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Towards a Viscosity and Density Correlation for Dairy Fluids - A Soft Sensor Approach

Tien-I Lin[a], Glauce De Souza[a] and Brent Young[a*]

[a]*Chemical and Materials Engineering Department, The University of Auckland*
20 Symonds Street, Private Bag 92019, Auckland 1142, New Zealand
**Author to whom correspondence should be addressed: b.young@auckland.ac.nz*

Abstract

In spray drying, the viscosity of the feed is a critical characteristic. It affects the droplet size formed at the atomiser, the end product properties (e.g. bulk density, moisture content and solubility) and the characteristics of the process (e.g. drying time and wall sticking behaviour.). Currently viscosity has been robustly measured online on a pilot scale plant but no robust online measurements were reported for industrial scale plant. However, it is possible that empirical correlations can be developed.

For a spray drying plant, the density is usually measured online and used to calculate the percentage of total solids. It has also been identified that there is a positive relationship between the viscosity of the feed concentrate and the density. However this correlation is not reported in the open literature. Therefore, it is of interest to attempt to identify and validate this correlation for model predictive control purposes. Due to the availability and reliability of on line density measurement at the industrial plant, the validation and application such a correlation is potentially relatively simple.

A pilot plant was set up with a Micro-Motion density meter and a Hydromotion vibrational online viscometer. The viscometer was placed in a water- jacketed tank, where the measurements took place. Milk concentrate for total solids (TS) contents between 35% to 45% was then re-circulated in the system at temperatures up to 70^{o}C at atmospheric pressure to obtain online measurements of both density and viscosity.

Keywords: *Viscosity, Milk powder, Process control, Soft sensor*.

1. Introduction

Milk powder is the most important processed dairy product for export due to its long shelf life and ease of transportation. However, there is a huge energy requirement in the milk power production process, predominately from evaporation and drying

Spray drying is the most common milk powder manufacturing method. In recent years, spray drying technology has been developing in two separate directions. The improvement of energy efficiency and the development of new functional products by applying spray drying techniques. However the future success of both directions lies with better understanding and controllability of the process [1].

In the 2004 International Symposium on Spraying of Milk Products, advanced process control (APC) was considered to be one of the top areas that required further development due to difficulty in operating larger spray drying units. This area is currently advancing in a significant rate [1]. However, the primary challenge in developing a model for model predictive control remains the of understanding of

properties and stickiness behaviour of the droplet and powder in the spray drying process [1].

Currently, many spray drying process plants adopt process control schemes that are very dependent on the operator's experience, where the set point of a process is determined by the operators [2, 3]. Some attempts have been made to incorporate advanced process control but fail to take physical attributes of the milk into the process model [4], [5].

For spray drying, temperature is normally measured in lieu of moisture content online [4, 6, 7]. This method assumes a correlation between the temperature measurement and the evaporation rate [6]. However there is a fundamental flaw in this method. The variation in inlet air, wet bulb temperature or equilibrium constant of inlet air can result in dramatic change in moisture content of the product but have no effect on the outlet air temperature [6]. This relationship changes depending on the physical attributes of the milk. Chen pointed out that viscosity of the milk concentrate is important in this process [3].

Before any process control is applied to a system, a process model must be developed. For spray drying the main challenge is developing a process model for the spray drying unit. The relationship between the process condition and the product characteristics must be known. Also thermodynamic/fluid dynamic equations must be described in terms of online heat and mass transfer measurements [8]. These measurements must be valid for a range of process conditions and product specifications. Given these requirements, viscosity must be measured in terms of total solids and shear rate for a range of the product specifications to give accurate fluid dynamic equations.

In many cases, development of process control is limited by the availability of robust and accurate sensors. This can be due to the cost effectiveness and the limitation of technology. In the case of viscosity in the diary industry, attempts of using online viscometers and soft sensor approaches have both been undertaken [8, 9] but only on pilot plant. Density of milk concentrate is normally monitored at the end of the evaporator and used to determine the total solid content of the milk concentrate. It has been reported by several sources that a correlation between density and viscosity of milk concentrate exits, however without details.

This paper intends to describe an experimental set up that is capable of finding such a correlation. The performance of the experimental rig was evaluated and preliminary results are also presented in this paper.

2. Experimental method

The set up of the experiment and the experimental method of testing for the viscosity-density correlation are described in this section.

2.1. Rig description

The basic experimental rig setup (Figure 1, [10]) consists of two tanks, henceforth referred to as the jacketed tank and the hot water tank, respectively. The jacketed tank is fitted with an external jacket in which hot water is re-circulated. The jacketed tank also contains the milk concentrate in an inner tank. The hot water tank is a temperature controlled heat source, providing heated water which is circulated through the jacket of jacketed tank.

A Hydro-motion vibrational viscometer (XL/7) is inserted into the jacketed tank. This allows the viscosity of the milk concentrate to be measured at this point. Afterwards, the milk concentrate from the jacketed tank is passed though a micro-motion coriolis meter (DL65) allowing the density and flow rate of the system to be measured. The milk concentrate is then returned to the jacketed tank.

In order to control the milk concentrate temperature, a control valve is fitted on the water line from the hot water tank to the jacketed tank. The temperature of the jacketed tank is controlled by a control valve which adjusts the amount of heated water allowed to enter the jacket of the jacketed tank. Inside the hot water tank, a 3kW heating element is fitted along with a PID temperature controller and temperature sensor to regulate the water temperature.
Temperature is monitored at several places in the rig, including the hot water inlet to the Jacketed tank, the Jacketed tank hot water outlet, inside Hot water tank, at the micro-motion coriolis meter and inside the Jacketed tank. Other variables such as flow and density are monitored at the micro-motion coriolis meter within the jacketed tank.

2.2. Operating Procedure

The heating element is turned on to heat the water to a desired temperature. A controller is fitted to the heating element which cuts power to the element when the water reaches the desired temperature. For a typical run, the controller is set to cut out at 80 °C to maintain heated water to the rig.
Whole milk powder was used to prepare samples of milk concentrate with total solid contents of 35%, 40% and 45%. Once prepared, the sample is poured into the hot water jacketed tank for viscosity testing and recirculation. Once the jacketed tank is loaded, the pump was switched on to start to re-circulate of milk concentrate within the milk loop; the pump from hot water loop is also started to heat up the sample to the desired temperature.
From offline experimental results to minimise protein denaturation and structure change, the operating temperature for pilot plant is limited to 70°C. Structure change was observed above 80°C and it has been reported in literature that excessive whey denaturisation could not be avoided if it is operating above approximately 71°C [11].
Measurements were made during both heating and cooling of the concentrate to investigate whether the thermal history is an important factor in viscosity measurement. Data was collected during heating and cooling in the following temperature runs: 40 °C, 50 °C, 60°C and 70 °C. Other than at the maximum temperature of 70 °C, results were all collected twice. These two sets of results were then compared to check for consistency of viscosity measurements.
At the end of the experimental run, hot water is used to clean the rig by re-circulating at 70°C for 10 minutes. This is followed by a Clean In Place (CIP) procedure where caustic and acid clearing was performed.

3. Results

3.1. Temperature sensitivity

Figure 2 presents the viscosity for both the heating and cooling cycles for whole milk concentrates of 35%,40% and 45% TS contents. The viscosity value was taken to be the average viscosity reading over a steady state of period of 10 minutes. Fluids were tested for both heating and cooling cycles to see if the fluid was still the same.
The results in **Figure 2** show that there is a strong downward trend, in which the viscosity of the fluid decreases as the temperature increases. The milk concentrate behaviour appears to observe the same trend when heating and cooling. This suggests that the milk concentrate is thermally stable over the experimental range.

3.2. Viscosity density relationship

The viscosity –density curves and correlations at temperatures between 40°C to 70°C are presented in Figure 3 and Table 1, in which we can see that there is a clear upward trend as expected.

4. Discussion

The method presented in this paper offers a potential way of formulating a viscosity-density correlation. The fluid appears to be thermally stable within the experimental conditions. There appears to be no difference in viscosity after the fluid is heated and cooled down.

Viscosity and density correlations were able to be identified, an increase-increase relationship was found at all temperatures tested. This relationship is in agreement with expectation and literature. However, the slopes of the correlations are very shallow (0.0041-0.0090 $g.cm^{-3}/cP$) and there is some scatter in the data (R^2 is 0.77-0.93). This indicates that further study is needed, either with respect to improving these correlations or researching direct on-line viscosity measurement.

The method presented in this paper offers a potential way of testing for viscosity-density correlation which is useful for the formulation of a process model for advanced process control. In this experiment, reconstituted milk was used instead of fresh milk concentrate. This is expected to have an influence on the viscosity value of the milk concentrate, as the milk powder is exposed to high temperature at the spray drying chamber and this can cause protein denaturisation and hence viscosity change.

5. Conclusions and Future work

The correlations developed by this method are empirical, and hence each different product specification (milk composition) needs to be tested if it is to be used for the model. It is expected that milk composition will affect the viscosity-density correlation. Further development is required in the search of density-viscosity correlation for milk. The approach presented above offers a way of identifying this correlation in steady state. However, for process control purposes, a dynamic correlation must be identified. This implies that the density and temperature needs to be changed and the response of viscosity needs to be observed.

Currently the sample used is reconstituted milk. It is expected that there is a difference in viscosity of milk concentrate between fresh milk concentrate and reconstituted milk concentrate. As this correlation is purely empirical correlation and highly possible site dependant. It is important perform the experiment with the feed from the particular plant.

In conclusion, this paper presents a proposed method for formulating a viscosity-density relation via an empirical approach. This correlation is potentially useful for the development of predictive control for milk concentrate when formulated, but requires further study and also testing on a specific site.

References

1. Kelly, P.M., Innovation in milk powder technology. 2006. 59(2): p. 70-75.
2. Setnes, M. and R. Babuska, Fuzzy decision support for the control of detergent production. International Journal of Adaptive Control and Signal Processing, 2001. 15(8): p. 769-785.
3. Chen, X.D., Towards a comprehensive model based control of milk drying processes. Drying Technology, 1994. 12(5): p. 1105-1130.
4. Govaerts, R., Johnson, A., Crezee, R., Reyman, G. and Swinkels, P.L.S., Control of an industrial spray drying unit. Control Engineering Practice, 1994. 2(1): p. 69-85.

5. Allen, R.M. and H.H.C. Bakker, Spray dryer control based on online particle size analysis. Chemical Engineering Research and Design, 1994. 72(A2): p. 251-254.
6. Zaror, C.A. and J.R. Perez-Correa, Model based control of centrifugal atomizer spray drying. Food Control, 1991. 2(3): p. 170-175.
7. Clement, K.H., et al., On the dynamic behaviour of spray dryers. Chemical Engineering Research & Design, 1991. 69(3): p. 245-252.
8. O'Callaghan, D. and P. Cunningham, Modern process control techniques in the production of dried milk products - A review. Lait, 2005. 85(4-5): p. 335-342.
9. Schuck, P., Mejean, S., Dolivet, A., Beacher, E. and Famelart, M.-H., Pump amperage: A new method for monitoring viscosity of dairy concentrates before spray drying. Lait, 2005. 85(4-5): p. 361-367.
10. De Souza, G.F., Lin T.-I.; Young, B.R., Dairy Fluid Viscosity Measurement and Process Control, Proceedngs of the International Conference of Agricultural Engineering, Iguazu Falls, Brazil, 2008.11.
11. O'Sullivan, A.C., Whey protein denaturation in heat processing of milk and dairy products. Journal of the Society of Dairy Technology, 1971. 24(1): p. 45-53.

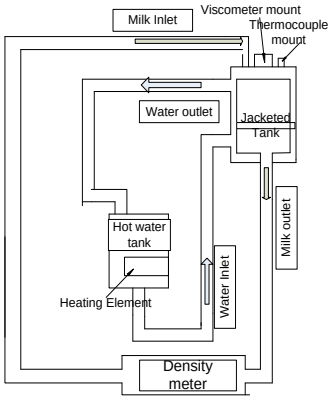


Figure 1. Schematic of the jacketed tank and hot water tank, adapted from [10]

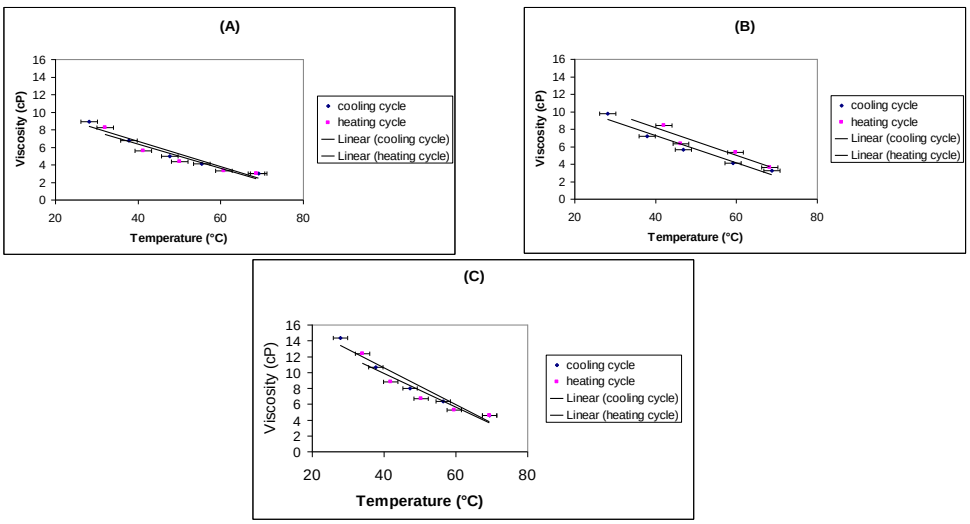


Figure 2. Viscosity versus Temperature for Skim Milk Concentrate for Heating/Cooling cycles, (A) At 35% TS, (B) At 40% TS, and (C) At TS 45% TS

Table 1. Density-viscosity correlation for whole milk between 35% to 45% total solids, where y= density (gcm^{-3}) and x= viscosity (cP)

Temperature (°C)	Correlation	R^2
40	y = 0.0041x + 1.0437	0.77
50	y = 0.0067x + 1.0324	0.93
60	y = 0.0064x + 1.0362	0.81
70	y = 0.0090x + 1.0259	0.87

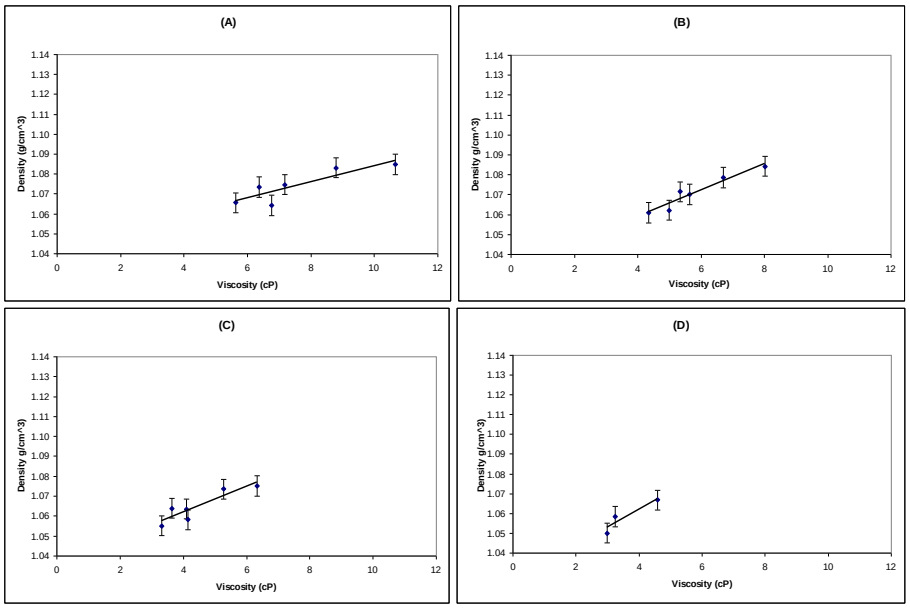


Figure 3. Viscosity-Density correlation for whole and skim milk over temperatures of 40 °C to 70°C: (A) At 40 °C, (B) At 50°C, (C) At 60 °C, and (D) At 70°C

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Learning to repair plans and schedules using a relational (deictic) representation

Jorge Palombarini[a], Ernesto Martínez[b]

[a]*UTN-Fac. Reg. V. María, Av. Universidad 450, Villa María 5900, Argentina*
[b]*INGAR (CONICET-UTN), Avellaneda 3657, Santa Fe S3002 GJC, Argentina*

Abstract

Unplanned and abnormal events may have a significant impact in the feasibility of plans and schedules which requires to repair them 'on-the-fly' to guarantee due date compliance of orders-in-progress and negotiating delivery conditions for new orders. In this work, a repair-based rescheduling approach based on the integration of intensive simulations with logical and relational reinforcement learning is proposed. Based on a deictic representation of schedule states a number of repair operators have been designed to guide the search for a goal state. The knowledge generated via simulation is encoded in a relational regressión tree for the *Q*-value function defining the utility of applying a given repair operator at a given schedule state. A prototype implementation is discussed using a representative example of 3 batch extruders processing orders for 4 different products. The learning curve for the problem of inserting a new order vividly illustrates the advantages of logical and relational learning in rescheduling.

Keywords: rescheduling, artificial intelligence, batch plants, relational modeling.

1. Introduction

Fast rescheduling in real-time is mandatory to account for unplanned and abnormal events including arrival of new orders, equipment breakdown, reprocessing and raw material delays (Vieira et al, 2003). Reactivity and responsiveness is a key issue in any rescheduling strategy which makes critical the capability of generating and representing knowledge about heuristics for repair-based scheduling using case-based reasoning (Miyashita, 2000). One such example is the **CABINS** framework for case-based rescheduling proposed by Miyashita and Sycara (1994) that heaviliy resorts to human experts. Along similar ideas another important work in the field of the so-called *intelligent scheduling* techniques are contributions by Zweben et al. (1994). In this work the novelty lies in integrating intensive simulations with automatic learning to overcome the issue of human experts for domain-specific scheduling problems. A novel methodology which combines a deictic representation of schedules states with relational reinforcement learning is proposed to develop a near-optimal policy for interactive scheduling repair bearing in mind different goals and scenarios. To this aim, learning to repair schedules based on simulation is undertaken using two general-purpose algorithms already available: TILDE and RRL (Džeroski et al, 2001; De Raedt, 2008).

2. Repair-based (re)scheduling

Fig. 1 depicts the repair-based optimization architecture where search control knowledge about repair operator selection is acquired through reinforcements using a schedule state simulator. In the simulation environment an instance of the schedule is

interactively modified by the learning agent using a set of repair operators until a goal is achieved or the impossibility of repairing the schedule is accepted. In each interaction, the learning agent receives information from the schedule situation or state *s* and selects a repair operator to be applied to the current schedule as an action *a*. The resulting quality of a schedule after the repair operator has been applied is evaluated using the simulation environment via an objective or reward function *r*(*s*). The learning agent then updates its action-value function *Q*(*s*,*a*) that estimates the value or utility of resorting to the chosen repair operator *a* in a given schedule state *s*. Such an update is made using a reinforcement learning algorithm (Sutton and Barto, 1998) such as the well-known *Q*-learning rule. By accumulating enough experiences over many simulated interactions the agent is able to learn an optimal policy for choosing the best repair operator at each schedule state. The main issue for learning is then how schedules states and actions must be represented for knowledge acquisition and iterative revision.

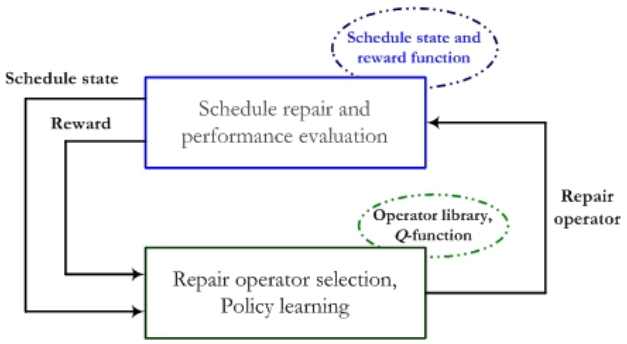


Fig. 1. Knowledge acquisition for schedule repair using reinforcement learning

For repairing a plan or schedule, the agent is given a *goal-state* $S :\rightarrow \{true, false\}$ defining which states in the planning world are target states. The objective of a repair task is phrased as: Given a starting state schedule s_1, find a sequence of repair operators $a_1, a_2, ..., a_n$ with $a_i \in A$ such that the $S \equiv true$. Usually a precondition function is used to specify which subset of repair operators can be applied at each state of the planning world to account for resource and precedence constraints. Also, a reward function is used by a learning algorithm to develop a repair policy using simulations (Martínez, 1999). The key issue is though how a schedule state must be best represented so that the repair policy can be applied to unseen schedule states more effectively. Propositional representations are not adequate for learning in open planning worlds defined by tasks, their properties and their relations to other tasks and resources. A *deictic* representation dealing with the varying number of tasks in the planning world by defining a focal point for referencing tasks is proposed here as a much powerful alternative. In a deictic representation, the rest of the planning world is then defined in relation to that focal point as it is shown in Fig. 2. These local repair operators move the position of a task alone, however due to the ripple effects caused by tight resource-sharing constraints other tasks may need to be moved which is not desirable. Whenever the goal-state for the schedule cannot be achieved using primitive repair operators more elaborated macro-operators can be used to implement a combination of basic repair operators such as `task-swapping`, `batch-split` or `batch-merge` until the goal state (e.g. order insertion without delaying other orders) is achieved.

3. Relational reinforcement learning (RRL)

RRL algorithms are concerned with reinforcement learning in domains that exhibit structural properties and in which different kinds of related objects such as tasks and resources exist (Džeroski et al, 2001; De Raedt, 2008). These kind of domains are usually characterized by a very large and possibly unbounded number of different possible states and actions. In this kind of environments, most traditional reinforcement learning techniques break down. One reason why traditional RL fails is that it stores the learned *Q*-values explicitly in an state-action table, with one value for each possible combination of states and actions. Rather than using an explicit state−action *Q*-table, RRL stores the *Q*-values in a logical regression tree (Blockeel and De Raedt, 1998).

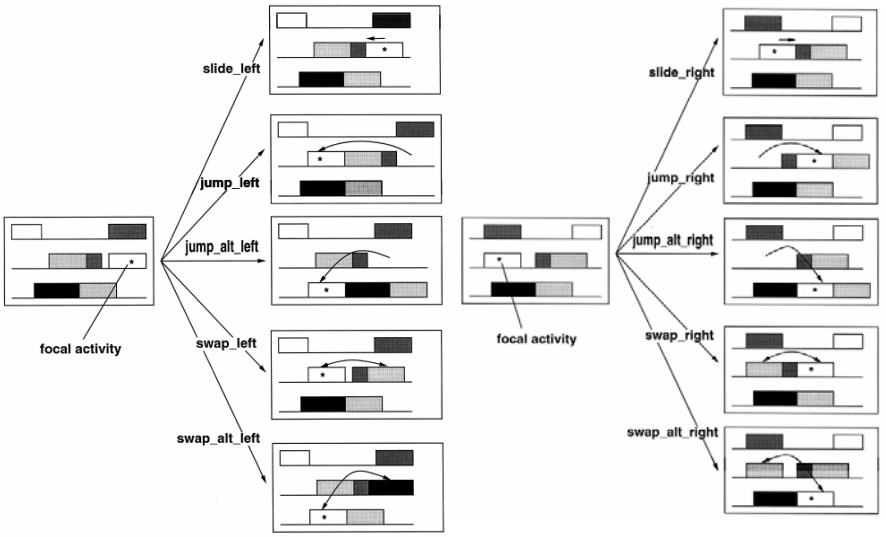


Fig. 2. Deictic representations of repair operators

In RRL states are represented as sets of first-order logical facts, and the learning algorithm can only see one state at a time. Actions are also represented relationally as predicates describing the action as a relationship between one or more variables. Because of the relational representation of states and actions and the inductive logic programming component of the RRL algorithm, there must exist some body of background knowledge which is generally true for the entire domain. After the *Q*-function hypothesis has been initialized, the RRL algorithm starts running learning episodes like any standard *Q*-learning algorithm (Sutton and Barto, 1998; Džeroski et al, 2001). During each learning episode, all the encountered states and the selected actions are stored, together with the rewards related to each encountered (state, action) pair. At the end of each episode, when the system encounters a *goal* state, it uses reward back-propagation and the current *Q*-function approximation to compute the appropriate *Q*-value approximation for each encountered (state, action) pair in the episode. The algorithm then presents the set of (*state*, *action*, *qvalue*) triplets to a relational regression engine, which will use this set of *Examples* set to update the current regression tree of the *Q*-function, and then the algorithm continues executing the next learning episode.

Initialize the *Q*-function hypothesis $\hat{Q}_0$
$e \leftarrow 0$
repeat
  Examples $\leftarrow \varnothing$
  Generate a starting schedule state s_0
  $i \leftarrow 0$
  repeat
    choose a repair operator a_i at s_i using a policy (e.g., ε-greedy) based on the current hypothesis $\hat{Q}_e$ implement operator a_i, observe r_i and the resulting schedule s_{i+1}
    $i \leftarrow i+1$
  until schedule state s_i is a *goal* state
  for $j = i$ -1 to 0 **do**
    generate example $x = (s_j, a_j, \hat{q}_j)$, where $\hat{q}_j \leftarrow r_j + \gamma \max_a \hat{Q}_e(s_{j+1}, a)$
    Examples $\leftarrow$ *Examples* $\cup \{x\}$
  end for
  Update $\hat{Q}_e$ to $\hat{Q}_{e+1}$ using *Examples* and a relational regression algorithm (e.g. TG in Fig. 4)
until no more learning episodes

Fig, 3. A RRL algorithm for learning to repair schedules through intensive simulations

The TG algorithm described in Fig. 4 is a relational regression algorithm that have been developed for policy representation in logical and relational learning (De Raedt, 2008). This incremental first order regression tree algorithm is used here for accumulating simulated experience in a compact representation of a repair-based policy based on *Q*-values for the repair operators available at the state *s*. Fig. 5 gives a small example of a first order regression tree for the *Q*-value function in a task (re) scheduling world to react to events and disturbances.

//initialize by creating a tree with a single leaf with empty statistics
for each learning example that becomes available **do**
  sort the example down the tree using the tests of the internal nodes **until** it reaches a leaf
  update the *Q-value* in the leaf according to the new example
  if the statistics in the leaf indicate that a new split is needed **then**
    generate an internal node using the indicated test
    grow 2 new leafs with empty *Q* statistics
  end if
end for

Fig. 4. TG algorithm for relational induction

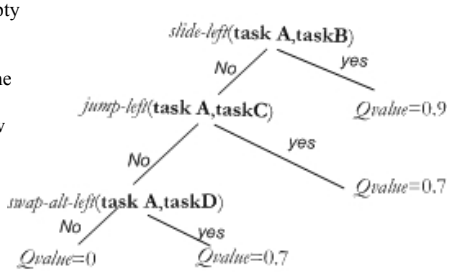


Fig. 5. A relational regression tree

4. Example

A small example problem proposed by Musier and Evans (1989) is considered to illustrate the use of repair operators for batch plant rescheduling. The plant is made up of 3 semicontinuous extruders which process customer orders for 4 products. Processing rates and cleanout requirements are detailed in the quoted reference.

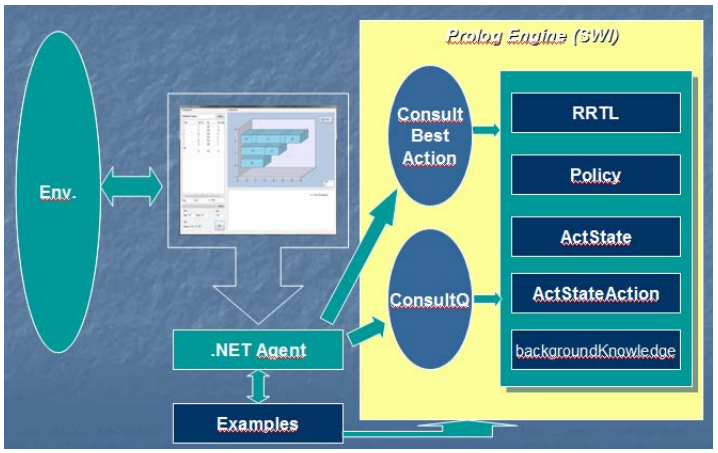


Fig. 6. SWI Prolog engine for implementing RRL in repair-based (re)scheduling

The prototype application has been implemented in Visual Basic.NET 2005 Development Framework 2.0 SP2 and SWI Prolog 5.6.61 running under Windows Vista. Also, the **TILDE** and **RRL** modules from The **ACE Datamining System** developed by the Machine Learning group at the University of Leuven have been used. Let's discuss only the specific situation where there exist 10 already scheduled orders in the plant and a new order #11 must be inserted so that the Total Tardiness (TT) in the schedule is minimized. Data for scheduled orders and the new order are given in Table 1. En each episode, a random schedule state for orders #1 through #10 is generated and a random insertion attempted for the new order which in turn serves as the focal point for defining repair operators. Also, a goal state for the resulting schedule is stated in terms of the TT, e.g. of 5% increase. Background knowledge such as "the number of orders scheduled for extruder #3 is larger than the number for extruder #2" is provided to speed up learning. In Fig. 7, the learning curve for the new order insertion rescheduling event is shown. Training episodes comprise of a number of steps, each one corresponding to the application of one local repair operator. After 15 training episodes only 3 repair steps are required on average to insert the 11th order. Fig. 8 provides an example of the optimal sequence of repair operators from the schedule in Fig. 8(a).

5. Final comments

A novel approach for simulation-based development of a relational policy for automatic repair of plans and schedules using reinforcement learning has been proposed. The policy allows generating a sequence of deictic (local) repair operators to achieve a goal such as order insertion with minimum tardiness based on deictic representations.

6. Acknowledgements

The authors thanks to the ANPCyT of Argentina which partially supports the work presented here through project PICT 1099/06.

References

H. Blockeel, L. De Raedt, 1998, Top-down induction of first order logical decision trees. *Artificial Intelligence*, 101, 1-2, 285–297.

K. Driessens, 2004, *Relational reinforcement learning*, Ph.D. dissertation, Department of Computer Science, K.U.Leuven, Leuven, Belgium..
S. Džeroski, L. De Raedt, K. Driessens, 2001, Relational reinforcement learning, *Machine Learning*, 43, 1/2, 7−52.
L. De Raedt, 2008, *Logical and relational learning*, Springer-Verlag, Berlin, Germany.
E. C. Martinez, 1999, Solving batch process scheduling/planning tasks using reinforcement learning, *Computers and Chemical Engineering*, 23, S527-531.
K. Miyashita, 2000, Learning scheduling control through reinforcements, *International. Transactionsin Operational Research* (Pergamon Press), 7, 125-138.
K Miyashita, K. Sycara, 1994, **CABINS**: a framework of knowledge acquisition and iterative revision for schedule improvement and iterative repair, *Artificial Intelligence*, 76, 377-426.
R. Musier, L. Evans, 1989, An approximate method for the production scheduling of industrial batch processes with parallel units, *Computers chem Engng*, 13, 1-2, 229-238.
R. Sutton, A. Barto, 1998, *Reinforcement Learning: An Introduction*, MIT Press, Boston, MA.
G. Vieira, J. Herrmann, E. Lin, 2003, Rescheduling manufacturing systems: a framework of strategies, policies and methods, *J. of Scheduling*, 6, 39-62,
M. Zweben, E. Davis, B. Doun, M. Deale, 1993, Iterative repair of scheduling and rescheduling, *IEEE. Trans. Syst. Man Cybern.*, 23, 1588-1596.

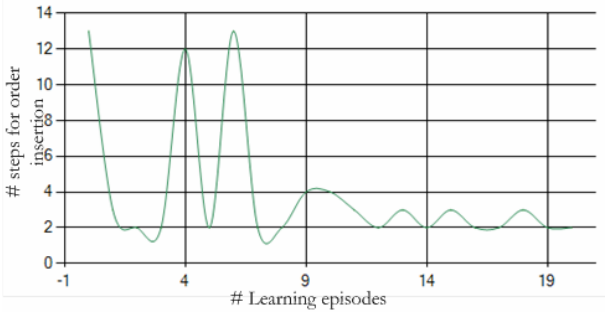
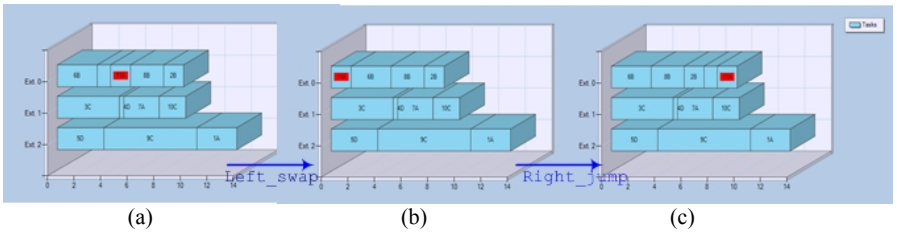


Fig. 7. Learning curve for the repair-based order insertion example

Table I. Data for the new order insertion example

#	Product	Size (Kg)	DD (h)	#	Product	Size (Kg)	DD (h)
1	A	300	6	**6**	B	600	5
2	B	300	5	**7**	A	400	6
3	C	700	3	**8**	B	500	12
4	D	100	2	**9**	C	700	17
5	D	700	10	**10**	C	300	8
11	A	150	10				



(a) (b) (c)

Fig. 8. Before inserting order #11 (in red) the is TT=17.5. The goal is to repair the schedule in (a) having a TT=20 h so that TT increase is minimized; (b) After a `Left_swap` (b) TT=19 h.; After a `Rigth_jump` in (c) TT=17.50, meaning an insertion is found without any increase in TT.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Development of multi-modal control programs for continuous-discrete process supervision

Mariano De Paula, Ernesto Martínez

[a]*INGAR (CONICET-UTN), Avellaneda 3657, Santa Fe S3002 GJC, Argentina*

Abstract

Continuos-discrete processes are characterized by the strong coupling of continuos state dynamics of recipe-driven discontinuos operations and the discrete state dynamics of logic-based controllers acting at the interface with continuos processes. Multi-modal control is a promising design tool for supervisory control of these type of systems by resorting to sequences of control modes each one comprising of a purposeful feedback law and a stopping condition. In this paper the problem of developing multi-modal control programs from a given mode alphabet using data from alternative disturbance scenarios is addressed. A case study related to maximizing the average productivity of a hybrid chemical plant through multi-modal control of a buffer tank is presented.

Keywords: Batch control, hybrid plant, supervisory control, mode-switching control.

1. Introduction

There has been a lot of interest in modelling and control of hybrid dynamical systems in recent years, motivated by a growing number of hybrid chemical plants involving tight interaction between continuous variables and discrete events/decisions (Engell, *et al*., 2000). A multi-modal system is a system whose dynamics switches among finitely many possibilities to achieve a desired goal or behavior. These switches can be in response to an occurrence of a specific event or a controlled decision (Liberzon, 2003; Mehta and Egerstedt, 2006). Optimal control of multi-modal dynamic systems is of paramount importance for designing a supervisory control structure (Koutsoukos, *et al*., 2000). In this context, a mode σ is specified as a triple $(\kappa(\bullet),\xi(\bullet),T)$, where $\kappa(\bullet)$ is a control law (closed-loop control policy) seeking to achieve a given sub-goal or to deploy a certain behavior, $\xi(\bullet)$ is the interrupt function mapping process observations to the set $\{0,1\}$ and T is the time interval over which the mode σ is active. The key issue in the design of a supervisory control system is to find the optimal concatenation of modes in a finite-length mode string π based on the rewards observed after each mode transition occurs.

2. Multi-modal control programs

Suppose the state x of a dynamic system follows:

$$\frac{dx}{dt} = f(x, u(t), z(t)),\ x \in X = R^n,\ u \in U = R^m \tag{1}$$

where $z(t)$ is a time-varying measurable disturbance which evolves arbitrarily as

$$\frac{dz}{dt} = g(z,t),\ z \in Z = R^d \tag{2}$$

If at a given time τ_0, where the state is $x(\tau_0)$ and the disturbance is $z(\tau_0)$, the system receives the input string $\pi = \{(\kappa_1, \xi_1),\ldots,(\kappa_q, \xi_q)\}$ which gives rise to the state transition

$$\delta(\sigma, x(\tau_0), z(\tau_0)) = x(\tau_0) + \int_{\tau_0}^{\tau_1} f(x(t), u(t), z(t)dt + \ldots + \int_{\tau_{q-1}}^{\tau_q} f(x(t), u(t), z(t)dt \tag{3}$$

If the length of the mode sequence π is bounded, only a finite set of states are reachable. Hence, by resorting to control programs with maximum length N a quantization of the system state results. Based on this multi-modal quantization a Lebesgue-sampled finite state machine $(X_N^Q, \Sigma, \tilde{\delta}, \tilde{x}_0, \tilde{z}_0)$, where Q stands for the value function (Sutton and Barto, 1998) and the state transition $\tilde{\delta}$ is defined as follows

$$\tilde{x}_{k+1} = \tilde{\delta}(\tilde{x}_k, \tilde{z}_k, \sigma_k) = \delta(\sigma, x(\tau_0), z(\tau_0)), k = 0,1,2,\ldots., x_0 = x(\tau_0), z_0 = z(\tau_0), \tag{4}$$

The discrete state space X_N^Q is given by the set of all states that are reachable from $\chi_0 = (x_0, z_0)$ using a sequence of modes $\sigma \in \Sigma$ of length less than or equal to N. The issue of concern here is finding a sequence of control-interrupt pairs that maximizes a cumulative reward for such state transitions. Based on the multi-modal quantization X_N^Q of the state space and the transition dynamics of the Lebesgue-sampled FSA, reinforcement learning (Sutton and Barto, 1998) can be used to learn the Q-values for a given reward function. It is worth noting that both state exploration and disturbance scenarios are required to discover which is the best mode σ^* for each pair $\chi \in X_N^Q$.

$\aleph := \{\tilde{\chi}_0, \delta(\tilde{\chi}_0, \sigma)\}$; *step* $(\tilde{\chi}_0) := 0$;
step $\tilde{\delta}(\tilde{\chi}_0, \sigma) := 1, \quad \forall \sigma \in \Sigma$
$k := 1$; *index for counting visits to state-disturbance pairs* $\chi \in \aleph$
$Q_k(\tilde{\chi}, \sigma) := const, \forall \chi \in \aleph, \sigma \in \Sigma$
repeat
  $k = k+1$
  $(\tilde{\chi}) := rand(\chi \in \aleph | step(\chi) < N)$
  $\sigma := rand(\Sigma)$
  $\tilde{\chi}' := \tilde{\delta}(\tilde{\chi}, \sigma)$
  if $\tilde{\chi}' \notin \aleph$ **then**
    $step(\tilde{\chi}') = step(\tilde{\chi}) + 1$
    $\aleph := \aleph \cup \tilde{\chi}'$
    $Q_k(\tilde{\chi}', \sigma) := const, \forall \chi \in \aleph, \sigma \in \Sigma$
  end if
  $Q_{k+1}(\tilde{\chi}, \sigma) := Q_k(\tilde{\chi}, \sigma) +$
  $\alpha_k \{r(\tilde{\chi}, \sigma) + \gamma \max_{\sigma' \in \Sigma} [Q_{k-1}(\tilde{\chi}', \sigma') - Q_{k-1}(\tilde{\chi}, \sigma)]\}$
until $\mathrm{mod}(k, L) = 0$ and $|Q_k(\tilde{\chi}, \sigma) - Q_{k-L}(\tilde{\chi}, \sigma)| < \varepsilon, \forall \chi \in \aleph, \sigma \in \Sigma$
$X_N^Q = \aleph$

Fig. 1. Q-learning algorithm for multi-modal control programs

In the *Q*-learning algorithm of Fig. 1, the quantized state-disturbance space X_N^Q is assumed to be unknown initially and we thus begin learning/exploring from selected states/disturbances pairs $\tilde{\chi}_0 = (\tilde{x}_0, \tilde{z}_0)$ and simulating transitions toward all the states which are reachable using the first mode of all control programs $\sigma \in \Sigma$. At each iteration of the learning process, a state-disturbance pair is randomly chosen from the set of known states-disturbance pairs and a control mode from the set Σ is applied to observe a state transition to $\tilde{x}'$; $r(\tilde{\chi}', \sigma)$ is the reward obtained for the transition observed once the execution of mode σ is finished.

In the algorithm, the function *step*($\tilde{x}$) represents the length of the shortest control program used so far to reach the state $\tilde{x}'$ from the initial state $\tilde{x}_0$. As a result, only states that are reachable from $\tilde{x}_0$ using mode strings of length less than or equal to *N* are explored, i.e. $\tilde{x}' \subset \tilde{\chi}' \in X_N^Q$ is guaranteed. As the transition to the next state $\tilde{x}'$ is observed it is necessary to determine if it is in the neighborhood of a previously visited state-disturbance pair $\tilde{\chi}$. If not the pair $\tilde{\chi}$ is added to the known state space-disturbance, the coresponding number of steps is increased by 1 and a new entry to *Q*-table is added. It is worth noting than when $\tilde{\chi}'$ is in the the neighborhood of a previously visited state the *Q*-values of all states-disturbances pairs in the neighborhood are updated. Exploration of the state-disturbance space and updates of the *Q*-table (value function) continues in this manner until the *Q*-table is stationary. Stopping conditions are imposed so as to guarantee that sufficiently many state-disturbance pairs are visited and all control modes $\sigma \in \Sigma$ are tried enough times for the *Q*-table to converge while the learning rate α is increasingly reduced.

3. Case study

3.1. Hybrid chemical plants

Buffer tanks play a key role in hybrid chemical plants such as the *Solvay* PVC production line (Melas, 2003) or sugar milling plants (Ghaeli et al., 2008) to smooth the interface between the batch and continuos sectors in the plant. To maximize the average plant productivity a buffer tank is used as shown in Fig. 2. This type of processes are generally made up of several parallel working units (e.g., batch reactors or crystallizers) ”sharing” common resources, typically utilities (e.g. cold water, power or steam). Batch reactors or crystallizers often operates by cyclically passing through a sequence of phases, namely loading of raw material, cooling, heating, reaction, etc.. Resource sharing by different process unit may increase the duration of these tasks in the recipe which will alter the whole schedule of batches in the plant. Also, as batch units discharge their final product in a buffer tank used to transfer it continuously to the continuous units in the plant (centrifuges, dryers, etc.) such delays may drastically the pattern of the inflow to the tank. There also exist some “waiting” times for a batch which can be used to compensate for deviations in the original schedule but at the cost of a productivity loss.

Maximization of average outflow from the tank is the main objective for productivity optimization in this hybrid chemical plant. However, another important component of the desired tank operation is that its outflow rate F_{out} must be changed smoothly despite significant variations in its inflow rate. To do so the tank capacity must be managed properly by allowing the tank level to vary within its minimum and maximum limits. Also, the outflow rate F_{out} is constrained by minimum and maximum values that must

be enforced due to throughput limitations downstream in the continuos processing sector. It is worth noting that in hybrid chemical plants the main role of the buffer tank is handling intermitent inflow rates which prevents resorting to averaging level control techniques (McDonald, et al., 1986). Also, for this type of processes the inflow rate variability pattern may be the result of significant deviations from periodic or optimal schedules due to resouce sharing, e.g. utilities (Simeonova, 2008).

How does multi-modal control of the buffer tank fits in the problem of productivity optimization of the hybrid plant? Let's take a quick look to what may be expected from the control system to elaborate on some intuitive insights in the search for a well-posed solution. Firstly, it is important that feedback laws allows that at any time the outflow rate must be a monotonic function of the tank level. This selection introduces a self-regulation property into the multi-modal control strategy. Also, modes must be designed so that despite peaks and valleys in inflow rate mode switching makes possible to maintain the average outflow rate as high as possible throughout. Thus, different modes should provide different degree of aggresiveness through their feedback laws when seeking a sub-goal (accumulate or drain-off inventory). Finally, rewards obtained after each mode application must be defined bearing in mind productivity maximization, e.g. the mean value of the outflow rate over the time inteval *T* over which the mode was applied. Large negative rewards are given to modes whose applications results in less than a minimum outflow rate or tank overflowing. To avoid undesirable disruptions in downstream units once a mode change occurrs the corresponding feedback laws must be applied for a minimum number of time intervals, say 5 min. This avoids excessive mode switching without compromising mode sequence optimality.

3.2. Tank inflow data and control modes

As a representative example of the application of the algorithm in Fig. 1 to a buffer tank, let's assume the buffer tank has a maximum height of 1 m which is equal to its diameter. The inflow rate variability is modeled here as the time-series generated using the well-known Mackey-Glass differential delay equation using random initial values

$$\frac{dF_{in}}{dt} = \frac{0.2F_{in}(t-30)}{1+\left(F_{in}(t-30)\right)^{10}} - 0.9F_{in}(t)\ [\mathrm{m^3/h}] \tag{5}$$

Time series generated using Eq. (5) are very sensitive to initial conditions while exhibiting a non-periodic/non-convergent pattern as can be seen from portions of the series shown in Fig. 3. This type of inflow variability may look far from real but it will serve as an extreme benchmark for developing multi-modal control programs using the algorithm of Fig. 1 to handle disturbance scenarios where significant variations in the upstream parallel unit discharge schedule are present. Modes are designed so as to achieve either the sub-goal *accumulate* (A1 and A2) or the sub-goal *drain-off* (D1 and D2) using for both the following the two simple feedback laws:

$$\kappa_1 : F_{out}(t) = \frac{0.4.[h(t)]^{0.5}}{1.2-h(t)} - 0.17 \quad \text{and} \quad \kappa_2 : F_{out}(t) = 0.934.h(t) - 0.15 \tag{6}$$

where *h* is the tank level (see Fig. 4). The built-in smoothness of the feedback laws in (6) is important to reduce the amplitude of outflow rate variations as much as possible. Once the optimal sequence of modes is found feedback laws and their parameters can be optimized using a variational formulation with the objective of lowering the complexity

of π and smoothing the outflow rate. Stopping conditions are defined based on the failure to achieve the corresponding sub-goal over time. For example, for any accumulation (drain-off) mode its stopping event triggers the corresponding variable change from ξ=0 to ξ=1 as soon as the level stop increasing (decreasing). For describing the disturbance scenarios for inflow rate a 2-entries vector $z(t)=(\rho(t),F_{in}(t))$ is used where $-1<\rho(t)<1$ is the local correlation coefficient calculated based on the current and 3 precedent values of $F_{in}(t)$. After each mode implementation, rewards given are calculated as the avergage outflow rate over the time interval between transitions as long as the tank does not overflow nor becomes empty. When an overflow or shut-down event occurrs a large negative reward is given to the state transition.

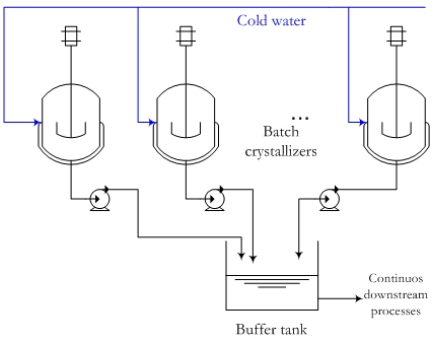


Fig. 2. Hybrid chemical plant

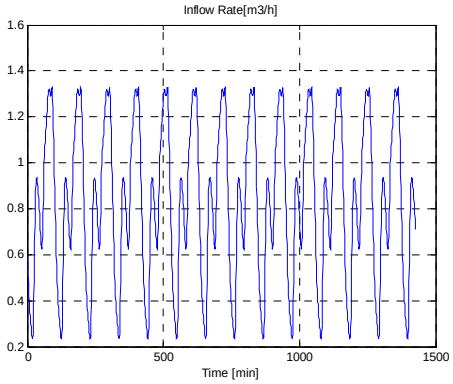


Fig. 3. Inflow rate variability pattern

In Eq. (6) above the feedback laws used for both to accumulate and drain-off are given. As a result four different modes are available to maximize throuhgput whilst avoiding tank overflow or drownstream feeding interruptions. Results obtaned for some input disturbance scenarios are summarized in Fig. 4 through 7. It's noteworthy in Fig. 4 and 5 that despite abrupt changes in this scenario for the inflow rate (see Fig. 3) the multi-modal control based on the learned *Q*-values is able to maintain the outflow rate significantly high. Note that the tank is operated quite close the maximum level to maximize the average productivity whereas mode switching is used to avoid overflow. The corresponding control program is the mode sequence π: (A1-D1-A1-D2).

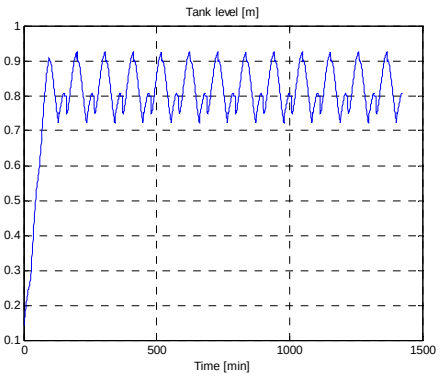


Fig. 4. Tank level under multi-modal control

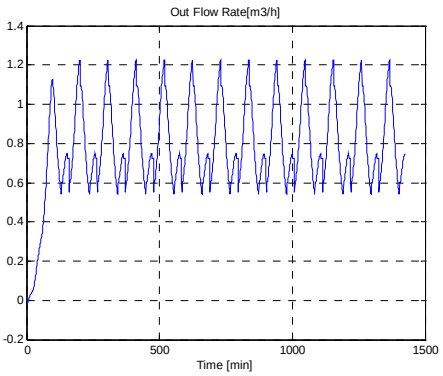


Fig. 5. Outflow rate under multi-modal control

Finally, let's consider a rather different scenario for the inflow rate and tank situation when its initial level is very low. As shown in Fig. 6 and 7 the multi-modal control program restores the maximum productivity condition in the tank using only mode A1.

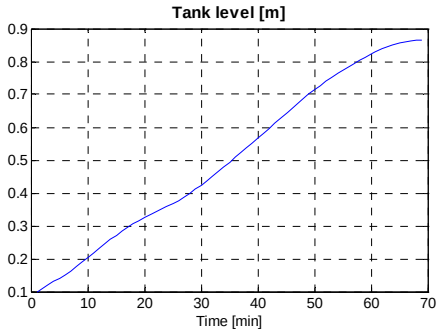


Fig. 6. Tank level recovery

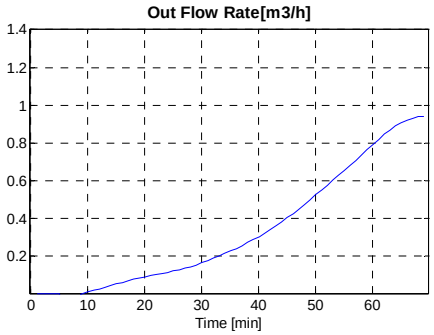


Fig. 7. Outflow rate maximization

4. Final remarks

A new approach for automatic generation of a multi-modal control program based on simulations and reinforcement learning has been presented using buffer tank control for productivity maximiazation in a hybrid chemical plant. The main advantage of the proposed methodology is that if the inflow rate pattern changes the *Q*-function for state-modes in the program will change and the multi-modal control policy will be updated.

5. Acknowledgements

The authors thanks to the ANPCyT of Argentina which partially supports the work presented here through project PICT 1099/06.

References

P. Barton, C. Lee, M. Yunt, 2006. Optimization of hybrid systems, Computers chem Engng, 30, 1576-1589.

S. Engell, S. Kowalewski, C. Schulz and O. Stursberg, 2000, Continuous-Discrete Interactions in Chemical Processing Plants, Proceedings of the IEEE 88, 1050-1068.

M. Ghaeli, P. Bahri, P. Lee, 2008, Scheduling of a mixed batch/continuous sugar milling plant using Petri nets, Computers chem Engng, 32, 580-589.

X. Koutsoukos, P. Antsaklis, J. Stiver and M. Lemmon, 2000, Supervisory Control of Hybrid Systems, Proceedings of the IEEE 88, 1026-1049.

D. Liberzon, 2003, Switching in Systems and Control, Systems & Control: Foundations & Applications, Birkhäuser Boston Inc., Boston, MA.

K. McDonald, T. McAvoy, and A. Tits, 1986, Optimal averaging level control, AIChE J., 32, 75–86.

T. Mehta and M. Egerstedt, 2006, An optimal control approach to mode generation in hybrid systems, Nonlinear Analysis 65, 963–983.

S. Melas, 2003, Pvc line predictive inventory control: a production rate optimization for a hybrid system, Solvay-Group report, 1–34.

I. Simeonova, 2008, On-line periodic scheduling of hybrid chemical plants with parallel production lines and shared resources, PhD Thesis, Université Catholique de Louvain.

R. Sutton and A. Barto, 1998, Reinforcement learning:An Introduction, The MIT Press, Cambridge, MA.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Advanced Temperature Tracking Control for High Quality Wines using a Phenomenological Model

Oscar A. Ortiz[a], Martha D. Vallejo[b], Gustavo J. E. Scaglia[a], Carmen A. Mengual[a] and Pablo M. Aballay[a]

[a]*Instituto de Ingeniería Química, Universidad Nacional de San Juan, Av. Lib. San Martín Oeste 1109, J5400ARL San Juan, Argentina*
[b]*Instituto de Biotecnología, Universidad Nacional de San Juan, Av. Lib. San Martín Oeste 1109, J5400ARL San Juan, Argentina.*

Abstract

The fermentation in winemaking has high complexity due to interactions between cell biokinetics and bioreactor hydrodynamics. Hence, the new technological options to obtain high quality wines with outstanding organoleptic characteristics require more strict process monitoring and control based on rigorous models. This work presents an advanced temperature control system based on an improved non-isothermal phenomenological model that allows tracking complex temperature profiles to achieve optimal quality of wine. The controller has been performed in discrete-time, so that the current disturbance effect at the output is computed as the difference between the current measured value of the output and the predicted one. The obtained results are satisfactory for experimental data from literature.

Keywords: Dynamic inverse control**,** temperature tracking, wine fermentation.

1. Introduction

The growing customers demand for high quality wines and its marked preferences for outstanding wine organoleptic properties, presents new challenges for the winemaking. Hence, new technologies are recently emerging. Some of them include variable temperature trajectories throughout the fermentation. Indeed, there are strong evidences about the contribution of temperature to the release of some varietal aroma precursors and to the synthesis of volatile compounds by the wine yeast. Temperature affects metabolism and gene expression of the yeast, and is also an important variable in the maceration phase (extraction and diffusion of sugars, acids, aroma/taste precursors, etc., from the grape), that may be carried out in a pre-fermentation stage or during the fermentation. Temperature is seen as a fine tool to a better regulation of the fermentation progress, in accordance with a winemaking strategy depending on the desired wine (Torija *et al*., 2003; Molina *et al*., 2007; Romero Cascales, 2008). On the other hand, fermentation presents great interactions between cell biokinetics and bioreactor hydrodynamic conditions, which lead to models with non-linear and unsteady characteristics. Therefore, advanced optimization and control tools for monitoring and controlling strictly the process, are required. The implementation of advanced control strategies needs appropriate dynamic models and reliable on-line measurements (Henson, 2003). So, more rigorous and accurate models are necessary. They must be at the same time precise and simple, from a mathematical point of view, to be used in on-line control algorithms. In previous contributions, the authors have addressed isothermal

and non-isothermal first principles and hybrid models, and an improved isothermal phenomenological model with satisfactory capability to approximate the wine fermentation profiles (Vallejo *et al.*, 2005; Ortiz *et al.*, 2006; Aballay *et al.*, 2008; Scaglia *et al.*, 2009).

In this work, an advanced temperature control system for an enological batch alcoholic fermentation is presented. The control system use a non-isothermal phenomenological model that couples mass and energy balances to predict the main process state variables: viable cells, sugar and ethanol concentrations, carbon dioxide released, and the bioreactor temperature. The energy balance takes into account, the heat exchanged between the bioreactor and its cooling jacket and, the cellular growth rate dependence on temperature. Variations of specific heat and density in the bulk have been included in the model. Section 2 presents a brief description of the fermentation in winemaking and its modelling. Also, the non-isothermal model with considerations on each one of its components and parameters, as well as on its main properties is described in this section. In section 3, the dynamic model performance is evaluated by means of a wide performance analysis. The most frequently found disturbances in actual plants are applied on operating process variables and main model parameters and, the corresponding responses of the state process variables are shown. The control problem is addressed in section 4 where the main control objective has been posed on temperature, which is required to change according to the fermentation behavior and technological options, such as maceration or to promote enzymatic activities. The proposed control on temperature is based on the manipulation of cooling water flow rate through a dynamic inverse controller. The corresponding simulations of the controller response, with and without model parameters errors, are shown. Also, the errors between the control system response and the reference trajectory are analyzed. Section 5 presents the discussion of the main obtained results, analyzes the contribution of the work and outlines some topics to be addressed in future works.

2. Non - isothermal model

The grape juice fermentation is the winemaking heart. Since it is operated on batch mode, yeasts must get rapidly adapted to highly variable environmental conditions, which greatly difficult the kinetic predictions of the bioprocess. Numerous and complex physical, chemical and biological changes occur at this stage and fix on the final wine quality (Boulton *et al.*, 1996; Cramer *et al.*, 2002; Rosenfeld *et al.*, 2003). On the other hand, some outstanding organoleptic properties only are achieved tracking rigorously predefined temperature trajectories. For example, maceration at around 28-30ºC is generally required at the beginning of the fermentation, because enzymes are affected by the increasing ethanol concentration and better enzymatic activities are achieved at these "high" temperatures; 28ºC is appropriate for a fast start-up of the yeast multiplication, favoring the implantation of the desired population and inducing the synthesis of flowery related aroma compounds; temperatures between 23 and 18ºC are standard in winemaking and avoid ethanol and volatile compound losses; or definite stages at low temperature (near 15ºC) often stated to produce compounds related to fresh and fruity aromas in the wines, maintaining varietal aromas (Esti and Tamborra, 2006; Beltrán *et al.*, 2008). Significant but controllable and not stressing changes occur in the fermentative activity when temperature is slightly changed. Such control and monitoring requirements, establish each more time the availability of more accurate non-isothermal models. The reductive metabolic pathway S ➔ X + P + CO_2, characterizes the yeast population growth in anaerobic conditions. X, S, P and CO_2

correspond to viable cells, carbon source (glucose and fructose) and ethanol concentrations, and carbon dioxide released [kg m^{-3}], respectively. The metabolite accumulation in the extra cellular medium has been modelled by a set of ODE based on mass balances on X, S, P and CO_2 like in the isothermal model of Scaglia *et al.* (2009). The mentioned variables progress in time t [h] and the (bioreactor) temperature T [K] influence on maximum specific cell growth rate μ_m [h^{-1}] are:

$$\{dX/dt,\ dS/dt,\ dCO_2/dt \text{ and } dP/dt\} = f(X, S, CO_2, \mu_m(T)) \quad (1)$$

The non-isothermal kinetic model is constituted by the previous mass balances and the energy balance in the reactor and its cooling jacket (Eq. 2), coupled by means of Arrhenius' equation (Aballay *et al.*, 2008). Assumptions are: heat transfers by radiation and conduction and heat losses due to CO_2 evolution, water evaporation and ethanol and flavor losses are negligible. Water properties and the bulk pH were considered constant.

$$\frac{d(\rho_r \cdot V_r \cdot Cp_r \cdot T)}{dt} = Y_{H/CO_2} \cdot V_r \cdot \frac{dCO_2}{dt} - F_w \cdot Cp_w \cdot (T_{w,out} - T_{w,in}) \quad (2)$$

V_r [m^3] is the volume.$Y_{H/CO2}$ [W·h produced/kg CO_2 released] is the energy due to the carbon dioxide released by the bio-reaction, and the second term in the right side represents the exchanged heat between the fermenting mass and the cooling jacket [W]. F_w is the cooling water flow rate [kg h^{-1}] and C_{pw} is specific heat of water [W·h kg^{-1} K^{-1}]; the cooling water temperature [K] in the jacket inlet is $T_{w,in}$ and in the jacket outlet is $T_{w,out}$ (see details in Aballay *et al.*, 2008). ρ_r [kg m^{-3}] and C_{pr} [W·h kg^{-1} K^{-1}] are density and specific heat of the fermenting mass, which were incorporated to the model as variables to achieve more accuracy in process state prediction, approximately a 10%.

3. Dynamic model performance analysis

For this purpose, steps changes on most probable manipulated variable in practice, such as the cooling water flow rate, were introduced. A decrease or shut-down in the cooling water flow can affect the most critical variable for wine quality, the temperature. Model performance was evaluated considering responses of parameter μ_m and, the T and CO_2 evolution to negative steps, -33% and -70% of nominal F_w (Fig. 1). The changes produce a positive increase on: maximum growth rate (up to +20%), (a), and maximum temperature (up to +25%), (b). Moreover, an early positive increase in carbon dioxide released (+14%) at t = 88 hours can be observed, (c). Therefore, the dynamic model performance is satisfactory.

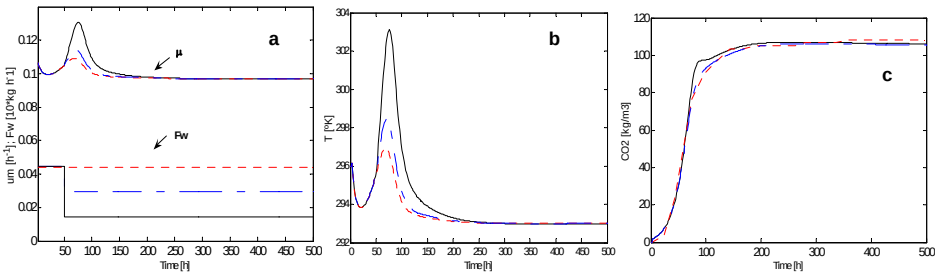


Figure 1. Model responses to steps on the cooling water flow rate. $F_{w,nominal}$ = 0.448 kg h^{-1} (- - -); $F_{w,}$ = 0.30 kg h^{-1} (— - —); $F_{w,}$ = 0.143 kg h^{-1} (—). (a) Maximum specific growth rate and water flow rate. (b) Bioreactor temperature. (c) Carbon dioxide released.

Most probable changes on other mass balances parameters in the model are shown in a wide performance analysis (Scaglia *et al.*, 2009). Regarding to the main energy balance parameters, they were varied in ±20% and the corresponding responses on the state process variables have been obtained with satisfactory results (not shown).

4. Control problem

The proposed control on temperature is based on the manipulation of cooling water flow rate through a dynamic inverse controller. Previously knowing the process model and the desired temperature profile ($T_{ref}(t)$), this control strategy allows to generate the corresponding control actions in such a way that the tracking error ($e(t)$) tend to zero. The controller has been performed in discrete-time, so that the current disturbance effect at the output is computed as the difference between the current measured value of the output and the predicted one.

The values of $T(t)$ at discrete time $t = n\ T0$, where $T0$ is the sampling period and $n \in \{0,1,2,\ldots\}$, will be denoted as $T(n)$. From Eq. 2, the value of $T(n+1)$ is,

$$T(n+1) = T(n) + \int_{nT0}^{(n+1)T0} \frac{1}{\rho_r V_r Cp_r}\left[Y_{H/CO_2} V_r \frac{dCO_2}{dt} - F_w Cp_w \left(T_{w,out} - T_{w,in} \right) \right] dt \qquad (3)$$

where, F_w remains constant through the interval $nT0 \leq t < (n+1)T0$. Using the Euler approximation to calculate the integral, we have:

$$T(n+1) \approx T(n) + \frac{T0}{\rho_r V_r Cp_r}\left[Y_{H/CO_2} V_r \frac{dCO_2(n)}{dt} - F_w(n) Cp_w \left(T_{w,out}(n) - T_{w,in}(n) \right) \right] \qquad (4)$$

$$F_w(n) = \left\{ \left[T_{ref}(n+1) - k_T \left(T_{ref}(n) - T(n) \right) - T(n) \right] \frac{\rho_r V_r Cp_r}{T0} - Y_{H/CO_2} V_r \frac{dCO_2}{dt} \right\} \Big/ \ldots\ldots - \left[Cp_w \left(T_{w,out} - T_{w,in} \right) \right] \qquad (5)$$

Equation 5 has only the error due to the integral approximation, whose maximum value is A (Eq. 6) (Boyce and DiPrima, 2005), where clearly, if the sampling time is small, also the error will be small. Replacing 5 in 4 we obtain 6:

$$A = \max_t \left(\frac{1}{2} \frac{d^2 T}{dt^2} T0^2 \right) \; ; \quad \underbrace{T_{ref}(n+1) - T(n+1)}_{e(n+1)} - k_T \underbrace{\left(T_{ref}(n) - T(n) \right)}_{e(n)} = 0 \qquad (6)$$

In Eq. 6, it can be seen that if $0 < k_T < 1$, then $e(n) \rightarrow 0$ when $n \rightarrow \infty$, where $e(n)$ is the tracking error in discrete form. Besides, since in each sampling period the process variables are measured, there are no accumulative errors. Nevertheless, a term ($E(n)$), that takes into account the approach error of the integral and the modeling error, is included in Eq. 4 (See Eq. 7). Here, the terms with cup (ˆ) represent the estimated values of the actual system parameters. Finally, considering $E(n)$ the control action is expressed by Eq. 8.

$$T(n+1) = T(n) + \frac{T0}{\hat{\rho}_r \hat{V}_r \hat{Cp}_r}\left[\hat{Y}_{H/CO_2} \hat{V}_r \frac{dCO_2}{dt} - F_w(n) \hat{Cp}_w \left(T_{w,out} - T_{w,in} \right) \right] + E(n) \qquad (7)$$

$$F_w(n)=\left\{\left[T_{ref}(n+1)-k_T\left(T_{ref}(n)-T(n)\right)-T(n)-E(n)\right]\frac{\hat{\rho}_r\hat{V}_r\hat{C}p_r}{T0}-\hat{Y}_{H/CO_2}V_r\frac{dCO_2}{dt}\right\}\Big/ \\ \ldots\ldots-\left[\hat{C}p_w\left(T_{w,out}-T_{w,in}\right)\right] \tag{8}$$

Figures 2a and 2b show the temperature change and the control action respectively, when the latter is calculated with Eq. 5 for a small sampling period. It can be seen how the system temperature tracks closely the reference trajectory. Also, it is observed that such tracking is achieved with a finite control action (Fig. 2b). On the other hand, errors in model parameters significantly influence the process states estimation, carried out by the multi-parametric and good performance model (Scaglia *et al.*, 2009; Aballay *et al.*, 2008), and consequently on the control law. Figure 2c shows the bioreactor temperature change when some model parameters contain error. In this case, the control signal was calculated with Eq. 8, but with $E(n) = 0$. If Figs. 2a and 2c are compared it is observed that the parameters errors produce a considerable deviation between the bioreactor and reference temperature profiles (Fig. 2c). The inclusion of term $E(n) \neq 0$ in the proposed control signal (Eq. 8) improves significantly the temperature tracking (Fig. 3a). Besides, the good performance of the system can be seen in Fig. 3b that shows the comparison with one of the literature data. X profile for the T_{ref} tracking control shows a satisfactory reduction in the cellular death (from 150 h) with respect to experimental fermentation FER1 in Aballay *et al.* (2008). In this latter case without a T_{ref} profile tracking, cellular death achieved is total.

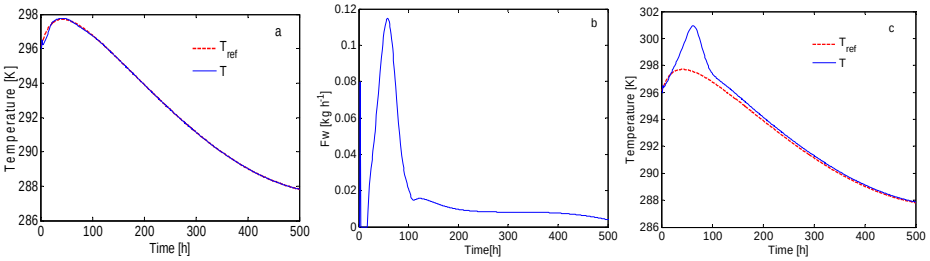


Figure 2. Temperature tracking. (a) Reference and bioreactor temperature profiles. (b) Control action on cooling water flow rate. (c) Reference and bioreactor temperature profiles when there are errors in the model parameters.

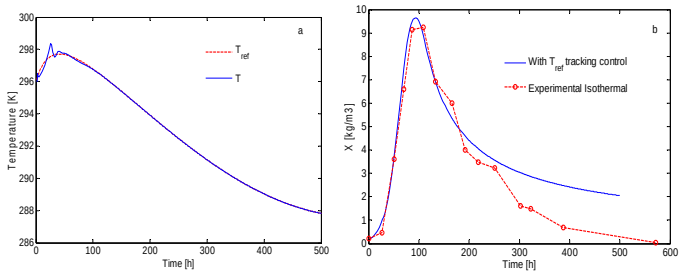


Figure 3. (a) Reference and bioreactor temperature profiles when the modelling errors are considered. (b) Viable cells profile: experimental isothermal fermentation and with the proposed control scheme for the reference temperature profile.

As can be seen from Figs. 2 and 3, the main control objective is achieved using a finite control signal, in spite of the model can have modeling errors.

5. Conclusions

An advanced temperature tracking control for high quality wines have been presented. The control system is based on an improved non-isothermal model that predicts closely the main fermentation state variables and allows a strict process control, when non-constant temperature profiles must be tracked to achieve distinctive fine wines. The proposed control manipulates the cooling water flow rate through a dynamic inverse controller. The system shows a good performance tracking narrowly a predefined reference trajectory with a finite control signal, even though the model can present errors in its parameters. The system was satisfactorily tested with literature experimental data. In next contributions, the system will be extended and improved to tracks other important process variables such as density and cell concentration.

References

P.M. Aballay, G.J.E. Scaglia, M.D. Vallejo, and O.A. Ortiz, 2008, Non isothermal phenomenological model of an enological fermentation: modelling and performance analysis, 10th International Chemical and Biological Engineering Conference CHEMPOR 2008, Braga, Portugal, 4-6 September.

G. Beltran, M. Novo, J.M. Guillamón, A. Mas, and N. Rozès, 2008, Effect of fermentation temperature and culture media on the yeast lipid composition and wine volatile compounds, International Journal of Food Microbiology, 121, 169–177.

R. Boulton, V.L. Singleton, L. Bisson, and R. Kunkee, 1996, Yeast and biochemistry of wine fermentation, Principles and practices of winemaking, Chapmann and Hall, New York.

W.E. Boyce, and R. C. DiPrima, 2005, Elementary Differential Equations and Boundary Value Problems, 8th Edition, John Wiley & Sons Limited, Australia.

A.C. Cramer, S. Vlassides, and D.E. Block, 2002, Kinetic model for nitrogen-limited wine fermentation, Biotechnol. Bioeng., 77, 49-60.

M. Esti, and P. Tamborra, 2006, Influence of winemaking techniques on aroma precursors, Analytica Chimica Acta, 563, 173–179.

M.A. Henson, 2003, Dynamic modeling and control of yeast cell populations in continuous biochemical reactors, Computers & Chemical Engineering, 27, 8-9, 1185-1199.

A.M. Molina, J.H. Swiegers, C. Varela, I.S. Pretorius, and E. Agosin, 2007, Influence of wine fermentation temperature on the synthesis of yeast-derived volatile aroma compounds, Appl. Microbiol. Biotechnol., 77, 675–687.

O.A. Ortiz, P.M. Aballay, and M.D. Vallejo, 2006, Modelling of the Killer Yeasts Growth in an Enological Fermentation by means of a Hybrid Model, XXII Interamerican Congress of Chemical Engineering, A. 13b-224, 451-452, Buenos Aires, Argentina, 1-4 October.

I. Romero Cascales, 2008, Extracción de compuestos fenólicos de la uva al vino. Papel de las enzimas de maceración, Tesis Doctoral (PhD Thesis), Universidad de Murcia, Spain.

E. Rosenfeld, B. Beauvoit, B. Blondin, and J. Salmon, 2003, Oxygen consumption by anaerobic *Saccharomyces cerevisiae* under enological conditions: effect on fermentation kinetics, Appl. Environ. Microbiol., 69, 113-121.

G.J.E. Scaglia, P.M. Aballay, C.A. Mengual, M.D. Vallejo, and O.A. Ortiz, 2009, Improved phenomenological model for an isothermal winemaking fermentation, Food Control, doi: 10.1016/j.foodcont.2008.12.012

M.J. Torija, G. Beltran, M. Novo, M. Poblet, J.M. Guillamón, A. Mas, and N. Rozès, 2003, Effects of fermentation temperature and Saccharomyces species on the cell fatty acid composition and presence of volatile compounds in wine, Int. J. Food Microbiol., 85, 127–136.

M.D. Vallejo, P.M. Aballay, M.E. Toro, F. Vazquez, G.I. Suarez, and O.A. Ortiz, 2005, Hybrid Modeling and Neural Prediction of the Wild Killer Yeast Fermentation Performance in a Winemaking Process, 2nd Mercosur Congress on Chemical Engineering and 4th Mercosur Congress on Process Systems Engineering, Paper code 230, 1-10, Rio de Janeiro, Brazil, 14-18 August.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

MPC of a Four-Stage Grape Juice Evaporator based on an Adaptive Neural Model

Graciela I. Suárez, Gustavo J.E. Scaglia, Pablo M. Aballay and Oscar A. Ortiz

Instituto de Ingeniería Química, Universidad Nacional de San Juan, Av. Lib. San Martín Oeste 1109, J5400ARL San Juan, Argentina

Abstract

The four-stage evaporator is the core of the process in the manufacture of concentrated grape juice. The dynamic features of this process are very complex due to inputs/outputs constraints, time delays, loop interactions and the persistent unmeasured disturbances that affect it. This paper addresses the non-linear control of an industrial-scale multiple-effect evaporator by means of a model predictive control (MPC) based on a neural network model (NNM). This strategy allows it on one hand, overcome the classical control systems limitations and, by the other that the neural controller continue learning during the plant operation. The achieved results allow us to conclude that the developed neural model predictive control is adequate to control effectively a four-effect evaporator to concentrate grape juice with different characteristics; its application at industrial scale is possible in the future.

Keywords: Adaptive model-based predictive control; artificial neural networks; grape juice concentration process

1. Introduction

The multi-step evaporation (M-SE) is the most important unit operation used in the food industry to concentrate juices of grapes and others fruits. Even when the main objective of this process is to produce a concentrated product, this should also possess certain organoleptic properties that are critical with respect to its quality and acceptance grade by the customers.

Product requirements and the complex characteristics of the process such as non-linear behavior, input and output constraints, time delays and loop interactions justify the use of an advanced control system. In this sense, the control of industrial evaporation processes with such constrains remains as a challenge for researchers.

Diverse control strategies have been intended for multi-stage evaporators. Among others the following works can be mentioned: conventional multi-loop linear feedback control, differential geometric control, generic model control GMC, multivariable generalized predictive control GPC, input-output linearization, differential geometric control using input-output linearization. Some authors have applied with success the differential geometric control to evaporation systems in the alumina industry (Kam and Tadé, 1999; Kam and Tadé, 2000). On the other hand, the model based predictive control (MPC) is today widely used in the industry and recognized as a powerful tool for high performance control of constrained multivariable processes. The main features of some of these industrial applications are available in Camacho and Bordons (1999).

In this work, a model predictive control (MPC) based on a neural network model (NNM) is presented. The neural network model is trained with data generated by simulation of a previously developed phenomenological model and with a portion of

plant data. The phenomenological model validation as well as the plant data used for NN training corresponds to a 1370 kg/h-production four-effect evaporator. The simulation data used for NN training include a range of most likely plant operating conditions introduced by means of pseudo-stochastic disturbances in the form of white noise on the manipulated variables; they are: feed flow rate and/or the heating steam temperature and, the range of variation is about ±10%. Afterwards, the NN model is validated with the other portion of the input-output plant data. This adaptive strategy is adopted, in order to obtain a more accurate process description without introducing disturbances in the real plant, which supposes a high cost, to obtain relevant data sets for training the NN.

In spite of a phenomenological model is available it is preferred the use of a NN model trained with data from simulation. The lack of sufficient amount of plant data does not permit an adequate training of the NN, however this approach has the advantage that at the beginning, the NN trained by means of simulation, at least imitate the performance of the phenomenological model. Therefore, when the NN model is coupled with the real plant in a MPC it can continue learning from the process and improve its performance through time. The first principles model is not used in the control strategy proposed because it has the inherent errors to the modelling of complex processes and, it does not have the capacity to adapt to the changing operation conditions of this kind of processes (Psichogios and Ungar, 1991; Hussain et al., 2003). NNs of different sizes are trained and those ones with the best performance with respect to the first principles model are selected.

The paper is organized as follow. In Section 2, a short description of a four-effect evaporator is presented, which is described by a set of non-linear ODEs. Section 3 presents the proposed control scheme, where mathematical and neural models are used to compute the control signal with the objective of considering the model variations due to changes in composition, organoleptic and physical properties of the grape-juice to be processed. Section 4 presents the discussion of the main obtained results, analyzes the contribution of the work and outlines some topics to be addressed in future works.

2. Process Description

Figure 1a shows the input and output streams in a vertical generic effect evaporator with long tubes. The solution to be concentrated circulates inside the tubes, while the steam, used to heat the solution, circulates inside the shell around the tubes.

The evaporator operates in co current. The solution to be concentrated and the steam are fed to the first effect by the bottom and by the upper section of the shell, respectively. Later on, the concentrated solution from the first effect is pumped to the bottom of the second effect, and so on until the fourth effect. On the other hand, the vapor from each effect serves as heater in the next one. Finally, the solution leaving the fourth effect attains the desired concentration.

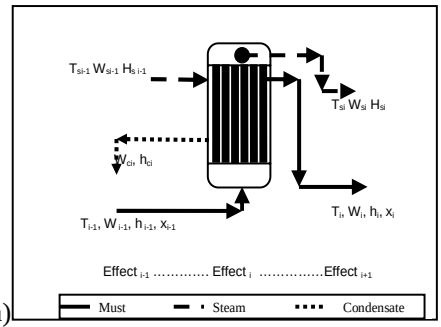
Each effect has a baffle in the upper section that serves as a drops splitter for the solution dragged by the vapor. The vapor from the fourth effect is sent to a condenser and leaves the process as a liquid. The concentrated solution coming from the fourth effect is sent to a storage tank.

Stefanov and Hoo (2003, 2004) have developed a rigorous model with distributed parameters based on partial differential equations for a falling-film evaporator, in which the open-loop stability of the model to disturbances is verified. On the other hand, various methods have been proposed in order to obtain reduced-order models to solve such problems (Christofides, 1998; El-Farra, Armaou and Christofides, 2003; Armaou

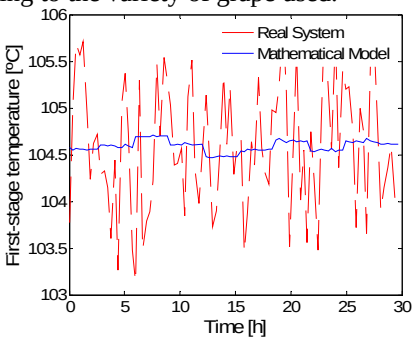
and Christofides, 2002; Hoo and Zheng, 2001; Zheng and Hoo, 2002). However, there is not a general framework yet, which assure an effective implementation of a control strategy in a multiple effect evaporator.

In practice, due to a lack of measurements to characterize the distributed nature of the process and actuators to implement such a solution, the control of systems represented by partial differential equation (PDE) in the grape juice evaporator, is carried out neglecting the spatial variation of parameters and applying lumped systems methods. However, a distributed parameters model must be developed in order to be used as a real plant to test advance control strategies by simulation.

In this work, it is used the mathematical model of the evaporator developed by Ortiz *et al.* (2006), which is constituted by mass and energy balances in each effect. The assumptions are: the main variables in the gas phase have a very fast dynamical behavior, therefore the corresponding energy and mass balances are not considered (Newell and Lee, 1989). Heat losses to surroundings are neglected and the flow regime inside each effect is considered as completely mixed (Newell and Fisher, 1972). Figure 1b shows real and simulated responses of the system, where mathematical model tracks accurately the real system. The model parameters were adjusted for one variety of grape juice. The advantage of working with on-line adaptive neural networks is that the network learns the model with new organoleptic properties and adapts to the characteristics of each grape juice variety, according to the variety of grape used.



a)



b)

Figure 1. (a) Effect i in the four-stage evaporator flowsheet, i = 1,…, 4. (b) Real system and mathematical model responses.

3. Model-based predictive control (MPC)

The MPC is composed by two main components: the plant model and the controller – optimizer. Figure 2a shows the scheme implemented in the process. For the MPC (Camacho and Bordons, 1999), the following variables that determine its operation are defined: prediction horizon, the penalty coefficient of the control signal, and the control horizon.

The MPC algorithm allows defining future control actions while the next function is minimized.

$$J = \sum_{N_{1k}}^{N_{2k}} \gamma_k \left(\hat{y}_k(t-j) - Yref_k(t+j)\right)^2 + \sum_{1}^{Nu_k} \beta_k \left(u_k(t+j-1) - u_k(t+j-2)\right)^2 \qquad (1)$$

N_{1k} and N_{2k} constitute the initial prediction horizon for each output and N_{uk} is the control horizon for each input. y is the first-stage temperature, Y_{ref} is the temperature profile to

be tracked by the system and *u* is the heating steam temperature that enters to evaporator.

The success of Model Predictive Control (MPC) control performance is highly dependent on the accuracy of the open-loop predictions, which in turn depends on the accuracy of the process models. It is possible for the predicted trajectory to differ, perhaps considerably, from the actual trajectory followed by the plant being controlled (Berber, 1995). The difference between the plant and the model is known as plant-model mismatch, which can cause the control performance to be sluggish, overly conservative or, in the worst-case scenario, unstable. On the other hand, closed-loop performance may be poor if the plant differs significantly from the model. In the case of the evaporator, grape juice physicochemical parameters change with the varietal type to be processed. In Fig. 2b shows the control system response, when the phenomenological model is used in the MPC scheme, for a step change introduced in the reference, when the used model accurately describes the plant and when there are errors in the model. As can be seen, when the model precisely describes the plant to be controlled, the output signal tracks the reference signal in accurate way, whereas when there are errors in the model, the plant increases the difference between reference and system response signals, respectively. A neural network is used as model to solve this problem. The plant model identification was carried out using a radial basis function neural network (RBF-NN), due to its complexity. The training of the network was carried out by means of the Neural Network Toolbox of Matlab^{TM}. For the training process, the Levenberg - Marquardt Learning Law was used with 2000 pair of input – output values obtained by simulation. The simulator was excited with a pseudo-stochastic multi-binary signal (PRMS) over the manipulated variables (heating steam temperature), to obtain signals which produce an enough wide variation in the system operation conditions.

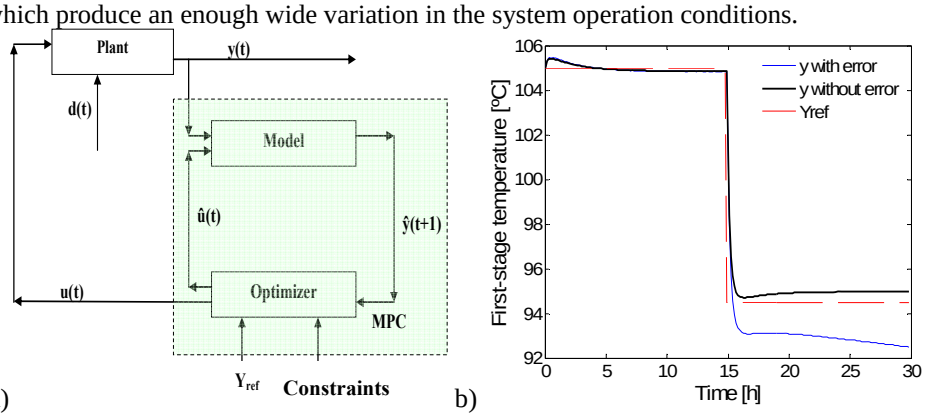
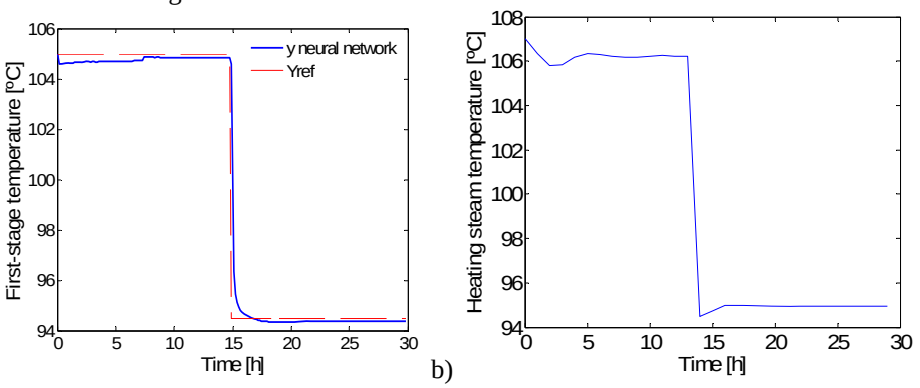


Figure 2. (a) MPC scheme. (b) System response with and without modelling errors when a step change occurs in the reference output.

A data set with 1000 pairs of values was used for training and other 1000 pairs were used in the validation stage. The sum of the quadratic errors between the NN model output and the simulation results from the validation data set have been considered in order to select the most suitable NN. The network with the better performance has 15 neurons in the hidden layer and a linear function in its output layer. The obtained neural model presents a maximum error less than 4% with respect to validation data. Then, this model is used for implementing the predictive control strategy so that the NN learns the plant behavior. Figure 3a shows the system response when the NN is used as the plant model. It can be seen that, when there are errors between the model used to train the NN and the plant to be controlled, the network can learn that difference and improve the

system performance. Moreover, it can be seen how the error is decreasing while the NN learns the system behavior and when a step change in the reference occurs, the system shows a better behavior than the observed in Fig. 2b, where the used model has errors in its parameters. In Fig. 3b it can be seen that the control objective is achieved with a bounded control signal.



a) b)

Figure 3. (a) System response when the model is a NN. (b) Control action.

4. Conclusions

In this work, a model-based predictive controller for an industrial-scale four-stage grape juice evaporator is presented, considering two alternative scenarios. First, the MPC was implemented with a phenomenological model with its parameters adjusted for a specific grape juice. Second, a NN model, obtained with data from simulation of a phenomenological model validated with industrial data, was used for the MPC. When there are no modelling errors both control alternatives show a suitable performance. In case of the model has errors, i.e., due to organoleptic and physicochemical properties variations in the grape juice, it is observed that by using a NN model, the MPC controller has a better system performance, since the NN can learn and correct the difference between the phenomenological model and the process to be controlled. The use of a MPC based on an on-line adaptive NN model allows that the controller responds adequately when the process is fed with grape juices that contain different organoleptic properties. These properties vary with the grape variety to be used. The development of an advanced control algorithm using a NN which takes variables coupling (first-stage temperature and final product concentration) into account as process model, will be addressed in a next work.

References

A. Armaou, and P.D. Christofides, 2002, Dynamic Optimization of Dissipative PDE Systems Using Nonlinear Order Reduction. Chemical Engineering Science 57, 24, 5083-5114.

R. Berber, 1995, Methods of Model Based Process Control, in NATO ASI Series. Series E, Applied Sciences, 293, Kluwer Academic Publishers, Dordrecht (The Netherlands).

E.F. Camacho, and C. Bordons, 1999, Model Predictive Control, Springer-Verlag.

P.D. Christofides, 1998, Robust Control of Parabolic PDE Systems, Chemical Engineering Science, 53, 16, 2949-2965.

N.H. El-Farra, A. Armaou, and P.D. Christofides, 2003, Analysis and Control of Parabolic PDE Systems with Input Constraints, Automatica, 39, 4, 715-725.

K.A. Hoo, and D. Zheng, 2001, Low-Order Control Relevant Model for a Class of Distributed Parameter Systems, Chemical Engineering Science, 56, 23, 6683-6710.

M.A. Hussain, 2003, in Cornelius T. Leondes (Ed.), Neural Network Techniques and Application in Chemical Process Control System, 5, V327–V360, CRC Press.

K.M. Kam, and M.O. Tadé, 1999, Case studies on the modelling and control of evaporation systems, XIX Interamerican Congress of Chemical Engineering COBEQ.

K.M. Kam, and M.O. Tadé, 2000, Simulated Nonlinear Control Studies of Five Effect Evaporator Models, Computers and Chemical Engineering, 23, 1795 - 1810.

R.B. Newell, and P.L. Lee, 1989, Applied process control: A case study, Prentice Hall.

R.B. Newell, and D.G. Fisher, 1972, Model development, reduction and experimental evaluation for an evaporator, Eng. Chemical Process Res. Develop., 11, 2, 213-221.

O.A. Ortiz, G. I. Suárez, and C. A. Mengual, 2006, Adaptive Neural Model Predictive Control for the Grape Juice Concentration Process, in Proceedings of XXII Interamerican Chemical Engineering Congress and V Argentinian Congress of Chemical Engineering. Innovation and Management for Sustainable Development, Buenos Aires, Argentina, 1-4 October.

D.C. Psichogios, and L.H. Ungar, 1991, Direct and indirect model based control using artificial neural networks, Ind. Eng. Chem. Res., 30, 12, 2564-2573.

Z.I. Stefanov, and K.A. Hoo, 2003, A Distributed-Parameter Model of Black Liquor Falling Film Evaporators. Part I. Modeling of Single Plate, Ind. Eng. Chem. Res., 42, 1925-1937.

Z.I. Stefanov, and K.A. Hoo, 2004, Distributed Parameter Model of Black Liquor Falling-Film Evaporators. 2. Modeling of a Multiple-Effect Evaporator Plant, Ind. Eng. Chem. Res., 43, 8117-8132.

D. Zheng, and K.A. Hoo, 2002, Low-Order Model Identification for Implementable Control Solutions of Distributed Parameter Systems, Computers and Chemical Engineering, 26, 7-8, 1049-1076.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

On l_1-Predictive Control of Mixed-Logical Dynamical Systems

David L. de Souza Júnior[a], Luís C. Oliveira-Lopes[a]

[a]*FEQUI/UFU – School of Chemical Engineering, Federal University of Uberlandia, Av. João Naves de Ávila, 2121, Santa Mônica, CEP 38408-902, Uberlândia – MG , Brazil*

Abstract

Constant technological innovations and environmental and economical requirements have created new challenges in industrial process control. The hybrid system predictive control is a representative example of great interest in this modern scenario. Hybrid behavior is found by continuous operation dynamics inserted with discrete event changes that might occur by internal and external actions of the system. Internal actions that generate hybrid behavior can be faults, changes between operation modes and perturbations. External actions can be represented by logical and decisory aspects and safety constraints. The chemical process operation often deals with controlled transitions among operation modes due to changes in the raw materials, energy sources, product specification and market demand. In this paper, a hybrid systems representation of linear systems known as mixed logical dynamical models (MLD) is investigated. The paper addresses the MLD l_1-Model Predictive Control (MPC) problem. Structurally, the MLD model control problem is described through a linear mixed integer optimization problem (MILP), which is computationally demanding, but they can be satisfactorily applied to systems of certain speed and dimension. The paper presents a study on the characteristics of the controller. A system of interest is used to illustrate the study. The results show that the application of MLD in the MPC control framework can have a positive impact specially when there is the inclusion of logical decisions aspects and operational knowledge in the MPC problem.

Keywords: Hybrid systems, MPC, MLD.

1. Introduction to Hybrid Systems

The term *hybrid* refers basically to a mixture of two or more methods or objects. The expression *hybrid system* was originally proposed by Witsenhausen (1966) to describe a combination between continuous dynamical systems and discrete event systems. Hybrid Systems come from the coupling between digital controllers, computers and subsystems built as finite-state machines, and plants modeled by partial or ordinary differential equations, or difference equations (Lazar, 2006).

Hybrid systems can be composed by several operating modes, endowed with a proper equations set, therefore they are capable of showing various dynamical behaviors simultaneously. Amongst these behaviours are continuous and discrete-time dynamics, jump discontinuity and transitions driven by logical commands. These systems are capable to select the activation of a specific mode when a particular event occurs. Figure 1 shows this fact. In a transition among operating modes, discontinuities may happen.

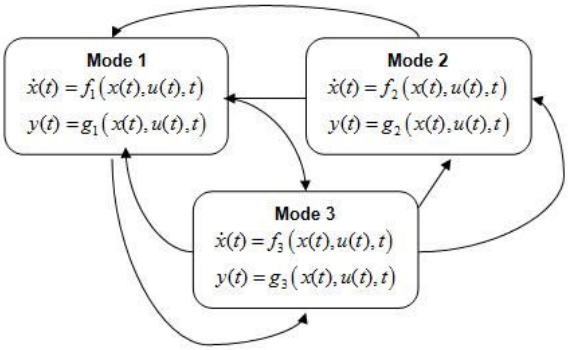


Figure 1 – Hybrid system sketch composed by three operational modes. The transitions are caused by events.

An event can be classified according to its origin. A *state event* occurs when variables reach a specific value. A *time event* occurs by elapsing a definite time period. And an *input event* occurs when external inputs influence the system (Lazar, 2006).

Although hybrid systems are capable to retain an appreciable variety of dynamic behaviors, a hybrid system general formulation introduces high levels of complexity for analysis purposes, application of techniques to design of controllers, simulation and validation. This fact is responsible for the form as the research in hybrid systems is led. Instead of a general formalism, specific classes of systems are adopted to represent, with relative simplicity and fidelity, the process of interest. Distinguished classes of hybrid systems are: *Piecewise affine systems*, PWA (Sontag, 1981, Johansson and Rantzer, 1998, Ferrari-Trecate et al., 2002); *Mixed-logical dynamical systems*, MLD (Bemporad and Morari, 1999, Bemporad, 2004); *Linear complementarity systems*, LC (van der Schaft and Schumacher, 1998; Heemels et al., 2000); and *Max-min-plus-scaling* systems (MMPS) (De Schutter and van den Boom, 2001).

1.1. Mixed-logical dynamical systems

Mixed-logical dynamical systems are described by equations set 1:

$$x(k+1) = A_k x(k) + B_{1k} u(k) + B_{2k} \delta(k) + B_{3k} z(k) \tag{1.a}$$

$$y(k) = C_k x(k) + D_{1k} u(k) + D_{2k} \delta(k) + D_{3k} z(k) \tag{1.b}$$

$$E_{2k} \delta(k) + E_{3k} z(k) \geq E_{1k} u(k) + E_{4k} x(k) + E_{5k} \tag{1.c}$$

where $k \in \mathbb{Z}$; A_k, $B_{1k,2k,3k}$, C_k, $D_{1k,2k,3k}$, $E_{1k,2k,3k,4k,5k}$ are matrixes with appropriate dimensions. And also, the following domains are defined by the set 2:

$$\begin{aligned} x &= \left[x_c \;\; x_l\right]^T, x_c \in \mathbb{R}^{n_c}, x_l \in \{0,1\}^{n_l}, n = n_c + n_l \\ y &= \left[y_c \; y_l\right]^T, y_c \in \mathbb{R}^{p_c}, y_l \in \{0,1\}^{p_l}, p = p_c + p_l \\ u &= \left[u_c \;\; u_l\right]^T, u_c \in \mathbb{R}^{m_c}, u_l \in \{0,1\}^{m_l}, m = m_c + m_l \end{aligned} \tag{2}$$

Vector **x** denotes the states, which are composed with continuous states ("c" index), and logical binary states, 0-1 ("l" index). The vector **y** represents the model output data. The input vector **u** encloses continuous commands ("c" index) and binary commands, 0-1 ("l" index). The MLD system is said to be well posed if **δ**(k) and **z**(k) are uniquely defined by inequality 1.c, in its domain, once **x**(k) and **u**(k) are specified. Thus, from

equation 1.a and 1.b, **x**(k+1) e **y**(k) are exclusively functions of **x**(k) and **u**(k). An interesting advantage is, in certain conditions, MLD systems can be equivalent to the other model classes in the section 1 (Heemels et al., 2001). This fact means that systems modeled by this kind of description can be converted into MLD systems that possess well definite techniques that favor its use for optimal and predictive control purposes.

In the MLD formulation, logical propositions can be rewritten in the format of linear inequalities by means of the transformation of Boolean variables in binary integer variables. Previous operational knowledge of a process can also be added systematically under the form of inequalities. Together with the equations of type 1.a and 1.b, which dictate the dynamics of the process, these linear inequalities form the set of constraints of an optimization problem which is base for optimal control problem synthesis (Bemporad e Morari, 1999; Krilavicius, 2006).

2. MLD Systems Optimal Control Aspects

Optimal control target is to make possible the synthesis of control law which deals with certain purposes. Such purposes are congregated in a performance index. It is desirable that this index is minimized subject to the process constraints. It is possible to penalize efforts of control (energy) and input/output reference deviation.

Control optimal synthesis yields an optimization problem. The resulting optimization problem for a MLD system can have difficult solution; therefore it involves mixed-integer programming (MILP, MIQP, and MINLP). It is also complicated the performance index selection and this issue demands an iterative procedure for its adequacy to the problem.

2.1. Definition of MLD system predictive control

Given to an initial state $\mathbf{x}_0$ and a final instant T, it is desirable to find, if possible, a sequence of control $\boldsymbol{u}$, defined as: $\boldsymbol{u} \equiv \{u(0), u(1),..., u(T-1)\}$, that leads the system of the state $\boldsymbol{x}_0$ to $\boldsymbol{x}_f$, minimizing the following performance index:

$$J\left(u_0^{T-1}, x_0\right) = \sum_{k=0}^{T-1} \left\| \Delta u(k) - \Delta u_f \right\|_{Q_1}^p + \left\| \delta\left(k, x_0, u_0^k\right) - \delta_f \right\|_{Q_2}^p + \left\| z\left(k, x_0, u_0^k\right) - z_f \right\|_{Q_3}^p + \left\| x\left(k, x_0, u_0^{k-1}\right) - x_f \right\|_{Q_4}^p + \left\| y\left(k, x_0, u_0^{k-1}\right) - y_f \right\|_{Q_5}^p \quad (3)$$

Subject to the MLD system and its constraints, and also to a terminal/stability condition, such as, for instance:

$$x\left(T, x_0, u_0^{T-1}\right) = x_f \quad (4)$$

$\left\| x \right\|_{Q_i}^p$ represents the l_p-norm application under vector **x**. $Q_i = Q_i^{'} \geq 0$, i = 1,2,..., 5, weight matrixes and $\mathbf{x}_f$, $\mathbf{u}_f$, δ_f, $\mathbf{z}_f$, $\mathbf{y}_f$ represent a system equilibrium point that satisfies the equations set 1. From l_p-norm definition, l_1-norm is established below, where $\mathbf{x}^{(i)}$ represents the ith entry of a given vector **x**:

$$\| x \|^p = \left(\sum_{i=1}^{n} \left| x^{(i)} \right|^p \right)^{1/p} = \sum_{i=1}^{n} \left| x^{(i)} \right| \quad (5)$$

The biggest advantage of the use of a linear performance index is that the optimization problem can be written as a linear programming problem (LP). Solving LP problems is less demanding, in a computational effort point of view, than the quadratic

programming (QP) for same size and complexity systems (Rao and Rawlings, 2000). The same can be said for problems of mixed-integer programming (MIP). A mixed-integer linear programming (MILP) problem is much less demanding than mixed-integer quadratic programming (MIQP) problems. While for simple LP and QP problems this might not be that relevant, for MILP and MIQP problems this difference is somewhat more important.

2.2. MLD optimal control as MILP

When l_1-norm is applied, the equation 3 gives the following performance index:

$$J\left(u_0^{T-1}, x_0\right) = \sum_{k=0}^{N-1} \sum_i Q_1 \left|\Delta u^{(i)}(k) - \Delta u^{(i)}{}_f\right| + \sum_i Q_2 \left|\delta^{(i)}(k) - \delta^{(i)}{}_f\right| + \sum_i Q_3 \left|z^{(i)}(k) - z^{(i)}{}_f\right| + \sum_i Q_4 \left|x^{(i)}(k) - x^{(i)}{}_f\right| + \sum_i Q_5 \left|y^{(i)}(k) - y^{(i)}{}_f\right| \tag{6}$$

where index "i" is the i[th] entry of the respective vector. It is necessary, therefore, to write the optimal control problem, enunciated in section 2.1, in the MILP problem format. Thus, auxiliary project variables are adopted to transfer the modular terms to the constraints and the rewritten performance index is:

$$J\left(u_0^{T-1}, x_0\right) = \sum_{k=0}^{T-1} \left(\sum_i \alpha_i + \sum_i \beta_i + \sum_i \eta_i + \sum_i \mu_i + \sum_i \rho_i \right) \tag{7}$$

The optimization problem solved in this work may be written as given by Eq. (8).

$$\begin{aligned} &\min \ \gamma \\ &subject\ to: -\alpha_k \le Q_1\left(\Delta u(k) - \Delta u_f\right) \le \alpha_k \\ &\qquad -\beta_k \le Q_2\left(\delta(k) - \delta_f\right) \le \beta_k \\ &\qquad -\eta_k \le Q_3\left(z(k) - z_f\right) \le \eta_k \\ &\qquad -\mu_k \le Q_4\left(x(k) - x_f\right) \le \mu_k \\ &\qquad -\rho_k \le Q_5\left(y(k) - y_f\right) \le \rho_k \\ &0 \le \sum_{k=0}^{N-1} \left(\sum_i \alpha_i + \sum_i \beta_i + \sum_i \eta_i + \sum_i \mu_i + \sum_i \rho_i \right) \le \gamma \\ &\qquad and\ MLD\ system\ equations \end{aligned} \tag{8}$$

For this predictive control problem solution, the free software Scilab 4.1.2® was used in association with LPSOLVE optimization applicative interface, (LPSOLVE 5.5 API for Scilab). LPSOLVE code congregates several linear optimization techniques, amongst them, the Branch and Bound technique, which is capable to handle with mixed-integer linear programming with easy.

It is clear that there is already software available in the literature for formulating similar problems, such as the Hybrid Toolbox (Bemporad, 2004a) that uses HYSDEL models. Nevertheless, this paper aims to use the l_1-norm out of that context, aiming to show that it is easy to formulate and viable to be implemented for small processes. For classic MPC formulation, there are important differences between LP and QP MPC

implementations, perhaps the fact that the LP problem is easier to be solved than the QP problem is not of major concern because QP algorithms are quite developed nowadays. This is not quite true for mixed integer optimization problems. Besides that, another point deserves attention, the solution of the MPC problem for different norms are indeed different, and therefore, one needs to verify which properties are mostly wanted overall. One great advantage of QP-MPC formulations is the fact that it gives linear behavior as long as the constraints are inactive. In the hybrid description, where one has a large number of constraints, the l_1-norm keeps the solution faster, and in the intersections of hyperplanes that gives optimal solutions.

3. Application and Results

For illustrating the application of such a controller, it will be used a toy example presented by Bemporad and Morari (1999):

$$\begin{aligned} x(k+1) &= 0.8\begin{bmatrix}\cos\alpha(k) & -\sin\alpha(k)\\ \sin\alpha(k) & \cos\alpha(k)\end{bmatrix}x(k)+\begin{bmatrix}0\\1\end{bmatrix}u(k),\\ y(k) &= \begin{bmatrix}1 & 0\end{bmatrix}x(k),\\ \alpha(k) &= \begin{cases}\dfrac{\pi}{3} & if \quad \begin{bmatrix}1 & 0\end{bmatrix}x(k)\geq 0\\ -\dfrac{\pi}{3} & if \quad \begin{bmatrix}1 & 0\end{bmatrix}x(k)<0\end{cases},\\ x(k) &\in [-10,10]\times[-10,10],\\ u(k) &\in [-1,1] \end{aligned} \tag{9}$$

This system can be interpreted as a piecewise linear time invariant system (PWLTI) that contains two operational modes. The activation of the modes depends upon a state event, defined by α(k). Following the same lines presented in Bemporad and Morari (1999), this system can be expressed by an equivalent MLD formulation. For the sake of brevity, this transformation is omitted. Figure 2 shows the dynamical simulation of the studied system.

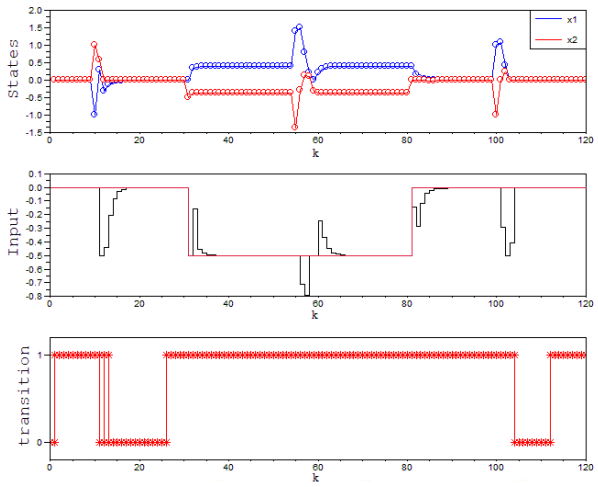


Figure 2 – Dynamical simulation: States behavior, input and its setpoint, and transition between modes.

The proposed MPC controller uses a horizon prediction H_p=3. The optimal control problem, equation set (8), presents 122 linear constraints and 53 project variables, amongst them, 6 variables are integer. The speed of control is limited in the interval: $\Delta u \in [-0,5;0,5]$. In the instants k=10, k=55 and k=100, perturbations affect the system configuring a regulatory control problem. In the instants k=30 and k=80, the setpoint is changed performing a servo control problem. The controller was also able to drive the system to its setpoint in a tracking control problem

4. Conclusions

It is known that MLD formulation allows the inclusion of logical aspects and constraints in an efficient way and that the resultant models can be used for prediction in MPC control algorithms. The selection of l_1-norm performance index makes possible the use of mixed-integer linear programming, MILP. Thus, l_1-predictive control for MLD systems can be considered a promising alternative for complex problems.

In this paper, the speed of control, Δu, is limited. This fact favors the use of l_1-norm based controllers avoiding typical problems like dead-beat and idle control. LPSOLVE 5.5 has been very efficient in MILP calculations for the simulations. Fast and stable, this free code has, basically, no limitations. The used technique for the attainment of algorithm control yields formulations that might bring problems due to the high dimensional space for optimization. However, for the current computers, this fact does not have a relevant impact for small systems and this might have an important impact when selecting or not when to use an l_1-norm or l_1-norm cost function.

References

M. Lazar, Model Predictive Control of Hybrid Systems: Stability and Robustness. Ph.D Thesis. Technische Universiteit Eindhoven, Holanda, 2006.

T. Krilavicius, Hybrid Techniques for Hybrid Systems. PhD Thesis. University of Twente, Holanda, 2003.

E. D. Sontag, Nonlinear regulation: the piecewise linear approach. IEEE Transactions on Automatic Control 26 (2), 346–357, 1981.

M. Johansson, A. Rantzer, Computation of piecewise quadratic Lyapunov functions for hybrid systems. IEEE Transactions on Automatic Control 43 (4), 555–559, 1998.

Ferrari-Trecate et al. Analysis of discrete-time piecewise affine and hybrid systems. Automatica 38 (12), 2139–2146, 2002.

A. Bemporad, M. Morari, Control of systems integrating logic, dynamics, and constraints. Automatica 35 (3), 407–427, 1999.

A. Bemporad. Efficient conversion of mixed logical dynamical systems into an equivalent piecewise affine form. IEEE Transactions on Automatic Control 49 (5), 832–838, 2004.

A. Bemporad. Hybrid Toolbox - User's Guide, availabe in http://www.dii.unisi.it/hybrid/toolbox, (2004a).

Heemels et al. Equivalence of hybrid dynamical models. Automatica 37 (7), 1085–1091, 2001.

Heemels et al. Linear complementarity systems. SIAM journal on applied mathematics 60 (4), 1234–1269, 2000.

B. de Schutter, T. J. J. Van den Boom, Model predictive control for max-plus-linear discrete event systems. Automatica 37 (7), 1049–1056, 2001.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Fuzzy Model Predictive Control Applied to Piecewise Linear Systems

Thiago V. Costa[a], Lívia M. Tizzo[b], and Luís C. Oliveira-Lopes[c]

[a]*LCAP/DESQ - Department of Chemical Systems Engineering, School of Chemical Engineering, University of Campinas, Av. Albert Einstein, 500, CP 6066, CEP13083-970, Campinas-SP, Brazil*
[b,c]*FEQUI/UFU –School of Chemical Engineering, Federal University of Uberlândia, Av. João Naves de Ávila, 2121, Santa Mônica, CEP 38408-902, Uberlândia – MG, Brazil*

Abstract

This paper presents the application of a combined control strategy applied to nonlinear systems and grounded in a MPC structure using a fuzzy model description. The nonlinear models were treated as several sub-models with linear behavior, called Piecewise Linear systems (PWL). The advantages of this methodology were visualized in the temperature control of a continuous tank reactor with output multiplicity behavior. Comparisons with classical control approaches showed that the PWL MPC is an attractive and practical strategy.

Keywords: fuzzy control, MPC, piecewise linear systems (PWL)

1. Introduction

Model Predictive Control (MPC) has called a great deal of attention in the industry, and today is the controller of choice for many areas of chemical and petrochemical industry. The great advantage of the linear model based predictive control (LMBPC) approach consists in solving a convex cost function instead of a non convex one, which has no guarantees to be solved in time. However, due to the intrinsic nonlinearity of chemical industries, obtaining good first-principles model of a process is a non-trivial task. Furthermore, an input-output data based identified model has its quality severely reduced when it is moved away from its designed operating point. An efficient alternative is the utilization of fuzzy modeling that can process numerical or language information and thus have the possibility of inclusion of qualitative information in the description of the process plant to be controlled.

Literature on fuzzy models (Takagi & Sugeno, 1985) for model predictive control covers both the utilization of fuzzy black-box models (Jang, 1993; Babuska et al., 1998) and the approximation of nonlinearities characteristics using multiple linear models. Fischer et al. (1997) and Espinosa et al. (1998) investigated the utilization of an identified fuzzy model in parallel with the MPC structure which generates step response coefficients at every sampling time. Roubos et al. (1999) showed the utilization of TS models as linear models with state dependent parameters. Huang et al. (2000) reported the approximation of nonlinear models using multiple step response convolution models; Marusak (2007), in a similar but more efficient way, designed a fuzzy model predictive controller applied to a non-minimum phase reaction system.

This paper aims to design and investigate the predictive control based on fuzzy logic models. The multiple step response approach presented by Marusak (2007) is examined

and discussed; it is known that the great drawback of this formulation is the over parameterized models and its restriction to describe only asymptotically stable plants. As an alternative for this problem, a piecewise linear state space model approach is proposed. Although less intuitive than the step response models, state space models are more informative, can easily treat multiple input – multiple output (MIMO) problems and are well suited for non-stable plants as well. The strategy and the developed controllers are illustrated in the control of a highly nonlinear continuous reactor. Results and comparisons with classical techniques of process control are addressed in the study of the MPC introduced in this paper.

2. Piecewise Linear Model Predictive Control

In the basic structure of MPC, both the model and optimizer act together in order to handle future and past inputs and outputs, cost function and reference trajectory generating the future errors and dealing with constraints. In order to compute de optimal control signal, an objective function *(1)* is minimized, subject to constraints in the input, input speed, and output.

$$J = \sum_{k=Hw}^{Hp} \left\| \hat{y}(t+k) - w(t+k) \right\|_Q^2 + \sum_{k=0}^{Hc-1} \left\| \Delta \hat{u} \right\|_R^2 \qquad (1)$$

Here H_w is the process delay, H_p is the predicted horizon, H_c is the control horizon, $\hat{y}$ is the output of the plant model, w is the reference trajectory and $\Delta\hat{u}$ is the future control increment. For a control horizon $Hc>1$, a vector of inputs are returned by the optimization, but only the first term is really sent to the process. For a detailed revision of MPC, see Maciejowski (2002).

Let us assume a nonlinear process represented by *(2),* where x is the state vector and u is the input of the process.

$$\frac{dx}{dt} = f(x,u) \quad \text{and} \quad y = g(x,u) \qquad (2)$$

a linear model could only represent the dynamic characteristics of the process in the neighborhood of a specified equilibrium point; restricting the use of the predictive controller to the boundaries of the linear model. As an alternative, the nonlinear model can be described by several linear sub-models known as piecewise linear models (PWL), which in conjunction represents the whole dynamic behavior of the nonlinear system. The PWL methodology allows the switching over the entire representation of the process (see Figure 2.1), minimizing the prediction quality losses caused by movements over the operational trajectory.

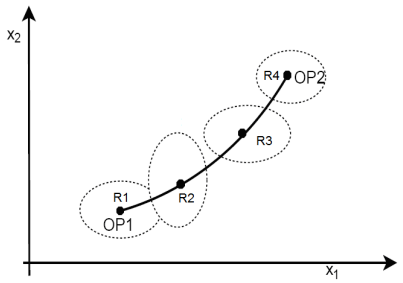


Figure 2.1 Transition of piecewise linear models.

These transitions can be adjusted by higher order TS models, with rules representing each linear region of an overall nonlinear model as an *LTV* (linear time variant) system. In this case, the estimated output is either a mean value between the two models without intersection (R3 and R4) or the weighted model response of two close operational points (R1 and R2).

2.1. Takagi-Sugeno Fuzzy Models

Takagi-Sugeno fuzzy models can be described for a multiple input single output (MISO) system as a set of $i= (1, 2, 3\ldots l)$ rules in the following form:

$$R_i\text{: IF } u_1 \text{ is } A_1 \text{ AND } u_2 \text{ is } A_2 \text{ AND} \ldots \text{AND } u_n \text{ is } A_n \text{ THEN } \hat{y}_i(k) = f(u_1, u_2, \ldots, u_n) \qquad (3)$$

where $u_1 \ldots u_n$ are inputs of the fuzzy model, $A_1 \ldots A_n$ are fuzzy sets represented by membership functions which weights a crisp input u_n in a degree of fulfilment such as $w_{jn}(u_n)$: $\mathbb{R} \rightarrow [0,1]$; $\hat{y}_i$ is the linear model of the i^{th} rule. The resultant model *(4)* is given by a *t-norm* operation of the inputs and the normalized weights *(5)* in the composition of the l rules.

$$\hat{y} = \Phi_1 y_1 + \Phi_2 y_2 + \ldots + \Phi_l y_l \qquad (4)$$

$$\Phi_i = \frac{w_i}{\sum_{j=1}^{l} w_j}; \qquad \sum_{i=1}^{l} \Phi_i = 1 \qquad (5)$$

For a piecewise step response model (*6*) with a horizon H, the weighting function Φ is applied by means of superposition principle over the step response coefficients of each linear model.

$$\hat{y}(t+k) = \sum_{m=k+1}^{H-1} S_m^* \Delta u(t+k-m) + S_T^* u(t+k-H); \quad S_m^* = \sum_{j=1}^{l} \Phi_j S_j \qquad (6)$$

Here S^* is the weighted toeplitz matrix of the step response coefficients and S_j is the matrix representing the j^{th} sub-model of the PWL or PWA (Piecewise Affine) system.

The PWA representation of a discrete state space model *(7)* is similar to the above step response model. The state matrix A^* and the input matrix B^* are also weighted by the function Φ, while A_j and B_j matrix *(8)* represents the local linearization at the j^{th} equilibrium point of a nonlinear model.

$$\begin{cases} \hat{x}(k+1) = A^* \hat{x}(k) + B^* u(k) \\ \hat{y}(k) = C^* \hat{x}(k) + D^* u(k) \end{cases} \qquad (7)$$

$$A^* = \sum_{j=1}^{l} \Phi_j A_j ; B^* = \sum_{j=1}^{l} \Phi_j B_j ; C^* = \sum_{j=1}^{l} \Phi_j C_j ; D^* = \sum_{j=1}^{l} \Phi_j D_j \qquad (8)$$

In Figure 2.2, a simple representation of the PWL Model Predictive control is placed in the sense of an internal model structure (Roubos et al, 1999). The bias correction ($y(t) - \hat{y}(t)$) is used to prevent modeling errors and in the estimation of unmeasured disturbances. The sub-models are composed by means of fuzzy reasoning to a single

model. The predicted output vector is sent to the optimizer and the quadratic convex function is solved for $\Delta u(t)$.

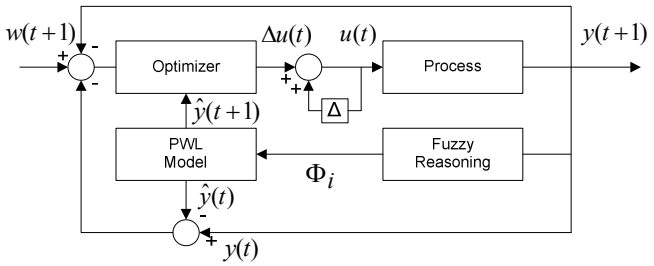
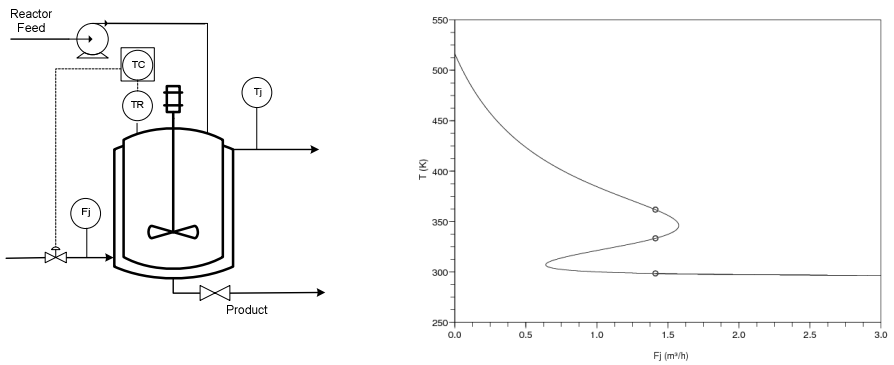


Figure 2.2 PWL Model Predictive control scheme.

3. Control Problem

A non-isothermal Continuous Stirred Tank Reactor (CSTR) was utilized as a control benchmark problem (see Figure 3.1a) for the PWL MPC. The nonlinear system presented by Luyben (1995) consists of an irreversible, exothermic reaction ($A \rightarrow B$) carried out in a perfectly mixed CSTR. The three state model is given by the reactor temperature, T (K), the reactor feed concentration, CA($kmol.m^{-3}$), and the jacket reactor temperature T_j(K). The problem consists in controlling the reactor temperature manipulating the makeup jacket flow, F_j(m^3/hr), bringing the process to a desirable product concentration, CB($kmol.m^{-3}$), value.

The motivation in using this example is due to the characteristics it presents with the output multiplicity behavior of the system. Here, a single input can lead the process to three different output values as shown in Figure 3.1b. For this system, the intermediate region is unstable in open-loop. For the open-loop case, depending on the input signal, the system can go to an ignition temperature (higher temperature region) or to an extinction temperature (lower temperature region) (Bequete, 2007).



(a) (b)

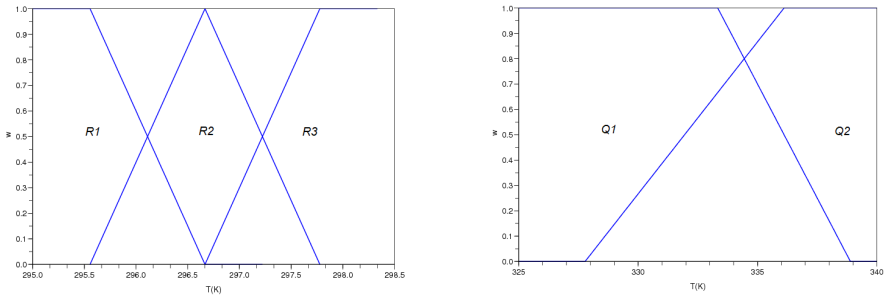
Figure 3.1 (a) Representation of the Continuous Stirred Tank Reactor (CSTR) problem; (b) Multiple steady-state behavior for the proposed control problem.

4. Results

Two PWL MPC controllers were designed for the temperature control of the proposed problem starting with $F_{jss}=1.41m^3/hr$ at a sampling time of *0.08hr*; A dynamic matrix controller (DMC) was developed for the stable low temperature region ($CA_{ss}=7.6kmol/m^3$, $T_{ss}=298.42K$) and a state space model predictive controller (SSMPC) for the intermediate unstable region ($CA_{ss}=3.92kmol/m^3$, $T_{ss}=333.34K$). For the comparisons, a standard MPC of each type and a PI (proportional-integral) controller optimally tuned for the reference trajectory were utilized.

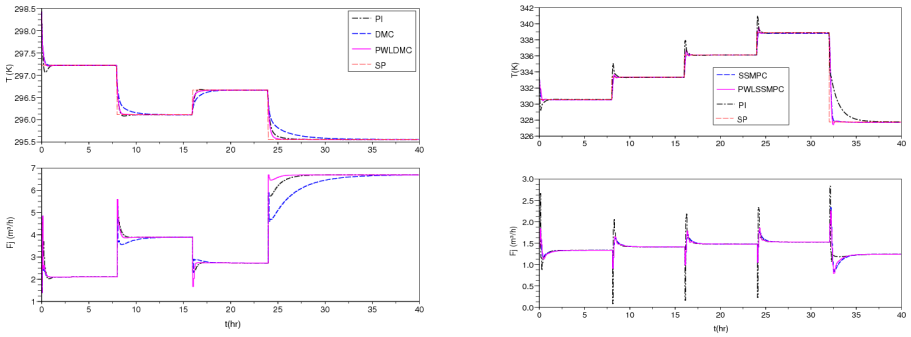
For the standard DMC algorithm, a single step response is taken in the region *R2*. The PWL DMC takes two additional regions into account covering a large area of operation given by *R1* and *R3*. For the SSMPC the model *Q1* is taken by linearization of the nonlinear model in the limits of *T=333.05K*. The PWL SSMPC takes an additional model *Q2* in the limits of the temperature *T=335.83K*. The fuzzy sets for the PWL controllers were chosen heuristically and are given in Figure 4.1.

In both simulations, the system was submitted to setpoint changes making the reactor temperature pass along the regions mentioned above. Control movements were penalized by a weighing factor *R=0.01* for both configurations of MPC. SSMPC was subject to constraints in the changing of control signal ($\Delta u \leq 0.425m^3/hr$) preventing large temperature overshoots.



(a) (b)

Figure 4.1 (a) Fuzzy sets for the PWL DMC; (b) Fuzzy sets for the PWL SSMPC.



(a) (b)

Figure 4.2 Closed-loop behavior of reactor temperature: (a) PWL DMC controller (Hc=2, Hp=4); (b) PWL SSMPC controller (Hc=10, Hp=15).

The closed-loop behavior of the system is given in Figure 4.2. Results showed that, for the lower temperature region (see Figure 4.2a), the PWL DMC had a better response due to the correction provided by the multiple sub-models approach on the final gain, leading the controller to a faster response. In Figure 4.3b the PWL SSMPC controller showed a slightly better response, the similar behavior were due to the high interpolation between the two models *Q1* and *Q2*. A small interpolation of the models was found to be unsatisfactory, due to the high nonlinearity and instability of that region. Furthermore, the great advantage of the proposed approach is the possibility to walk in the entire trajectory of the intermediate region, which cannot be done with a single model, depending on the nature of the operating point. It is shown that the PI controller is able to perform reasonably well, at the great cost of control moves and overshoot.

5. Conclusions

The PWL MPC based on a fuzzy logic description is introduced and illustrated with success in nonlinear process with output multiplicity. The overall performance of the controller depends not only on the number of local models considered, but also how the membership function is defined. Perhaps complex systems will require an optimization procedure for designing the membership region for each model. The structure presented herein is valid and can be easily adapted to any linear MPC strategy. The controller can be easily adapted to several cost function norms and can incorporate robustness constraints with easy.

References

B. Bequette, 2007, Non-Linear Model Predictive Control: A Personal Retrospective, The Canadian Journal of Chamical Engineering, 85, 408-415.

J. Espinosa & J. Vandewalle, 1998, Predictive Control Using Fuzzy Models Applied to a Steam Generating Unit, Proceedings of the Third International FLINS Workshop.

J. Jang, 1993, Anfis: Adaptive network based fuzzy inference system, IEEE Transactions on Systems, Man, and Cybernetics, 23, 665-685.

J. Maciejowski, 2002, Predictive Control with Constraints, Tokyo Denki University Press.

J. Roubos, S. Mollov, R. Babuska & H. Verbruggen, 1999, Fuzzy model-based predictive control using Takagi-Sugeno models, International Journal of Approximate Reasoning, 22, 3-30.

M. Fischer, M. Schmidt & K. Biasizzo, 1997, Nonlinear Predictive Control Based on the Extraction of Step Response Models from Takagi-Sugeno Fuzzy Systems, Proceedings of the American Control Conference, 5, 2878-2882.

P. Marusak, 2007, Advantages of an Easy to Design Fuzzy Predictive Algorithm: Application to a Nonlinear Chemical Reactor, Proceedings of the International Multiconference on Computer Science and Information Technology, 327 – 336.

R. Babuska, J. Roubos & H. B. Verbruggen, 1998, Identification of MIMO systems by input-output TS fuzzy models, IEEE World Congress on Computational Intelligence, 1, 657-662.

T. Takagi & M. Sugeno, 1985, Fuzzy identification of systems and its applications to modeling and control, IEEE transactions on systems, man, and cybernetics, 15, 116-132.

W. Luyben, 1996, Process Modeling, Simulation and control for Chemical Engineers, 2ed., McGraw-Hill, Singapore.

Y. Huang, H. Lou, J. Gong & T. Edgar, 2000, Fuzzy model predictive control, IEEE Transactions on Fuzzy Systems, 6, 665-678.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal operation of an industrial smelter furnace off-gas cleaning system

Antonio C. B. Araujo[a], Romildo P. Brito[a], Helen Shang[b]

[a]*Federal University of Campina Grande, Av. Aprigio Veloso 882, Campina Grande PB 58429-900, Brazil*
[b]*Laurentian University, 935 Ramsey Lake Road, Sudbury ON, P3E 2C6, Canada*

Abstract

Off-gas cleaning systems extract and treat hazardous emissions, and ensure that smelter operations are in accordance with environmental and industrial hygiene regulations. To this end, it is paramount that a well designed control structure be incorporated into the system. We first approach the problem by conducting a steady-state analysis based on a nonlinear model of the process, where the objective is to achieve safe, clean, and economic operation. Results reveal that this is accomplished by keeping the temperature in the furnace and in the two louvers at their upper bounds (active constraints). The unconstrained variable is found by applying the self-optimizing control technique, and the results indicate that a small loss is acceptable when one of the manipulated variables is fixed. The bottleneck of the system is identified as the fans' discharge pressures when we allow the feed rate as a degree of freedom. A control structure is then designed to keep pressures in the system well within the negative region such that acceptable dynamic performance in face of known, deterministic disturbance is achieved. Nonlinear dynamic simulations showed that very good dynamic performance is achieved with decentralized PI control schemes for both regulatory and supervisory designs.

Keywords: smelter, furnace, off-gas, optimization, control structure.

1. Introduction

Smelter furnaces provide the key operations in obtaining metal products from mineral concentrates. However, a byproduct of the process is the hazardous emission of off-gases which constitute a major source of atmospheric and plant environment pollutions. This gas is ladened with significant concentrations of CO, CO_2, and small amounts of SO_2. Continued efforts on process control and optimization is necessary to further improve the operations of smelter off-gas cleaning systems. In this work, the decision of which variables to control in an industrial furnace smelter off-gas system is addressed, and the issue is to ensure the operational objectives are achieved at the lowest possible cost, while at the same time avoiding atmospheric emissions of hazardous gases. A throughput analysis is carried out to determine where in the process the bottleneck is located, and hence where production rate should set. These decisions support the design of control configurations for the process, both at the regulatory and supervisory layers.

2. Process description

A schematic representation of the process is given in Figure 1. Hot calcines from other parts of the process, together with dusts from the off-gas cyclones and Cottrell plant are

fed to the furnace, to which coke and air are also added. Each furnace is equipped with electrodes to provide the heat required for complex reactions to take place (Celmer et al., 1987). The off-gas generated in the furnace freeboard consisting of CO, CO_2, SO_2, N_2 and O_2, first enters a cyclone to remove fine particles, which are recycled back to the furnace. It then passes a louvre for air cooling and through a fan that discharges the cooled gas into the Cottrell plant. The fans in the gas system are essential in maintaining a slightly negative pressure and suitable gas temperature in the furnace freeboards. Converters post-process the electric furnace product to remove the remainder of the iron as a slag byproduct.

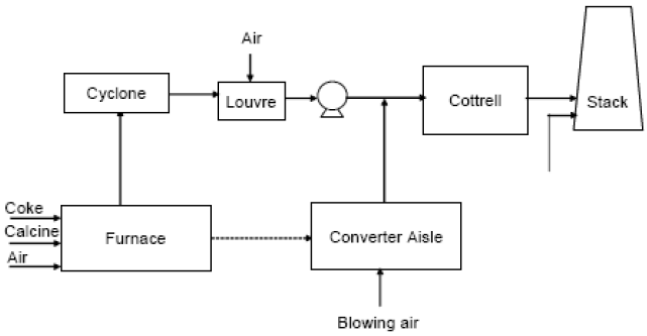


Figure 1: Schematic diagram of an industrial smelter furnace off-gas system.

3. Control structure design of the off-gas cleaning system

In this section, we design a control structure for the process based on the procedure of Skogestad (2004a), starting with a steady-state analysis.

3.1. Definition of optimal operation and degree of freedom analysis

The objective is to use as little power as possible in the two fans, subject to avoiding gas out-leakage. This is equivalent to minimizing the air in-leakage in the furnace while still preventing off-gas from finding a way out of the system. Therefore, the scaled cost function is given by (1).

$$J = pK/(\varepsilon\rho C_q)\ [(P_{fan1} - P_{lv1})W_{fan1} + (P_{fan2} - P_{lv2})W_{fan2}], \quad (1)$$

where p is the price of electricity; K is a coefficient that accounts for the compressibility of the gas, assumed constant for the range of pressures considered in this work; ε is the fan efficiency, assumed the same for all fans; ρ is the gas density, assumed constant for the range of pressures considered in this work; C_q is a constant that adjusts the dimension of the units used; the other variables are defined below. The constraints to operation are listed in Table 1. The process has 5 manipulated variables - coke added to the furnace (W_{coke}), louver #1 and #2 vane openings (O_1 and O_2), fan #1 and #2 rotation speeds (N_{fan1} and N_{fan2}) - and therefore 5 dynamic degrees of freedom. We also consider 16 candidate measurements: furnace freeboard pressure (P_f) and temperature (T_f); cyclone #1 and #2 pressures (P_{c1} and P_{c2}); louver #1 and #2 pressures (P_{lv1} and P_{lv2}); louver #1 and #2 temperatures (T_{lv1} and T_{lv2}); Cottrell inlet pressure (P_{iCt}) and temperature (T_{iCt}); Fan #1 and #2 outlet pressures (P_{fan1} and P_{fan2}); louver #1 and #2 vane openings (O_1 and O_2); and fan #1 and #2 rotation speeds (N_{fan1} and N_{fan2}). The disturbances in Table 2 include the effects of feed flow rate on the process and changes in the system's structure as represented by the effective in-leakage area in the furnace.

Table 1: Steady-state constraints to the process.

	Constraint description	Bound
P_f	Furnace freeboard pressure [Pa]	< -25
P_{c1} and P_{c2}	Cyclone #1 and #2 pressures [Pa]	< 0
P_{lv1} and P_{lv2}	Louver #1 and # 2 pressures [Pa]	< 0
P_{fan1} and P_{fan2}	Fan #1 and # 2 outlet pressures [Pa]	< 0
P_{iCt}	Cottrell inlet pressure [Pa]	< 0
T_f	Furnace freeboard temperature [K]	< 923.15
T_{lv1} and T_{lv2}	Louver #1 and #2 temperatures [K]	< 643.15
T_{iCt}	Cottrell inlet temperature [K]	< 643.15
O_1 and O_2	Louver #1 and #2 vane openings [%]	[0 100]
N_{fan1} and N_{fan1}	Fan #1 and #2 rotation speeds [rpm]	< 1500

3.2. Primary controlled variable selection

The starting point for the selection of primary (economic) variables is the optimization of the process for the disturbances in Table 2. The model developed by Shang et al. (2008) is implemented and optimizations with fixed W_{coke} are performed. Figure 2 shows the effect of the disturbances on the cost. 3 constraints are always active, namely, T_f (upper bound), T_{lv1} (upper bound), and T_{lv2} (upper bound). As these 3 active constraints must be implemented to ensure optimal operation (Maarleveld and Rijnsdorp, 1970), we are left with 1 degree of freedom.

Table 2: Disturbances to the process.

Disturbance description	Nominal	Disturbance (Δ)
Coke added to the furnace [kg/s]	2	+0.4 (D1) and -0.2 (D2)
Equivalent CO temperature in the furnace [K]	1573.15	+200 (D3) and -200 (D4)
Flow rate from converters [kg/s]	300	+60 (D5) and -60 (D6)
Converter outlet temperature [K]	473.15	+100 (D7) and -100 (D8)
Effective in-leakage area in the furnace [m2]	1.50	+0.75 (D9)
Room temperature [K]	273	-30 (D10) and +30 (D11)

It can be readily pointed out from Figure 2 that the disturbances which have the largest influence on the cost are those that affect the mass balance, i.e., D1 and D2. By increasing the flow rate from the converters the cost is reduced, and vice-versa. This is due to the reduction in the head from the louver to the fan. When the equivalent CO temperature in the furnace increases (D3 and D4), pressure in the furnace freeboard is reduced allowing more air into the system to cool down the off-gas, and vice-versa. The converter temperature (D7 and D8) has no economical effect on the system as it has no influence in upstream units. Disturbance D9 shows that it is “cheaper” to operate the furnace at higher pressures but within a safety margin. The seasonal change in temperature (D10 and D11) affects the economics of the system by allowing more or less cooling load.

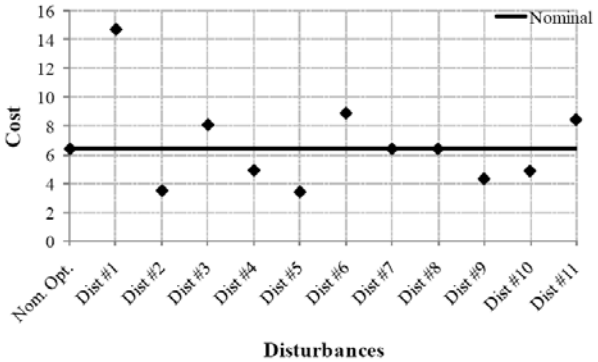


Figure 2: Effect of disturbances on optimal operation.

We now need to find a suitable unconstrained controlled variable, and to this end we use the self-optimizing control technology (Skogestad, 2000) which consists of finding a variable c that when kept at its nominal optimal setpoint minimizes the loss $L = J(c,d) - J_{opt}(d)$, where $J(c,d)$ is the actual value of the cost function for a given c and $J_{opt}(d)$ is the truly optimal value if the process was re-optimized for each disturbance d. The results for the various runs showed that a good candidate would be either N_{fan1} or N_{fan2}.

3.3. Production rate

The maximum throughput analysis determines the evolution of the static behavior of the system as W_{coke} is increased. It is usually optimal to increase the production rate as much as possible because the prices may be such that the actual profit of the enterprise would increase with W_{coke}. To find $W_{coke,max}$ (Mode II), and hence the bottleneck of the process, we use W_{coke} as a degree of freedom and re-optimize the process, using the profit function J' as given in (2), where p_f is the price of the feed.

$$J' = pK/(\varepsilon\rho C_q)\ [(P_{fan1} - P_{lv1})W_{fan1} + (P_{fan2} - P_{lv2})W_{fan2}] - p_f W_{coke} \qquad (2)$$

The cost as a function of W_{coke} is shown graphically in Figure 3. The active constraints were found to be the same as per Mode I, i.e., T_f, T_{lv1}, and T_{lv2} are active at their upper bounds. Operation becomes infeasible when W_{coke} = 3.15 kg/s since both P_{fan1} and P_{fan2} become active at their respective upper bounds; consequently, production cannot be further increased, and the bottleneck of the process is either P_{fan1} or P_{fan2}. However, when disturbances are considered $W_{coke,max}$ will change accordingly. Assuming the plant is subject to disturbances D3 to D11 in Table 3, the maximum throughput results show that the converters largely affect the steady-state economic operation of the off-gas system, and for these disturbances, the throughput varies largely from W_{coke} = 2.52 to 3.78 kg/s.

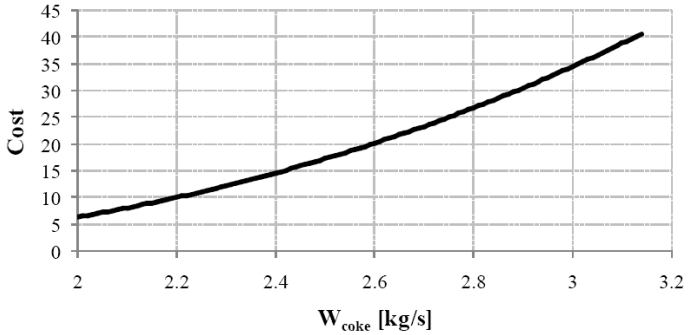


Figure 3: Optimization of the smelter furnace system with variable feed rate.

4. Bottom-up design of the cleaning off-gas system

The main objective of the regulatory layer is to provide sufficient control quality to enable a trained operator to keep the plant running safely without the need for using higher layers in the control system. In addition, the control task at this layer is to prevent the plant from drifting away from its desired operating point on the short time scale. As the smelter furnace off-gas system has no unstable mathematical modes, no stabilization of this kind is indeed necessary. However, as pressure dynamics are generally very fast, drift in these variables due to disturbances is avoided by controlling pressure at selected locations in the plant. In Mode I, P_f is selected to be controlled using either N_{fan1} or N_{fan2}, this minimizes the impact of disturbances in the primary controlled temperatures because of the direct relation between these two variables. As for Mode II, P_f is also controlled, and we use W_{coke} as manipulated variable; in addition, we control P_{fan1} and P_{fan2} using N_{fan1} and N_{fan2}, respectively. The intended aim of the supervisory control layer is to keep the active constraints and unconstrained controlled variables at their constant optimum setpoints. For the unconstrained variable in Mode I, N_{fan1} is fixed at its nominal optimum set point. For Mode II, T_f is controlled using as manipulated variable the setpoint of the furnace pressure controller, $P_{f,sp}$. These two strategies favor optimal economic performance with minimal losses. Moreover, for both modes of operation T_{lv1} and T_{lv2} are controlled using O_1 and O_2, respectively. The production rate manipulator is selected to be W_{coke}, where it is fixed in Mode I and adjusted to give the desired maximum throughput in Mode II. The resulting PI controllers are tuned using the SIMC tuning rules described in Skogestad (2004b). Note that optimal operation in both modes always lies at process constraints. However, during transients these constraints may eventually be violated for very short periods of time when using conventional control configurations, since PID controllers do not account for constraints in their algorithms. Nonetheless, the variables associated to these constraints must always return to their bounds irrespective of disturbances, and this can always be achieved by adding integral action to the controllers. One way to minimize these dynamic violations is to resort to multivariable controllers which explicitly make up for constraints in their formulations like, for instance, MPC controllers.

5. Validation of the proposed control structure

We here perform dynamic simulations to validate the proposed control structures for the smelter furnace off-gas system when facing fast changing disturbances. For the sake of brevity, only simulations involving Mode I with disturbances D1 and D2 are considered. Figure 4 illustrates the responses of the system where it can be seen that the proposed control structure is effective in rejecting disturbances, showing quick dynamic response, with pressures strictly within bounds.

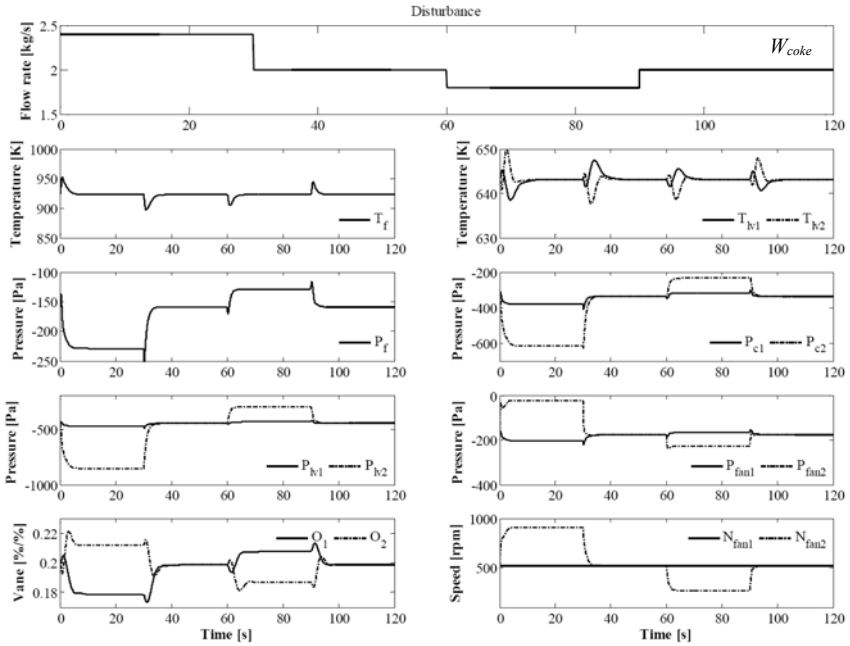


Figure 4: Responses for disturbances in W_{coke} (top of the figure) in Mode I.

The responses for Mode II are not shown but they yield good dynamic performance, with pressures well within their negative ranges.

6. Conclusion

The use of a systematic procedure to the design of a simple decentralized control structure shows it is possible to operate an industrial smelter furnace off-gas cleaning system in a (near) optimal economic fashion that entirely complies with environmental regulations by avoiding out-leakage of hazardous off-gases to the atmosphere. The resulting dynamic performances for both modes of operation ratify the topnotch efficiency of the proposed control configurations when facing large, fast disturbances affecting the process.

References

Celmer, R. S., Kaiura, G. H., Toguri J. M., 1987, Chemical reactions during the electric smelting of nickel-copper calcines, Canadian metallurgical Quarterly, 26(4), 277-284.

Maarleveld, A., Rijnsdorp, J. E., 1970, Constraint control on distillation columns. Automatica, 6, 51-58.

Shang, H., Megan, D., Nelson, P., Salt, B, 2008, Dynamic modelling of an industrial smelter furnace and converter off-gas system, American Journal of Environmental Sciences, 40(1), 22-30.

Skogestad, S., 2000, Plantwide Control: The search for the self-optimizing control structure, Journal of Process Control, 10, 487-507.

Skogestad, S., 2004a, Control structure design for complete chemical plants, Computers and Chemical Engineering, 28, 219-234.

Skogestad, S., 2004b, Simple analytic rules for model reduction and PID controller tuning, Modeling, Identification and Control, 25(2), 85-120.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal operation of an industrial PVC dryer

Antonio C. B. Araujo[a], Jose J. N. Neto[a], Helen Shang[b]

[a]*Federal University of Campina Grande, Av. Aprigio Veloso 882, Campina Grande PB 58429-900, Brazil*
[b]*Laurentian University, 935 Ramsey Lake Road, Sudbury ON, P3E 2C6, Canada*

Abstract

In this work, we devise a control structure architecture for an industrial polyvinyl chloride (PVC) dryer currently operating at Braskem Company, Brazil. The motivation is the optimally economic operation of the process as well as the prompt rejection of disturbances at lower layers in the control hierarchy. Optimization of a simplified model of the process for various important disturbances indicates that it is optimal to control the temperature level of the utilities serving the dryer and the final PVC moisture contents at their respective upper bounds. In addition, acceptable economic loss is achieved by keeping the flow of air to the dryer at a fixed optimum setpoint. Almost perfect indirect control of the final PVC moisture is achieved by tightly controlling a temperature difference in the dryer. The resulting decentralized control configuration leads to good dynamic performance for important disturbances affecting the system.

Keywords: drying, PVC, optimization, control structure.

1. Introduction

Fluidized bed dryers are widely used in both continuous and batch processes by the food, pharmaceutical, agricultural, and chemical industries. However, systematic design of control structures is still lacking. This leads to the fact that literature on control of dryers is surprisingly scarce, in addition to being based on heuristics arguments. We here name only a few results found in most relevant papers. Mujumdar (2006) described control strategies for fluidized bed dryers, and quotes that the most adequate is when the exhaust-air temperature is controlled by the inlet air temperature. Harbert (1974) and Alden *et al.* (1988) studied a technique where the temperature difference between the material being dried and the wet-bulb temperature is the selected controlled variable since it infers the end-point of drying. Abdel-Jabbar *et al.* (2005), on the other hand, proposed to control the final moisture content indirectly estimated via a linearly designed Kalman filter. Panda and Rao (1994) derived and compared the performance of internal model (IMC) and dynamic matrix (DMC) controllers in an experimental apparatus for fluidized-bed drying of sand. Most of the papers refer to vertical pneumatic fluidized bed dryers in which the sole heating media is the fluidizing air. In this paper, a different arrangement is considered for the drying of PVC where heating elements are inserted in the bed aiming at reducing capital costs.

2. Description of the industrial unit

Figure 1 depicts a schematic of the PVC drying operation. After polymerization, a mixture containing 30% (wt/wt) of PVC and 70% (wt/wt) water is sent to centrifuges to mechanically reduce the moisture of the slurry to 20-30% wt/wt water content. The resin is then fed into a fluidized bed dryer to remove the residual water. The energy

required by the drying process is supplied by hot air blown to the bottom of the dryer as well as by hot water circulating inside the bed through a series of heat exchangers.

3. Modeling of an industrial PVC dryer

We base the model developed in this paper on the work by Levi and Borde (1999) and Groenewold and Tsotsas (2007) due to its simplicity and accuracy against plant data. The process is described by a two-phase model. Considering a single cell *j*, we assume that the gas phase is forced vertically to flow through the fluidized particles in a plug-flow pattern. We then divide the gas path into *M* sections along the cell *j* where each section *i* is modeled as a CSTR. In other words, the model is distributed for the gas phase and lumped for the particle phase in a given cell.

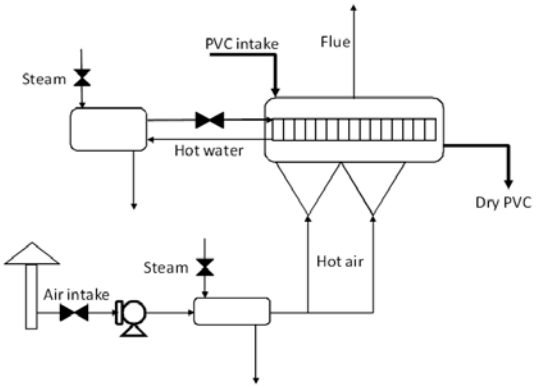


Figure 1: Schematic of the continuous PVC fluidized bed drying process.

A comparison between the proposed model and plant data was performed and the results showed very good agreement.

4. Control structure design of the PVC dryer

In this section, we design a control structure for the process based on the procedure of Skogestad (2004a), starting with a steady-state analysis.

4.1. Definition of optimal operation and degree of freedom analysis

The objective is to reduce the cost of energy as much as possible, which is equivalent to minimizing the amount of steam used to heat up the air and hot water, while still delivering the final PVC at the right moisture content specification. Therefore, the cost function *J* to be minimized is given by (1).

$$J = \frac{8000}{D_{scl}} \left(F_G \int_{T_{room}}^{T_{G,in}} C_{PG} dT + m_W \rho_W \int_{T_{W,out}}^{T_{W,in}} C_{PL} dT \right) \qquad (1)$$

where F_G is the air flow rate; T_{room} is the room temperature; $T_{G,in}$ is the inlet air temperature; $T_{W,in}$ is the inlet hot water temperature; $T_{W,out}$ is the outlet hot water temperature; m_W is the hot water flow rate; ρ_W is the liquid water density; $C_{PG} = C_{PW} + Y_G C_{PA}$ is the gas phase specific heat (C_{PW} and C_{PA} are the specific heat of water vapor and the specific heat of air, respectively for which equations were found experimentally from plant experiments); and C_{PL} is the liquid water specific heat. We here consider operation over a period of 8000 h/year. For optimization purposes the cost as given by (1) is properly scaled as given by the scaling factor D_{scl}. The process is subject to the constraints in Table 1. These constraints reflect two unmodeled features of the process:

The lower bound on the final particle moisture content, $X_{S,out}$, prevents electrostatic charge formation, and the bounds on the air flux ($f_{G,j}$) in each cell j refer to maximum and minimum fluidization velocities through the maximum ($f^{max}_{G,j}$) and minimum ($f^{min}_{G,j}$) air fluxes, for which equations are found in Howard (1993). The process has 5 manipulated variables - PVC feed rate in dry basis (F_S), inlet air flow rate in dry basis (F_G), hot water flow rate (m_W), inlet air temperature ($T_{G,in}$), and inlet hot water temperature ($T_{W,in}$) - and therefore 5 dynamic degrees of freedom. We also consider 10 candidate measurements: gas temperature at the bottom of cell #1 ($T_{G1,1}$), #2 ($T_{G1,3}$), and #3 ($T_{G1,4}$) and respective differences $T_{G1,1}$ - $T_{G1,3}$; $T_{G1,1}$ - $T_{G1,4}$; and $T_{G1,3}$ - $T_{G1,4}$; outlet gas temperature ($T_{G,out}$); outlet hot water temperature ($T_{W,out}$); inlet air flow rate (F_G); and hot water flow rate (m_W). Disturbances in Table 2 represent changes in the inlet PVC.

Table 1: Constraints to the process.

Constraint description	Bound
Final particle moisture content in dry basis [kg/kg]	$0.0018 \le X_{S,out} \le 0.0020$
Air flux in cell j [kg/(m^2 s)]	$f^{min}_{G,j} \le f_{G,j} \le f^{max}_{G,j}$
Inlet air temperature [K]	$303.15 \le T_{G,in} \le 383.15$
Inlet hot water temperature [K]	$303.15 \le T_{W,in} \le 383.15$
Inlet air flow rate in dry basis [kg/h]	$22370 \le F_G \le 78300$
Hot water flow rate [m^3/h]	$0 \le m_W \le 300$

4.2. Primary controlled variable selection

The starting point for the selection is the optimization of the process for the disturbances in Table 2. The proposed model is implemented and optimizations with fixed F_S are performed. The results show that 3 constraints are always active at their upper bounds, namely, $X_{S,out}$, $T_{G,in}$, and $T_{W,in}$. $X_{S,out}$ is active since it is always optimal to operate the plant at the minimum specification. $T_{G,in}$ and $T_{W,in}$ are active because drying processes are basically favored by high energy transfer rates. Therefore, what is left for the optimization is to minimize the flow rate of utilities to trade off cost reduction with feasibility, while satisfying the specification on $X_{S,out}$. As these 3 active constraints must be implemented to ensure optimal operation (Maarleveld and Rijnsdorp, 1970), we are left with 1 degree of freedom.

Table 2: Disturbances to the process.

Disturbance description	Nominal	Disturbance (Δ)
Particulate phase mass flow rate in dry basis [kg/h]	7370	+1630 (D1) and -2370 (D2)
Initial particle moisture content in dry basis [kg/kg]	0.233	+0.282 (D3) and -0.043 (D4)
Initial particulate phase temperature [K]	335	+23 (D5) and -2 (D6)

It can be seen from Figure 2 that the disturbances which most influence the cost is when $X_{S,in}$ is increased because it is rather difficult to remove the moisture inside the particle since more air would be necessary, and this increase in air flow rate will eventually hit the upper bound on f_G. Therefore, hot water is used to provide the complementary energy requirements, which is a more expensive heat source.

We now need to find a suitable unconstrained controlled variable, and to this end we use the self-optimizing control technology (Skogestad, 2000) which consists of finding a variable c that when kept at its nominal optimal setpoint minimizes the loss $L = J(c,d) - J_{opt}(d)$, where $J(c,d)$ is the actual value of the cost function for a given c and $J_{opt}(d)$ is the truly optimal value if the process was re-optimized for each disturbance d.

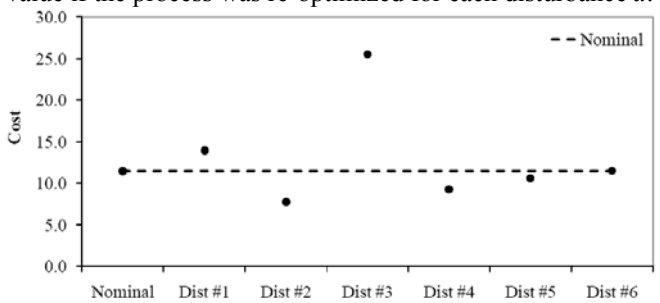


Figure 2: Effect of disturbances on optimal operation.

The results for the various runs show that operation is infeasible for any candidate controlled variable. We then "back off" the variables' optimal setpoints to resolve this problem until finding a suitable value for the selected candidate variables such that operation is always feasible, at the expense of a loss penalty. Two very attractive candidate variables are F_G and m_W since it is always a good practice to operate with minimal manipulation handling. The only value for F_G such that operation is still "cheap" would be the one calculated for disturbance D3, i.e., F_G = 26300 kg/h, because anything smaller would result in infeasible operation since m_W is at its maximum. m_W can only assume values such that $m_W \leq 5.32$ m^3/h, otherwise operation would become infeasible for D2 as F_G is already at its lower bound. The seemingly optimal value would then be m_W = 5.32 m3/h; however, under these conditions it is certain that operation for D3 will be infeasible since with such a small flow of hot water, either the upper bound on f_G or F_G would be violated. Therefore, F_G is chosen as the unconstrained variable.

Since $X_{S,out}$ is difficult to measure, it is necessary to find another suitable active constraint. The idea is to determine the value of a selected variable such that, when kept constant, minimizes the deviation in $X_{S,out}$ from its optimal value $X^{opt}_{S,out}$ when different disturbances are considered. The results of the optimizations show that only the candidate $T_{G1,3}$ - $T_{G1,4}$ gives feasible operation for the entire set of disturbances. Hence, it is select as the new active constraint.

4.3. Production rate

The maximum throughput analysis determines the evolution of the static behavior of the system as F_S is increased. It is usually optimal to increase the production rate as much as possible because the prices may be such that the actual profit increases with F_S. To find $F_{S,max}$ (Mode II), and hence the bottleneck of the process, we use F_S as a degree of freedom and re-optimize the process, using the profit function J as given in (1) (see Figure 3). Note that the active constraints are the same as per Mode I. When F_S = 12370 kg/h, f_G reaches its upper bound, and above F_S = 20810 kg/h operation is no longer feasible because m_W reaches its upper limit; therefore, m_W is the bottleneck of the process. However, when disturbances are considered $F_{S,max}$ reduces accordingly. By considering the worst case scenario, in which $X_{S,in}$ is increased by 10%, the ultimate throughput is found at $F_{S,max}$ = 18400 kg/h.

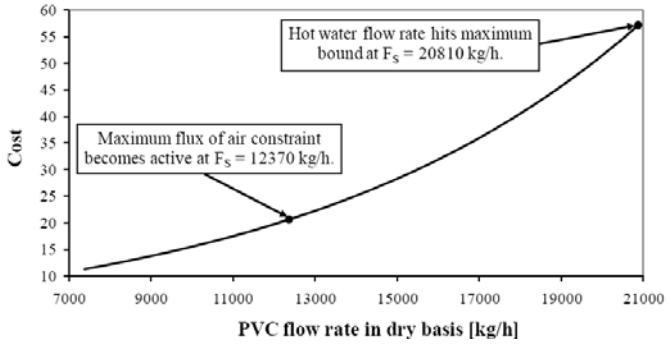


Figure 3: Optimization of the smelter furnace system with variable feed rate.

5. Regulatory and supervisory control structures design

The main objective of the regulatory layer is to provide sufficient control quality to enable a trained operator to keep the plant running safely and to prevent the plant from drifting away from its desired operating point. As the PVC dryer has no unstable mathematical modes, no stabilization of this kind is necessary. To reduce drift, the air blower vane downstream the unit controls pressure, and $T_{G,in}$ and $T_{W,in}$ are kept constant. The intended aim of the supervisory control layer is to keep the active constraints and unconstrained variables at their optimum setpoints. For this, $T_{G1,3}$ - $T_{G1,4}$ is used to indirectly control $X_{S,out}$. The resulting PI controllers are then properly tuned using the SIMC tuning rules described in Skogestad (2004b).

6. Validation of the proposed control structure

For Mode I of operation, we consider an increase of 0.1 kg/kg in $X_{S,in}$. As for Mode II, the disturbance consists of a 10 % step increase in $X_{S,in}$. Figures 4 and 5 illustrates the responses of the system, where it can be seen that disturbances can be quite effectively rejected with few oscillations and overshoots, as well as small offset in $X_{S,out}$.

7. Conclusion

The general guidelines to optimally operate a PVC dryer of the kind discussed in this work can be summarized as: operate the dryer at the maximum temperature level, i.e., with the temperature of the inlet air and hot water at their upper limits, and adjust the air flow rate to minimize the flow rate of hot water, while keeping the final PVC moisture content as close as possible to the upper specification. Having said that, the resulting control structure successfully fulfills the task of delivering PVC at the right moisture content while operating the unit at the lowest possible cost.

References

Abdel-Jabbar, M. N., Ali Jumah, R. Y., and Al-Haj Ali, M. Q., 2005, State estimation and state feedback control for continuous fluidized bed dryers, Journal of Food Engineering, 70, 197–203.

Alden, M., Torkington, P. and Strutt, A. C. R., 1988, Control and instrumentation of a fluidized-bed drier using the temperature-difference technique. I. Development of a working model, Powder Technology, 54, 15–25.

Groenewold, H. and E. Tsotsas, 2007, Drying in fluidized beds with immersed heating elements, Chemical Engineering Science, 62, 481–502.

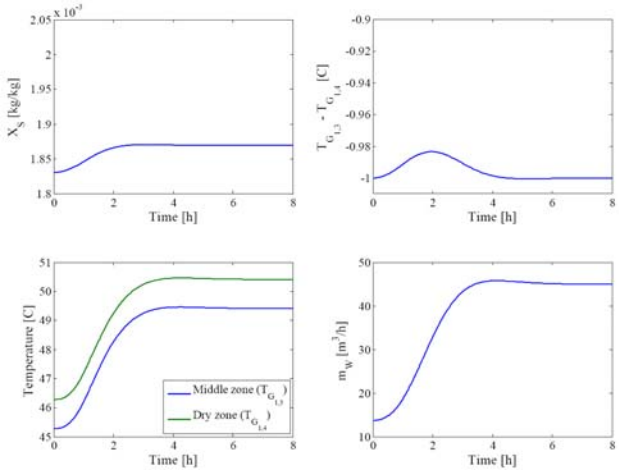


Figure 4: Responses for a disturbance in $X_{S,in}$ in Mode I.

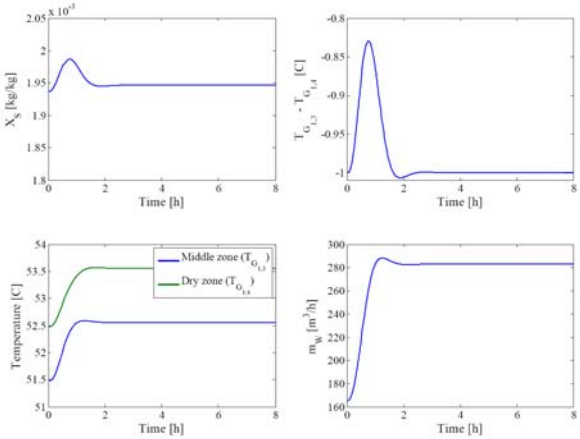


Figure 5: Responses for a disturbance in $X_{S,in}$ in Mode II.

Harbert, F. C., 1974, Automatic control of drying processes moisture measurement and control by the temperature difference method. Chemical Engineering Science, 29,888–890.

Howard, J. R., 1993, Fluidized Bed Technology: Principles and Applications, Adam Hilger.

Levi, A. and I. Borde, 1999, Steady state one dimensional flow model for a pneumatic dryer, Chemical Engineering Processing, 38(2), 121–130.

Maarleveld, A., Rijnsdorp, J. E., 1970, Constraint control on distillation columns. Automatica, 6, 51-58.

Mujumdar, A. S., 2006, Handbook of Industrial Drying, 3rd ed., CRC.

Panda, R. C. and Rao, V. S. R., 1994. Model based control of a continuous fluidized bed dryer. In Proceedings of the third IEEE conference on control applications, 507–511, Glasgow, Scotland.

Skogestad, S., 2000, Plantwide Control: The search for the self-optimizing control structure, Journal of Process Control, 10, 487-507.

Skogestad, S., 2004a, Control structure design for complete chemical plants, Computers and Chemical Engineering, 28, 219-234.

Skogestad, S., 2004b, Simple analytic rules for model reduction and PID controller tuning, Modeling, Identification and Control, 25(2), 85-120.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal Scheduling of a Multiproduct Continuous Paper Plant

Pedro P. Mariano,[a,b] Pedro M. Castro,[b] Ana P.F.D. Barbosa-Póvoa[a]

[a]*CEG-IST, Universidade Técnica de Lisboa, 1049-001 Lisboa, Portugal*
[b]*DMS, Inst. Nacional Engenharia Tecnologia e Inovação, 1649-038 Lisboa, Portugal*

Abstract

This paper addresses the solution of a real-life problem from a multiproduct paper plant, involving the scheduling of three machines in parallel for a time horizon of roughly one month. A decomposition strategy is proposed to reduce the complexity where orders are aggregated according to the product, density, size and due date. Furthermore, they are allocated to a particular machine, enabling the solution of three separate sequencing problems instead of one very large scheduling problem. Three alternative formulations from the literature are evaluated, with the results showing that it is better to use a unit-specific continuous-time formulation with 3-index binary variables. This is then used to find a better solution to the problem by releasing the order-unit assignments for a subset of the orders.

Keywords: Continuous-time, decomposition strategy, parallel machines, changeovers.

1. Introduction

Scheduling is concerned with the allocation of resources over time so has to execute the processing tasks required to manufacture a given set of products. Due to its critical role, this type of activity plays an important part in any organization. In particular, there is the case of industries of the production/transformation sectors with considerable amounts of inventory where an optimization of the associated schedule is required. Pulp and paper is a perfect example, with the Setubal paper plant of Portucel Soporcel, located in Portugal, providing the subject of this paper.

The plant's Production Department plans the budget for production and transformation, which includes cut, cut size and rewind, for a time horizon of one year. This plan is strongly influenced by the market behavior and the data acquired from the clients. Considering this plan and the orders received, the Sales & Marketing Department conceives the schedule for each month. A generic 30-day schedule (each with 3 daily shifts) features 22 changeovers between different products, distributed over the 3 paper machines in the following way: 12 in MP1, 5 in MP2, and 5 in MP3.

This paper aims to optimize the schedule of the plant by minimizing the total production time and the number of changeovers.

2. State-of-the-art

Scheduling has received considerable attention by the scientific community in the last twenty years, in large part due to the vital role it plays in the operations management area. Concerning time representation, two main types of formulations can be identified. Discrete time formulations (DT), where the time horizon is divided into a number of

intervals of equal duration, and continuous time formulations (CT) where the duration of the time intervals is going to be determined by the optimization solver.
The DT models, firstly based on a State-Task (STN) and later on a Resource-Task Network representation of process and/or production recipe, appeared earlier than their continuous-time counterparts. However, the latter have received most of the attention in the last decade (Méndez et al., 2006). Nowadays, researchers are mostly focusing on developing unit-specific approaches (Castro et al. 2006, Liu and Karimi, 2007, Shaik and Floudas, 2008, Castro and Novais, 2008) that can be as general as those based on global event points, since the former can be significantly more efficient computationally. In particular, multistage plants are an example of success. Overall, it is clear that there is not a single best approach for all problem types and researchers are still drawing the map that will tell which model to use as a function of the characteristics of the real problem. This paper evaluates the performance of two models (CT3I and CT4I) proposed by Castro et al. (2006) on a real-life industrial case.

3. Problem definition

In this paper, the short-term scheduling problem of a single stage, multiproduct continuous plant is considered. This is formed by a set of three machines ($m \in M$) that have to process within a month, 420 orders ($i \in I$) of 12 different products (P1-P12). Given are also the duration of the processing tasks and cleaning tasks, the due dates and the amounts to process. The objective is to minimize the total production time in all machines, which maximizes productivity by reducing the number of changeovers, each lasting between 10 to 30 minutes. It is required to meet (hard constraints) the due dates for all products as well as the maximum availability for storage in the warehouse.

4. Solution strategy

The original number of orders (420) is well above the current capabilities of state-of-the-art scheduling models, so a two-stage solution method (figure 1) was applied to reduce the complexity.
In the first stage, orders belonging to the same product and sharing the density (g/m^2), size and due date, were placed in the same group. This reduced the number of orders to 191, still too high to start to solve the current problem. Thus another aggregation step was needed, which consisted on grouping orders with total duration as close as possible to a day of production. The outcome was the reduction of the number of orders to 73, an acceptable value.
In the second stage, we assume a fixed product-unit assignment, which is stated by the Production and Sales & Marketing departments. In this way, orders belonging to P1-P6 are processed in MP1; orders belonging to P7-P9 in MP2; and those belonging to P10-P12 in MP3.
Overall, a trade-off is established between flexibility (more orders to process) and computer effort (less orders). Note that order aggregation and fixed order-units assignments will almost certainly remove the optimal solution from the feasible space. More importantly, the scheduling problem has been replaced by sequencing problems, which will be solved either simultaneously or separately.

5. Continuous-time model selection

The number in CT3/4I indicates the indices in the binary variables. Formulation CT4I identifies the execution of order *i* followed by the required changeover for order *i'* to immediately follow in unit *m* at event point *t*. This approach allows to model

changeovers explicitly. CT3I, on the other hand, identifies the execution of order *i* in unit *m* at time *t*, and handles changeovers implicitly through the model constraints. This is the main conceptual difference between the two formulations, which is reflected in the size of the mathematical problem by roughly one order of magnitude in the number of variables.

To evaluate which formulation is the most efficient for the present problem, a simplified version of the industrial problem was solved. It featured a half-month time horizon and 36 orders, corresponding to about 8500 tons of paper. One important decision concerns the number of event points to consider on each time grid, one per machine. Usually, an iterative procedure must be used to find the minimum value that ensures optimality but in this case there is no such problem since the order-unit assignment is fixed. Thus, the number of event points to use is equal to the number of orders assigned to the machine plus one. Furthermore, each machine can be scheduled separately since the problems are in fact independent. Nevertheless, we will be solving all machines simultaneously as well as separately. The objective function was the (sum of the) total production time and the results are listed in Table 1.

Formulation CT4I exhibited a better computational performance, which is consistent with past studies (Castro et al. 2008). We then went on to address the full set of orders (73, corresponding to 16752 tons of paper), only to find out that CT4I failed to reach even a feasible solution. In contrast, CT3I successfully found a very good solution (0.3% relative gap) worth 57.60 days in about 4 hours. The same solution was returned after scheduling the machines separately, in a total of 5576 s (3562, 4.0 and 2010 in MP1, MP2 and MP3, for relative gaps of 0.5, 0 and 0.6%, respectively). Note that the computational statistics given are for the largest instance solved (MP1), the one with the most orders. CT4I got just to 57.77 days in significantly more time, a total of 17292 s. Thus, CT3I was preferred over CT4I.

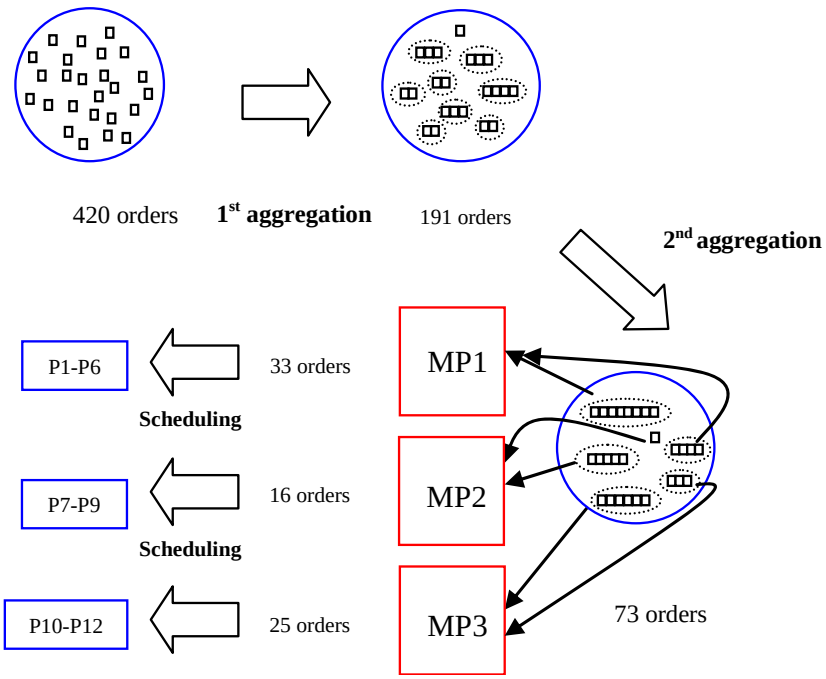


Figure 1. Two-stage solution strategy with order aggregation and fixed product-unit assignments

Table 1. Computational results for comparison CT3I vs. CT4I (CPLEX 10.2)

Model	# Orders	Machines	DV[a]	SV[b]	EQ[c]	OBJ (days)	CPUs
CT3I	36	Simultan.	1551	1595	1664	28.86	5042
CT4I	36	Simultan.	19212	20805	3085	28.86	3599
CT3I	73	Simultan.	2485	2567	5954	57.60	14800[d]
CT4I	73	Simultan.	59183	65032	6106	-	-
CT3I	73	Separately	1123	1158	1287	57.60	5576[d]
CT4I	73	Separately	34881	36071	1355	57.77	17292[d]

[a]Discrete variables; [b]Single variables; [c]Equations; [d]Maximum resource limit or computational time to reach relative optimality gap<1%

6. Comparison with a discrete-time approach

The schedule can also be derived with a discrete-time approach. One key difference is that in the second stage of the aggregation procedure described in section 4, the duration is now made as close as possible to a multiple of the interval length. The actual value, 0.25 days, results from a trade-off between total computational effort and data accuracy. Whenever the orders have combined duration lower than 0.25 days, more paper is produced to avoid the appearance of idle times, which is not tolerable in a continuous paper plant where some kind of paper is always being produced (excluding periods of maintenance. As a consequence, schedules from the discrete-time will produce 17.2% more paper (19428 vs. 16572 tons). The group of orders from the second aggregation procedure was also higher 124 vs. 73, which brings a natural increase in problem size.
The total production time was equal to 67 days, 16.3% higher than for continuous-time. All orders were produced before their due date and the maximum amount of paper in inventory was kept below the maximum capacity of the warehouse (6500 tons) even though no such constraint was in the model. In terms of the total number of changeovers, the continuous-time model was again the best performer (11 in MP1, 2 in MP2 and 4 in MP3 vs. 11, 5 and 8). More importantly, it reduced the number of changeovers in all three machines, when compared to the scheduling procedure at the plant for the same set of orders (12, 5, 5), indicating that a scheduling tool based on the continuous-time model can bring a higher level of productivity to the plant. Computationally, the higher complexity of the mixed-integer linear program (MILP) resulting from the discrete-time model, forced us to use a rolling-horizon scheme by further dividing the orders in two groups and scheduling them sequentially. Even with such scheme, the relative optimality gap was still significant for MP1 and MP3 after two hours of computational time (11.9 % and 2.9%) while MP2, featuring fewer orders, could be solved to optimality in 406 s.

7. From a sequencing to a scheduling problem

In order not to limit the study to the fixed order-unit assignment case decided by the plant, which significantly reduces the solution space and hence can severely compromise optimality, we released some products from this constraint. The choice took into consideration: (a) the number of orders; many would lead to a great increase in the computer effort;(b) frequency of occurrence of the maximum changeover time (30 minutes) between products allocated to the same machine. The outcome was to free P5 and P6, which represented 7 of the 33 orders allocated to MP1 by the aggregation procedure of the continuous-time model. We studied the effect of releasing either P5 or P6, and both products simultaneously. Naturally, the release process can only be tested if all machines are scheduled simultaneously and enough event points are added to the

time grids of MP2 and MP3 so that the orders can be relocated. Conceptually, we are making a switch from the solution of 3 sequencing problems to 1, more complex, scheduling problem. The results are given in Table 2.
In all three scenarios, the orders were reallocated to MP3, which was no surprise since this is the faster of the available machines. Its contribution to the total paper production increases from 62.1% to 66.7%. In terms of total production time, the reduction was significantly higher for P6 (4 orders) than P5 (3 orders) but was even greater for P5 and P6 (3.39%). It is particularly interesting to highlight that the P6 alternative is the best one in terms of makespan minimization (the maximum over all machines of the production time), leading to a makespan of 20.39 days. The computational effort has increased by over one order of magnitude. It is also interest to highlight that for the same relative optimality gap (1%) CPLEX 11.0 needs just 1269 s to solve the three separate problems, while CPLEX 10.2 required 14800 s. Note however, that the latter reached a slightly better solution (57.60 vs. 57.63).The optimal schedules are given in Figures 2-4.

8. Conclusions

This paper has addressed the optimal scheduling of a Portuguese paper plant. Given the high number of orders, in excess of four-hundred, a solution strategy was devised that consisted on order aggregation and fixed order-unit assignments followed by optimal sequencing of the three paper machines. Three alternative scheduling approaches were evaluated, two unit-specific continuous-time scheduling formulations for multistage multiproduct plants with sequence dependent changeovers and a discrete-time approach. The results showed that the continuous-time model handling changeovers implicitly was the best performer. Finally, we considered the case where a subset of the orders is no longer fixed to a particular unit to solve a more complex scheduling problem, where all machines are tackled simultaneously. Despite the order of magnitude increase in computational effort, we were able to decrease the total production time by 3.3% through the switching of production to the machine with the highest production rate.

References

P. Castro, I. Grossmann, A. Novais, 2006, Ind. Eng. Chem. Res. 45, 6210.
P. Castro, A. Novais, 2008, Ind. Eng. Chem. Res. 47, 6126.
Y. Liu, I. Karimi, 2007, Chem. Eng. Sci. 62, 1549.
C. Mendez, J. Cerdá, I. Grossmann, I. Harjunkoski, M. Fahl, 2006, Comput. Chem. Eng. 30, 913.
M. Shaik, C. Floudas, 2008, Comput. Chem. Eng. 32, 260.

Table 2. Computational results for comparison fixed vs. free P5/P6 assignments (CPLEX 11.0)

		Fixed	P5 free	P6 free	P5 & P6 free
Obj (days)	MP1	21.64	20.94 (-3.27%)	18.98 (-12.3%)	18.25 (-15.7%)
	MP2	16.80	16.80	16.80	16.80
	MP3	19.18	19.54 (+1.84%)	20.39 (+6.31%)	20.68 (+7.82%)
	Total	57.63	57.27 (-0.62%)	56.18 (-2.51%)	55.73 (-3.29%)
CPUs	Total	1269	15657	13325	20411
Production (ton)	MP1	5156	4989 (-3.23%)	4553 (-11.7%)	4386 (-14.9%)
	MP2	1192	1192	1192	1192
	MP3	10405	10571 (+1.60%)	11008 (+5.79%)	11174 (+7.40%)

Figure 2. Schedule for MP1 after releasing P5 & P6

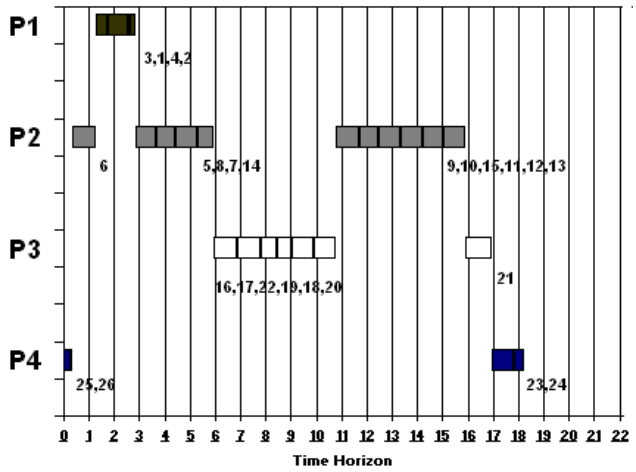


Figure 3. Schedule for MP2

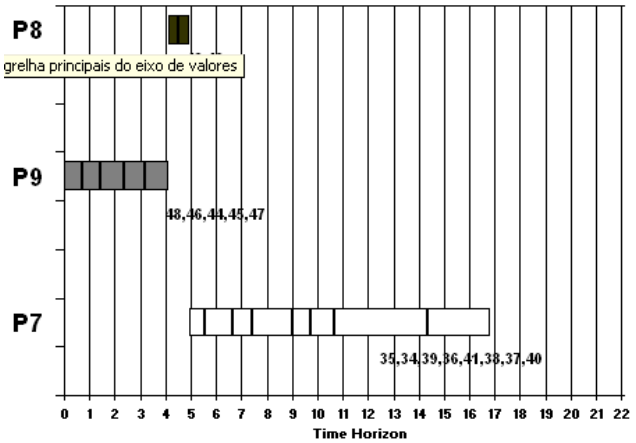
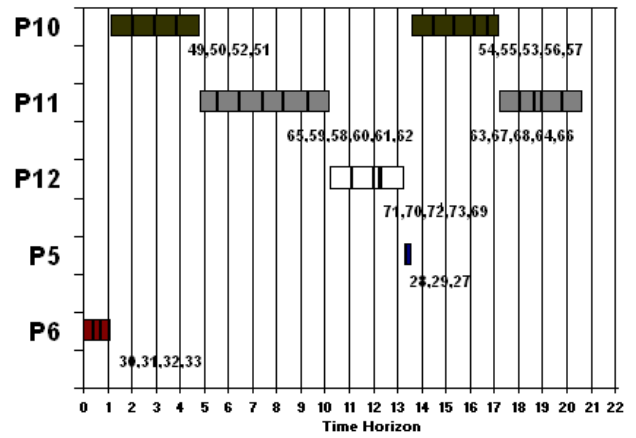


Figure 4. Schedule for MP3 after releasing P5 & P6



10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Entropic-Model-Based PI Controller

João T. Manzi, Antônio C.B. de Araújo, Romildo P. Brito and Heleno Bispo

Department of Chemical Engineering – Federal University of Campina Grande
Rua Aprígio Veloso n.882, Campina Grande-PB, CEP 58109-900, Brazil.

Abstract

Contributions of entropic modelling to the performance of reactive process control have been investigated. The modelling has been developed based on mass, energy and entropy balances and thermodynamics relations, resulting in a model for calculating the entropy production rate. Using the conventional optimization technique, a minimum for the entropy production rate has been found when a given relationship between the temperatures of the inlet stream and of the reaction is satisfied for a particular residence time in the reactor. A new class of nonlinear controller has been developed by means of introducing entropic models into classical PI algorithms designed by reference system synthesis. The results indicate that such a controller yields a superior performance when compared to classical feedback control strategies.

Keywords: PI controller, modeling, entropy, optimization.

1. Introduction

The use of models developed from a mechanistic (first-principles or phenomenological) point of view for the design, analysis, and control of chemical processes is a very powerful, underlying tool in process system engineering, and we are all used to considering such approaches as well-defined for the calculations. However, the thermodynamic treatment of such models are essentially based on the first law of thermodynamics, and do not usually consider the importance of the second law of thermodynamics as the limit factor on the direction of energy transformation.

Simultaneous use of the two laws of thermodynamics on the model building process has been the focus of important research studies. Recently, a procedure which incorporates direct minimization of entropy concepts as essential ingredients in the optimization of chemical process operations has been proposed (Manzi and Carrazzoni, 2008). The implications are such that operation at minimum entropy level can increase process yield while minimizing energy consumption. Furthermore, from a computational standpoint, the dimension of the optimization problem may also be reduced. Despite the advanced control strategies proposed for tackling industrial processes, the practical implementation of optimization policies, as in many other situations in the area of process engineering, usually relies upon the design of a decentralized control structure which makes use of the simplicity and recognized skills of PID controllers. However, it is well-known that these configurations do not yield satisfactory results in closed-loop for highly nonlinear processes when based on linear PID controllers.

Since entropic models can clearly reveal the optimal behavior and intrinsic relationships between the variables and parameters of the system, their introduction in PI algorithms designed by reference system synthesis, resulting in a nonlinear counterpart, can become a very attractive option to enhance closed-loop performance.

This article proposes the design of a class of nonlinear PI controllers for reactive systems using an entropic model in order to generate the control law.

2. Entropic modeling

Let a reactive system, depicted in Figure 1, be represented by the following generic reaction.

$$A_1 + A_2 + \cdots + A_i \rightarrow B_1 + B_2 + \cdots + B_i \qquad \Delta H < 0$$

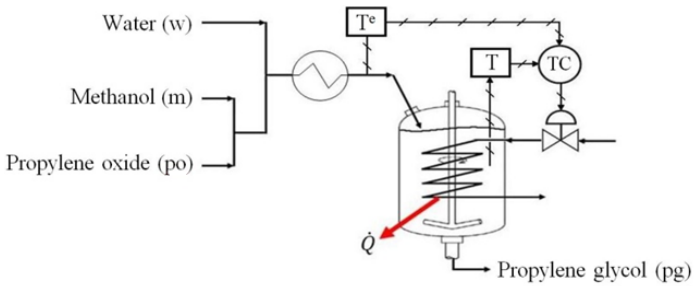


Figure 1: Diagram of a generic reactive system with a simplified feedback loop.

Consider a continuous stirred tank reactor (CSTR) as shown in Figure 1 in which a reactive system is presented consisting of A_i species reactants, resulting in B_i products. Such a reactive system has been mathematically described by means of the following equations representing the mass and energy balances respectively, besides kinetic considerations.

$$dn_{A_i}/dt = F^e C_{A_i}^e - F C_{A_i} - rV \tag{1}$$

$$dn_{B_i}/dt = -F C_{B_i} + rV \tag{2}$$

$$\left(\rho V c_p\right) \mathrm{d}T/\mathrm{dt} = -F^e\left(\sum C_{\mathrm{i}}^{\mathrm{e}} c_{p_i}\right)(T - T^e) + (-\Delta H_{\mathrm{R}})rV - \dot{Q} \tag{3}$$

Additionally, the entropy balance can be given by:

$$\mathrm{dS}/\mathrm{dt} = -\dot{Q}/\mathrm{T} + F^e \rho^e s^e(T^e) - F\rho s(T) + \dot{\sigma} \tag{4}$$

where r and $\dot{Q}$ are respectively the reaction rate and the heat transferred from or to the process.

Since entropy is a function of temperature and of the number of moles of the constituent species of the system, then based on the concept of differential total and taking into account that ΔG_R represents the Gibbs free energy change of the reactive system, the following equation can represent the entropy generation rate for the system considered, as shown by Manzi and Carrazzoni (2008).

$$\dot{\sigma} = -F^e\left(\sum C_{A_i}^e \, c_{p_{A_i}}\right)[(T - T^e)/\mathrm{T} + ln(T^e/\mathrm{T})] + rV(-\Delta G_R/\mathrm{T}) \tag{5}$$

The so-called driving force $(-\Delta G_R/T)$ can be performed by using the classical Gibbs-Helmholtz relationship expressed by:

$$[\partial(\Delta G_R/T)/\partial T]_p = -\Delta H_R/T^2 \tag{6}$$

3. The synthesis of the PI controller

3.1. *Modelling for control purposes*

Consider the arrangement shown in Figure1. Since the rate of heat transferred $\dot{Q}$ derived from the energy balance applied to the cooling medium can be given by its linearized form around the point $(\dot{m}_c^e, T^e)$, then

$$\dot{Q}(\dot{m}_c, T) = \dot{m}_c(T^e - T^c)\left[c_{p_c} - c_{p_c}e^{-\frac{UA}{\dot{m}_c^e c_{p_c}}} - (UA/\dot{m}_c^e)e^{-\frac{UA}{\dot{m}_c^e c_{p_c}}}\right] + UA(T^e - T^c)e^{-\frac{UA}{\dot{m}_c^e c_{p_c}}} + \dot{m}_c^e c_{p_c}(T - T^e)\left[1 - e^{-\frac{UA}{\dot{m}_c c_{p_c}}}\right] \quad (7)$$

Using the Eq.(7) into the Eq.(3) and since the reactor behaves isothermally, the inlet temperature T^e can be chosen as the temperature of reference in the steady state, then, an equation, as a function of the deviation variables defined by $\bar{T} = T - T^e$ and $\bar{\dot{m}}_c = \dot{m}_c - \dot{m}_c^e$, results in:

$$\left(\rho V c_p\right)\mathrm{d}\bar{T}/\mathrm{dt} = -F^e\left(\sum C_i^e\, c_{p_i}\right)\bar{T} + (-\Delta H_R)rV - \bar{\dot{m}}_c(T^e - T^c) \left[c_{p_c} - c_{p_c}e^{-\frac{UA}{\dot{m}_c^e c_{p_c}}} - (UA/\dot{m}_c^e)e^{-\frac{UA}{\dot{m}_c^e c_{p_c}}}\right] - \dot{m}_c^e c_{p_c}\bar{T}\left[1 - e^{-\frac{UA}{\dot{m}_c c_{p_c}}}\right] \quad (8)$$

3.2. The reference system synthesis (*RSS*)

The reference system synthesis, also known as the three steps synthesis (Bartusiak et al., 1989), can be outlined in the following stages, where *x, u, d* and *p* denote the state variable vector, the manipulated input, the disturbance and the model parameters respectively, *y* represents the controlled variable.

- *Stage 1*: *deriving the model.* $dy/dt = G_x f(x,u,d,p,t)\ where\ G_x = \partial g/\partial x$ (9)

- *Stage 2*: *specifying the reference system.* $dy/dt|_{ref} = G_x f(x^*, x, p, l, t)$ (10)

where x^* and l denote the set point and the controller parameters respectively.

- *Stage 3*: *minimizing the difference.*

$$min\left[dy/dt - dy/dt|_{ref}\right] \rightarrow u = h(x^*, x, d, p, l, t) \quad (11)$$

Using the modelling previously developed, as well as taking into consideration desirable behaviour by the control system, such as to return toward its set point and to be free of offset, the following reference trajectory can be specified:

$$\left(\rho V c_p\right)\mathrm{d}\bar{T}/\mathrm{dt}|_{ref} = -F^e\left(\sum C_i^e\, c_{p_i}\right)\bar{T} + (-\Delta H_R)rV - \left[k_c(\bar{T} - \bar{T}^{set}) + k_c/\tau_I\int(\bar{T} - \bar{T}^{set})dt\right](T^0 - T^c)\left[c_{p_c} - c_{p_c}e^{-\frac{UA}{\dot{m}_c^e c_{p_c}}} - (UA/\dot{m}_c^e)e^{-\frac{UA}{\dot{m}_c^e c_{p_c}}}\right] - \dot{m}_c^e c_{p_c}\bar{T}\left[1 - e^{-\frac{UA}{\dot{m}_c c_{p_c}}}\right] \quad (12)$$

Then, applying the reference system synthesis-based approach, the control algorithm can be found as:

$$\dot{m}_c = \dot{m}_c^e + k_c(\bar{T} - \bar{T}^{set}) + k_c/\tau_I\int(\bar{T} - \bar{T}^{set})dt \quad (13)$$

It is evident that Eq. (13) describes the PI control law. Since $\bar{T} = T - T^e$ can be related to the entropy production rate, the algorithm developed is in fact a controller based on the entropic concept.

4. Results and Discussion

To illustrate the analysis of such a control system when the process is submitted to a minimum entropy generation rate, consider, for simulation purposes, the production process of propylene glycol in a CSTR reactor reported by Manzi et al. (2009), and depicted in Figure 1. The reactive process occurs according to the reaction below, in which sulphuric acid has been used as a catalyst, while the maximum temperature should not exceed 324.8 K (Fogler, 1999). Table 1 provides the physical and chemical properties and the base steady state operating conditions applied to this system.

$$C_3H_6O + H_2O \rightarrow C_3H_8O_2 \qquad \Delta H < 0$$

Table 1 – Operating conditions and parameters for the CSTR applied to the production of propylene glycol

Variable or	Value	Variable or	Value
F^e	2.567 L s^{-1}	$c_{p_{pg}}$	192.59 J mol^{-1} K^{-1}
τ	5,842.44 s	c_{p_m}	81.64 J mol^{-1} K^{-1}
C_{op}	2.12 mol L^{-1}	k_0	47.11x10^8 s^{-1}
T_r	298 K	E	75,320 J mol^{-1}
T^c	302.8 K	R	8.314 J mol^{-1} K^{-1}
Θ_w	18.65 -	U	567.83 J s^{-1} m^{-2} K^{-1}
Θ_m	1.67 -	A	35.7 m^2
$c_{p_{op}}$	146.54 J mol^{-1} K^{-1}	$\Delta H_{R_r}(298K)$	-84,589.11 J mol^{-1}
c_{p_w}	75.36 J mol^{-1} K^{-1}	$\Delta G_{R_r}(298K)$	-68,274.08 J mol^{-1}

4.1. Entropy production rate

According to Manzi et al. (2009), the necessary condition for the minimum entropy production rate requires $d\dot{\sigma}/dT$ and $d\dot{\sigma}/d\tau$ to be equal to zero. Thus the minimum value of $\dot{\sigma}$ only is achieved when the following relationships are satisfied:

$$(T^e - T) = (\tau k_0)/\left(\sum \Theta_i c_{p_i}\right)\left\{(E/R)e^{\left(-\frac{E}{RT}\right)}\Big/\left[\left(1+\tau k_0 e^{\left(-\frac{E}{RT}\right)}\right)^2\right]\left(-\frac{\Delta G_r}{T_r}+\frac{\Delta H_r}{T_r}\right)\right.$$
$$\left.+\Delta H_r\left[e^{\left(-\frac{E}{RT}\right)}\Big/\left[1+\tau k_0 e^{\left(-\frac{E}{RT}\right)}\right]-(E/R)e^{\left(-\frac{E}{RT}\right)}\Big/\left(T\left(1+\tau k_0 e^{\left(-\frac{E}{RT}\right)}\right)^2\right)\right]\right\} = 0 \quad (14)$$

$$\varphi(\tau) = \left[\frac{1}{\left(\tau k_0 e^{(-E/RT)}+1\right)^2}\right] = 0 \quad (15)$$

where $\varphi(\tau)$ denotes the essential result of the differentiation procedure ($d\dot{\sigma}/d\tau$).

Using the data provided by Table 1, the behavior of the temperature difference $(T^e - T)$ given by Eq. (14) can be depicted as shown in Figure 2, revealing that the intersection point between the $(T^e - T)$ curve with the temperature axis is the optimal

solution for Eq. (14), which indicates this to be $T^e = T$ the optimum operating temperature. As shown by Manzi et al. (2009), the optimized set of operating conditions derived from the entropy production rate yields the best value of conversion for the system, when compared to the classical modus operandi.

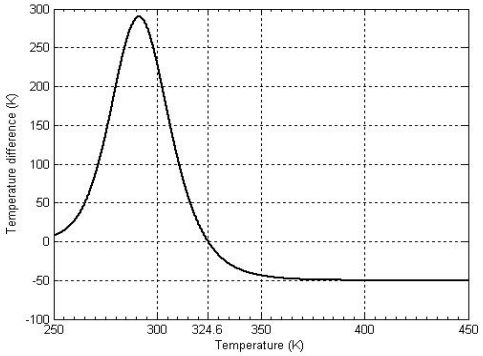


Figure 2: Temperature difference $(T^e - T)$ versus Temperature.

4.2. Entropic-model-based PI controller

The control structure, the so-called "entropic-model-based PI controller", results from Eq. (13) in which the signal of the error based on Eq. (14) is intrinsically related to the minimum entropy production rate. Since the temperature of the reactive system under the condition of the minimum entropy production rate requires $T^e = T$, then the input to the controller is given only by $\bar{T} = T - T^e$, that is, $\bar{T}^{set} = 0$, which results in the following nonlinear control law:

$$\dot{m}_c = \dot{m}_c^e + k_c\bar{T} + k_c/\tau_I \int \bar{T}dt \tag{16}$$

It is meaningful to observe that it is not necessary to specify the temperature desired for operating of the controller, because this temperature is implicitly determined by $\bar{T}^{set} = 0$, depicted in Figure 2, this being the unique solution for the data from table 1. It can be also observed that the search for the temperature desired follows a strategy which corresponds to the search for the minimum of the entropy production rate, and since, in practice, $T - T^e \geq 0$ is reflected straightforwardly in the settling time of the controller.

Figure 3 presents the comparative responses of the temperature of the reactive system when an entropic-model-based PI controller and a classical PI are used. In both cases, a disturbance has been introduced into the inlet temperature of the heat exchanger in addition to which the classical Ziegler-Nichols tuning procedure was applied with a fine adjustment based on trial and error in order to obtain the best set of tuning parameters. The results show that the entropic PI controller exhibits a much less pronounced oscillatory behavior, and has the lowest value for the integral of the absolute value of the error (IAE), with the response approaching its ultimate value asymptotically. It must also be emphasized that while the classical structure deals only with the controlled output, the entropic PI controller works under the condition of the minimum entropy production rate, and always yields the best performance for the whole reactive system.

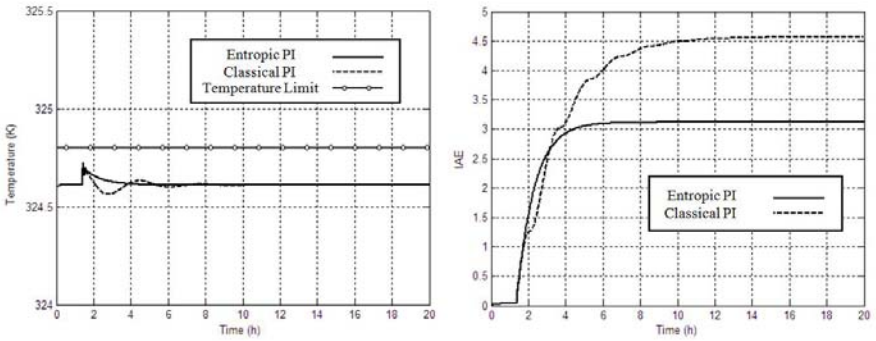


Figure 3: Temperature and IAE profiles from classical and entropic PI controllers when submitted to a disturbance in the inlet temperature of the heat exchanger from 290 to 295 K.

5. Conclusion

The entropic-model-based PI controller has been derived from relationships obtained by minimizing the entropy production rate, which allows the reactive system to reach the optimal conditions, in connection with the reference system synthesis. Since such optimal conditions have been introduced in the control structure, the control action moves the system towards the minimum entropy production rate, and satisfies the essential requirement $T = T^e$. Therefore, an advantage of such a control structure is due to the fact that it deals not only with the controlled output but above all with control of the whole process.

In addition, the entropic-model-based PI controller presents a shorter oscillating, faster response and the lowest IAE, and therefore reaches the desired value quickly and asymptotically. Thus, such results allow the conclusion to be reached that the entropic-model-basic PI presents a superior performance when compared with the conventional PI controller.

References

H. S. Fogler, 1999, Elements of chemical reaction engineering, Prentice Hall, New Jersey

J. Manzi, R. Viana and H. Bispo, 2009, Direct entropy minimization applied to the production of propylene glycol, Chemical Engineering and Processing: Process Intensification, 48, 1, 470-475

J. Manzi and E. Carrazzoni, 2008, Analysis and optimization of a CSTR by direct entropy minimization, Journal of Chemical Engineering of Japan, 41, 3, 194-199

R. D. Bartusiak, C. Georgakis and M. J. Reilly, 1989, Nonlinear feedforward/feedback control structures designed by reference system synthesis, 44, 9, 1837-1851

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

On-line Fault Detection on a Pilot Flotation Column Using Linear PCA Models

Luis G. Bergh and Sebastián Acosta.

Chemical Engineering Department, Santa Maria University, Valparaiso, Chile

Abstract

On-line fault detection, for instrumentation and process operation, has become important part of industrial programs leading to improve process operation and therefore product quality over time. Multivariate statistical projection methods, such as Principal Component Analysis (PCA), have been proposed to effectively deal with these situations.

In this work, a pilot flotation column is operated under distributed control of froth depth, gas hold up and bias, to experimentally collect operation data at steady state, to build a PCA model. The basic control is implemented in a PLC, and all data is communicated to a PC network for displaying and further processing, under Intouch software. The column is operated in a hybrid form, for the air water system, while concentrate and tailing grades are obtained by on line predictions by using a static metallurgical model. A steady state on-line detector has been implemented on the PC to test when the collected data met the requirements to be used to build a PCA model. Several examples are discussed, detecting both particular instrumentation failures and abnormal operating conditions, and how using the system suggestions the metallurgical objectives of the process are met again.

Keywords: Fault detection, modeling, projection methods, flotation columns, process control.

1. Introduction

In the last two decades the use of pneumatic flotation columns became wide-spread throughout the mineral processing industry of metallic, non-metallic and coal ores in the world. Columns out perform conventional mechanical cells in cleaning operations (better product grade) due to their particular froth operation (Finch and Dobby, 1990).

1.1 Flotation column control

The primary objectives are column recovery and concentrate grade, which represent the indices of process productivity and product quality. The on-line estimation of these indices usually requires a significant amount of work in maintenance and calibration of on-stream analyzers, in order to maintain good accuracy and high availability. Therefore, a common practice is to control secondary objectives, such as pH at the feed, froth depth, air flow rate and wash water flow rate. These are usually implemented as local controllers or under distributed control systems (DCS). Ideally, when primary objectives are measured, the control strategy is to change the set points of the

controllers under DCS, in order to achieve a good process performance. This is usually implemented in the form of expert systems (Bergh and Yianatos, 1993, 2003). On line analysers, tailings, feed flow rates and some other measurements are often incorporated into the system when a supervisory control strategy is implemented on top of a distributed control system.

1.2 Pilot flotation column

The pilot column was operated for the air-water (and frother) system coupled with an on-line steady state model to predict output stream grades. The convenience of the approach of combining on-line process measurements and models to empirically test strategies for process control, monitoring and diagnosis, was recently discussed by Bergh (2007). A flotation column phenomenological model was developed, following Finch and Dobby (1990). Following Figure 1, first the gas holdup, the bias rate and the kinetic constants for two mineralogical species are estimated from semi-empirical models, depending on operating variables such as feed flow rate, gas flow rate, wash water flow rate and froth depth. Then, dispersion number, residence times, froth and collection recovery are estimated. Finally concentrate and tailings grades are predicted. The semi-empirical model parameters were fitted using experimental data. More details can be found in Bergh et al. (1998).

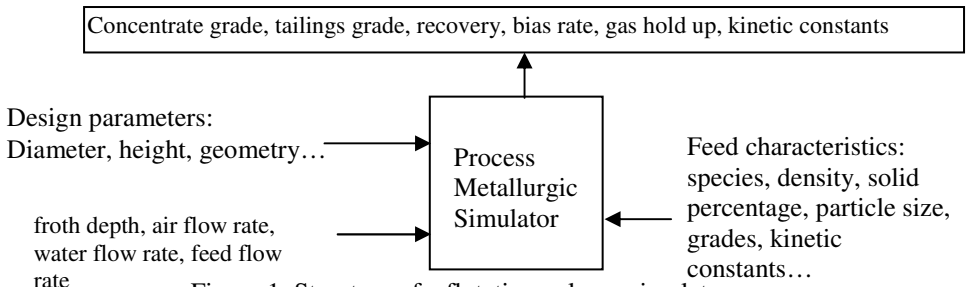


Figure 1. Structure of a flotation column simulator.

1.3 Pilot flotation column control

The column control is shown in Figure 2. There are three tertiary control loops: air flow rate, tailings flow rate and wash water flow rate. Feed flow rate is also measured and controlled. A hydrodynamic supervisory control is implemented to control gas hold up in cascade with air flow rate, froth depth in cascade with tailings flow rate and bias in cascade with wash water flow rate. All the operating variables are measured and communicated to a PLC, where the DCS has been implemented. All signals are communicated to a PC network, where the monitoring and hydrodynamic supervisory control is running under Intouch software. In the PC network, the steady state test is performed on-line. When the process reaches a steady state, the predicting model is solved on-line to estimate the concentrate and tailing grades. All data is displayed on PC screen.

2. PCA models

The key feature of PCA method is their ability to mathematically project high dimensional process and quality data into smaller dimensional, summary data sets via the development of linear models. The practical value of PCA modeling method is that

this technique allows for the systematic examination and interpretation of the model outputs. Examination of the model outputs can provide insight into the operation of an industrial process during monitoring and quality assurance activities. With PCA, the systematic interpretation of dominant patterns in the data and the isolation of the most important contributors to these patterns are possible. This allows the classification of data relationships according to normal and abnormal operation. Some of these numerous advantages of PCA method has over traditional monitoring and prediction technologies are: provision for data dimension reduction and robustness to highly correlated, noise and missing data (Kourti and MacGregor, 1995).

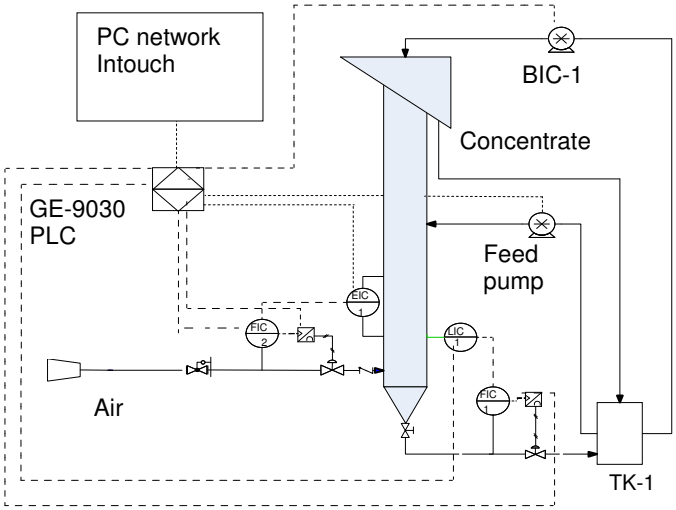


Fig. 2. Pilot flotation column control

The concept of a latent variable model is that the true dimension of a process is not defined by the number of measured variables, but by the underlying phenomena that drive the process. The latent variables themselves are modeled as mathematical combinations of the measured variables and describe directions of variation in the original data. A latent variable model can contain many fewer dimensions than the original data, it can provide a useful simplification of large data sets, and it can allow better interpretation of the measured data during analysis (MacGregor *et al*., 2007).

3. PCA model building

The original **X** matrix consisted of sixteen variables, shown in Table 1, and 2550 observations of steady state data. The experiments were designed to cover the maximum possible variation of the main independent variables, and were conducted under closed loop control. The data was processed using PLS_Toolbox from Eigenvalue Research.

A PCA model was built from 1800 sets of data corresponding to a normal condition. A model with 6 latent variables (scores) was found to explain at least 92 % of the variance in the centered and scaled pretreated data. The contribution of each operating and

quality variables (in the same order given in Table 1) to each PC is presented in Figure 3.

Table 1. Operating and quality variables considered.

N°	Variable	Tag	N°	Variable	Tag
1	Froth depth	z	9	Tailings superficial velocity	Jt
2	Gas hold up	E	10	Feed superficial velocity	Jf
3	Dp/cell low	LL	11	Wash water superficial velocity	Jw
4	Dp/cell high	LH	12	Cu recovery	R
5	Pressure to air control valve	PA	13	Concentrate Cu grade	CCG
6	Pressure to Tailings control valve	PT	14	Feed particle size d80	D
7	Bias superficial velocity	Jb	15	Feed Cu grade	FCG
8	Air superficial velocity	Jg	16	Feed solid percentage	S

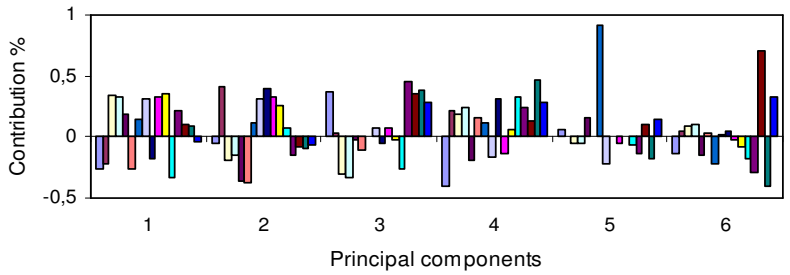


Fig. 3. Contribution of each sixteen operating and quality variables to each six PC

All variables were included in the model as a linear combination for each six scores. The first score showed the contribution of groups of correlated variables, while the last ones mostly represented the bias and the feed characteristics. For monitoring the process the Hotelling T^2 limit was found to be 12.6, while the Q residuals was 3.81.

On-line tests were implemented based on the following criteria:

(i) Normal operation if the new set of data satisfies the Q and T^2 test.
(ii) Abnormal operation if the T^2 test is failed. If the Q test is passed then the model is adequately representing the process. If the Q test is also failed then either the model is no longer appropriate or a measurement problem occurred.
(iii) Measurement problem or PCA model representation problem if only T^2 test is satisfied.

In this way a diagnosis of the operation can be accomplished for steady state data. Furthermore, the residuals are informative of the principal process variables affecting the abnormal situation.

4. Experimental results

Experiments were carried on to test when the process is out of control and an abnormal operating condition is met. Two results are presented: when the process is at steady state and during the transient period. One example is shown in Figure 4, where the T^2 and Q test has been followed for over 600 samples, taken every 5 (s).

One can see that most of the time the Q test is satisfied, while T^2 test is failed at intervals 130-200, 300-430 and 480-560. On these same periods, the concentrate grade is too low and recovery is high or concentrate grade is too high and recovery is low, then an abnormal operation has been detected. To identify which variables are causing this, the individual contribution to the T^2 residuals, for sample 512, are shown in Figure 5. One can see that the main contribution were the froth depth and the high and low dp/cells. All variables consistently showed that the problem is due to a low froth depth, causing high recovery and low concentrate grade. Figure 6 shows the froth depth changes during the whole period. If the froth depth were change from 50 to 100 cm, as is shown at sample 600, the column operation is driven back to a normal condition, as can be seen from the previous figures.

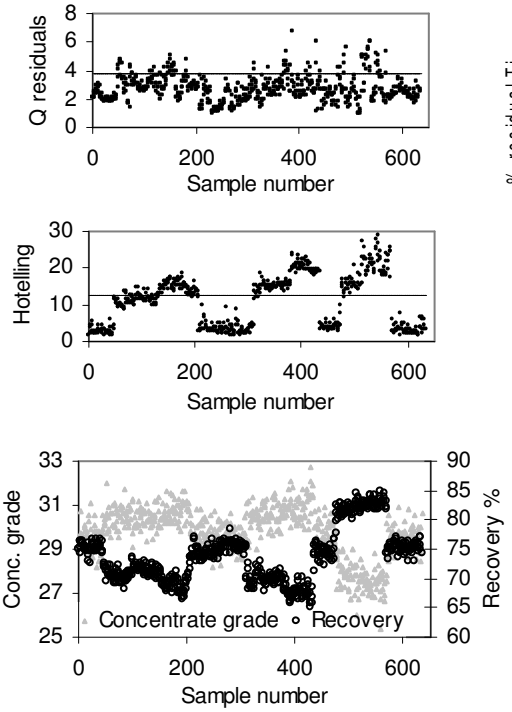


Fig. 4. Operating condition test

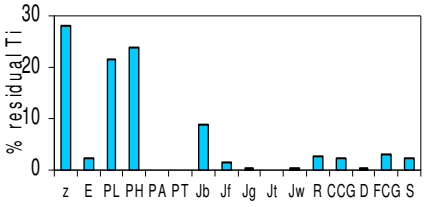


Fig. 5. Contributions to T^2

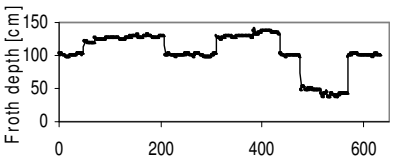


Fig. 6. Froth depth period

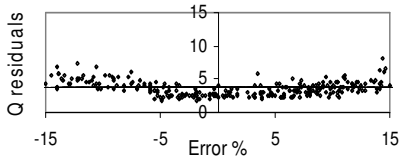


Fig. 7. Failure on concentrate grade

When only the Q residuals test fails, the device measuring the isolated variable must be recalibrated o replaced. Several tests were carried on to find the sensibility of the monitoring test to the extension of the fail, measured in percentage of error. Errors less than 5% on pressure to control valves, 7% on Dp/cells, 15% on flow meters and 10% on virtual measurements of concentrate grade were detected. These error limits were found for a large number of different operating conditions. One example is shown in Figure 7 for the virtual measurement of copper concentrate grade.

The same PCA model was used to test abnormal operation either because of decision based on failed sensors or process variable deviations. The PCA model relays on the selected data. If the data collected represents a narrow band of operation around the targets, it may be expected that abnormal conditions, as a result of a combination of

process variable deviations, will be easily detected. A model built on such selected data will be less useful to identify measurements problems. The model used in this work was based on data corresponding to a wide operation zone, favouring the detection of sensor failures. A best approach to be tested is the use of different PCA models, based on different data, for each purpose.

Another surprising aspect was that even when the process was in a transient, and the grades predictions were not entirely valid, the on-line application of the T^2 and Q test provided information about when an abnormal operation was in progress. Further research is needed to asses this fact.

5. Conclusions

Flotation column control quality is strongly depending on the accuracy of measurements and estimations. The flotation process is complex and it is a real challenge to decide which variables are to be changed in order to drive back the process to a normal operation.

The application of multivariate statistical methods, and particularly PCA, is a powerful tool to build linear models containing the essential of the process phenomena with the minimum number of latent variables. The application of PCA models to monitoring flotation columns has been demonstrated. These PCA models can be effectively used as part of a supervisory control strategy, specially when control decisions are infrequently made.

Acknowledgement

The author would like to thanks Conicyt (Project Fondecyt 1070106) and Santa Maria University (Project 270968) for their financial support.

References

Bergh L.G. and J.B. Yianatos (1993). Control Alternatives for Flotation Columns. *Minerals Engineering*, **6**, Nº 6, 631-642.

Bergh L.G. (2007). Combining on-line process measurements and models to empirically test strategies for process monitoring, diagnosis and control, *Proceedings 12th IFAC Symposium MMM 07, Quebec City, August 21-23,* 193-198.

Bergh L.G. and J.B. Yianatos (2003). Flotation Column Automation: state of the art. *Control Engineering Practice*, **11**, 67-72.

Bergh L.G., J.B. Yianatos and C. Leiva (1998). Fuzzy Supervisory of Flotation Columns. *Minerals Engineering*, **11**, N° 8, 739-748.

Finch J.A. and G.S. Dobby (1990). *Column Flotation*. Pergamon Press.

Kourti T. and J.F. MacGregor, (1995). Process Analysis Monitoring and Diagnosis using Multivariate Projection Methods – A Tutorial, *Chemometrics and Intelligent Laboratory Systems*, **28**, 3-21.

MacGregor J.F., T. Kourti, J. Liu, J. Bradley, K. Dunn, H. Yu (2007). Multivariate methods for the analysis of databases. process monitoring, and control in the material processing industries, *Proceedings 12th IFAC Symposium MMM 07, Quebec City, August 21-23,* 193-198.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Measurements of air quality using Lidar system

Juliana Steffens[a], Eduardo Landulfo[b], Roberto Guardani[a], Cláudio N. Oller[a], Andréia Moreira[c]

[a] *Universidade de São Paulo, Escola Politécnica, Departamento de Engenharia Química, Av. Luciano Gualberto Trav. 3 N. 380, Butantã, 05508-900, São Paulo - SP, Brazil.*
[b] *IPEN-Instituto de Pesquisas Energéticas e Nucleares, Avenida Prof. Lineu Prestes 2242,Cidade Universitária, 05508-000, São Paulo, Brazil*
[c] *PETROBRÁS, Petróleo Brasileiro S.A., Ilha do Fundão, Rio de Janeiro, Brazil*

Abstract

Air pollution is a widely recognized hazard to human health. In industrial cities the emission of toxic gases and particulate matter create hazardous public health situations. Is the case of the industrial complex of Cubatão, state of São Paulo that is one of the largest petrochemical and industrial in Brazil that has been subject of severe damage caused by massive emissions of pollutants, as a result of the progressive industrialization in the area. Therefore it is necessary to monitory the area to be able to control and to prevent ambient problems. In a partnership with the University of São Paulo (USP) the Brazilian oil company PETROBRAS has started off an Environmental Research Center - CEPEMA- located in the industrial site, in which the development of fieldwork will be carried out. The current joint R&D project focus on the development of real time acquisition system, together with automated multicomponent chemical analysis. Additionally fugitive emissions from oil processing and storage sites will be measured, together with the main greenhouse gases (CO_2, CH_4) and aerosols. Our first effort is to assess the potential chemical species coming out of an oil refinery site and to use Raman technique LIDAR for detecting and quantifying the particular gas. Raman lidar techniques have been demonstrated which provide most valuable descriptions of the evolution of air pollution events. The vibrational and rotational Raman LIDAR signals provide simultaneous profiles of meteorological data, ozone and measurements of airborne particulate matter.

Keywords: Raman LIDAR, air quality, fugitive gas.

1. Introduction

Methane plays an important role in the Earth's atmospheric chemistry and radioactive balance. It is the most abundant hydrocarbon in the atmosphere, and the second most important greenhouse gas after CO_2. It has a relatively short lifetime (8-12 years), but its strong IR absorption band at 7,66 μm, where water and CO_2 absorb weakly, makes methane an effective contributor to the radiative forcing with a global warming potential of 21 over a 100 year period (IPCC, 2001). Another important role in the atmosphere is the atmospheric aerosol, because is a complex mixture of particles from a large number of discrete sources (Ondov and Wexler, 1998). Its size distribution, composition, morphology, and source strenghts can vary significantly with meteorology, location, and time. This is especially true for urban atmospheres wherein emissions from

concentrated industrial and dispersed sources lead to urban pollutant excesses above rural background concentrations (IPCC, 1996).

Today, most industrial plants are obligated to measure their own emissions. Accurate measurements of the concentration and the flux of different gases are important, not only from an environmental point of view, but also for economical reasons, since a part of taxation may be dependent on the emission values. Different types of in situ instruments are frequently used to monitor the main emission sources, but external optical methods are becoming increasingly important. The advantages with optical methods are online evaluation of the measurement, no contamination of samples, measurement a long distances and the possibility of measuring large areas (Sigrist, 1994).

Certain industrial activities are of key influence on the air quality and nowadays have triggered the awareness over long term atmospheric changes as well. Their monitoring became not only a problem of mitigating human health impact due to air pollution but also a matter of climate change. Among the industries with influence on the air quality the petrochemical sector plays one of the most important roles since the economical growth has augmented the diversity of products and storage capacity.

Following this increase in activity the need for monitoring systems in these sites and their neighboring areas have also become tasks of great concern, especially because data acquisition and analysis in practically real time have turned into a sine qua non feature in such systems. Therefore the technique and equipments to be employed have to be capable of detecting in real time the emission of aerosols (particulate matter), polluting air gases, such as NOx, Ozone, Sulphur Compounds and fugitive gases in industrial and storage sites(namely Volatile Organic Compounds - VOCs). Recent advances of tunable all solid state laser systems and in the Lidar technique (Light Detection and Ranging) opened new perspectives in the 3D-analysis of atmospheric pollution dynamics. 3D mappings of concentrations of pollutants have been obtained, allowing a direct access to the physical and chemical dynamics of air pollution (Frejafon *et al*, 1998). Besides their detection, these species have to be tracked along neighboring areas by the use of dispersion models. Given the extension and quantities involved, meteorological sensors and even satellites have to be employed. Based on the situation described, the use of LIDAR-based technique is proposed in the region of Cubatão, Brazil, one of the largest petrochemical and industrial sites in Brazil, where petrochemical activities exist since about 50 years.

In a partnership with the University of São Paulo (USP) the Brazilian oil company PETROBRAS has created CEPEMA, an Environmental Research Center located in the industrial site, where the development of the monitoring technique is supposed to take place. It is intended to develop a system capable of real time acquisition and by the use of automated multicomponent chemical analysis techniques, this system should be able to identify fugitive gases from oil processing and storage sites, and also the main greenhouse gases, CO_2 and CH_4; and aerosols. The use of ancillary equipment should be carried out for validation and calibration purposes. The system should operate with a Fourth-Harmonic-Generated Nd:Yag UV laser source (266 nm) and detect the Raman scattered lines of the species mentioned above. Lower harmonics (355 nm, 532 nm and 1064 nm) should be used for the aerosol backscattered signal, and also capable of particle size determination. This paper should cover the specific needs for a air monitoring spectrocopy system in operation in situ in a refinery site.

2. Oil Refinery Gas Inventory

The air environment in an petrochemichal refinery is rich in hydrocarbon compounds which have rotationalvibrational transitions in the mid-IR wavelength region. Of the most popular carbon-containing compounds present in such an environment we highlight the class of alkanes or paraffins, such as Methane CH_4, Ethane C_2H_6, Hethane C_6H_{14} and Hexane C_4H_{10}, the class of alkenes, such as Ethene or ethylene C_2H_4 and Propene or propylene C_3H_6 , followed by the Alkynes with Acetylene C_2H_2, Aromatics Benzene C_6H_6 and Toluene C_7H_8 and finally the Aldehydes, Formalaldeyde $HCHO$ and Acetaldehyde CH_3CHO. Its is estimated that globally around 17.5 Tg of Nonmethane Volatile Organic Compounds is emitted to the atmosphere by the fuel production/distribution industry, from that about 30% is due the Oil refining branch (Seinfeld and Pandis, (1997).

3.Raman Lidar

3.1- Lidar remote sensing

The basis for lidar remote sensing lies in the interaction of light with gas molecules and particulate matter in suspension in the atmosphere (aerosols). More particularly, a lidar uses a laser (emitter) to send a pulse of light into the atmosphere and a telescope (receiver) to measure the intensity scattered back (backscattered) to the lidar. By measuring the scattering and attenuation experienced by the incident pulse of light, one can investigate the properties of the scatterers (concentration of gaseous species, aerosol distribution and optical properties, cloud height) located in the atmosphere. The light scattered back to the detector comes from various distances, or ranges, with respect to the lidar. Because the light takes longer to return to the receiver from targets located farther away, the time delay of the return is converted into a distance (range) between the scatterers and the lidar, since the speed of light is a well-known quantity. By pointing the laser beam in various directions and at various angles with respect to the ground surface (scanning), a ground-based lidar system can gather information about the three-dimensional distribution of aerosols in the atmosphere.

3.2 – Raman technique

Raman scattering is one of the processes that occurs when optical radiation is scattered from the molecules of the atmosphere. It is most useful because the vibrational Raman scattering provides distinct wavelength shifts for species specific vibrational energy states of the molecules and rotational Raman scattering provides a signal with a wavelength shift that depends directly upon the atmospheric temperature (Philbrick, 1994).

Figure 1(a) shows a diagram of the vibrational and rotational energy levels that are associated with Raman scatter. When a photon scatters from a molecule, the redistribution of the charge cloud results in a virtual energy state. Most of the atmospheric molecules reside in the ground vibrational level because the vibrational excitation corresponds to relatively large energy transitions (tenths of eV), for simple molecules like nitrogen and oxygen, compared to the thermalenergy available. After the scattering occurs, most of the events result in the return of the molecule to the ground state and the emitted photon has the energy of the initialphoton plus/minus the random thermal velocity of the molecule, that is the Doppler broadening. A small fraction of the

transitions (order of 0.1%) result in giving part of the photon energy to the molecule, and ending in the first vibrational level (a Stokes transition). The emitted photon energy is decreased by exactly the energy of the vibrational quanta for that mole cule. For the small fraction of molecules existing in the vibrational excited level, the unlikely anti-Stokes transition is possible. The relative sensitivity of the scattering from the vibrational and rotational states is indicated by the scattering cross-section values for scattering by a frequency doubled Nd:Yag laser at 532 nm shown in Figure 1(b). The wavelengths of vibrational Raman back scatter signals from the molecules of the water vapor and molecular nitrogen are widely separated from the exciting laser radiation and can be easily isolated for measurement using modern filter technology and sensitive photon counting detectors (Balsiger, and Philbrick, 1996).

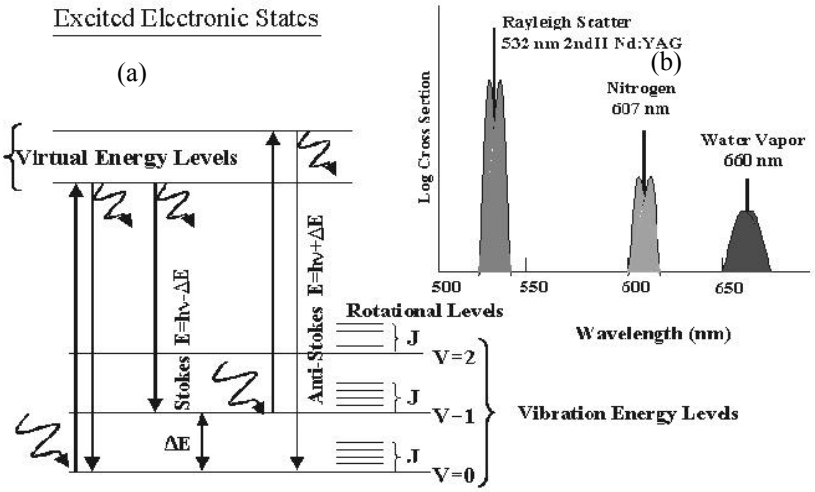


Figure 1 (a) The energy diagram of a molecule illustrates that the scattering of a photon raises the molecule to a virtual level which normally decays to ground (V=0) emitting a photon of the same energy as the incident energy, only broadened by thermal Doppler velocity. In a small fraction of cases, the return is Raman shifted to the first vibrational level (V=1), a Stokes shift. The relatively large vibrational energy (ΔE) compared with thermalenergy makes the Anti-Stokes vibrational transition unlikely,however, the rotational states (J-levels) are populated by thermal excitation, (b) The relative intensities of the Stokes vibrational Raman shifts of oxygen, nitrogen and water vapor are indicated for illumination of atmospheric molecules with the 532 nm laser. The expected line widths and the relative rotational states are indicated (Philbrick, 1994).

The ratio of rotational Raman signals at 528 nm and 530 nm provides a measurement which is sensitive to atmospheric temperature (Balsiger *et al.*, 1996). All of the molecules of the lower atmosphere are distributed in the rotational states according the temperature. By measuring the ratio scattered signals at two wavelengths in this distribution, the temperature can be directly measured. In order to push the lidar measurement capability into the daylight conditions, we have used the "solar blind" region of the spectrum between 260 and 300 nm. The “solar blind” region is darkened by the stratospheric ozone absorption of ultraviolet radiation. Night time measurements are made using the 660nm/607nm (H_2O/N2) signal ratio from the doubled Nd:Yag laser radiation at 532 nm. Daylight measurements are obtained using the 295nm/284nm

(H_2O/N_2) ratio from the quadruple Nd:YAGlaser radiation at 266 nm. A small correction for the tropospheric ozone must be applied. That correction can be obtained from the ratio of the O_2/N_2 signals 278nm/284nm, and from this analysis the ozone profile in the lower troposphere is also obtained (Esposito and Philbrick, 1998) The Raman techniques, which use ratios of the signals for measurements of water vapor and temperature, have the major advantage of removing essentially all of uncertainties, such as any requirement for knowledge of the absolute sensitivity and non-linear factors caused by aerosol and cloud scattering (Philbrick, 1998). Optical extinction is measured using the gradient of the measured molecular profile compared with that expected for the density gradient. Since the Raman signal is only scattered from the molecular component of the scattering volume, and difference in the gradient of the signal from that expected due to loss frommolecular scattering and absorption can be used to calculate the aerosol extinction (O'Brien *at al*.,1996).

4.Lidar system design

This system should be deployed at CEPEMA, Centro de Pesquisa do Meio Ambiente, Cubatão, in a neighboring region to a PETROBRAS refinery station. This system should be built not only to monitor the air quality but also to diagnose the refining process performance by inspecting the particle size distribution coming out of the chimneys. The telescope employed should be able to scan from 0^o to 90^o azimuthally and thus be capable of full spatial coverage.
The dates will be analyzed by Statistical technical to interpreted results.
At Figure 2 we present one Schematic diagram of Raman Lidar that should use in the experiments.

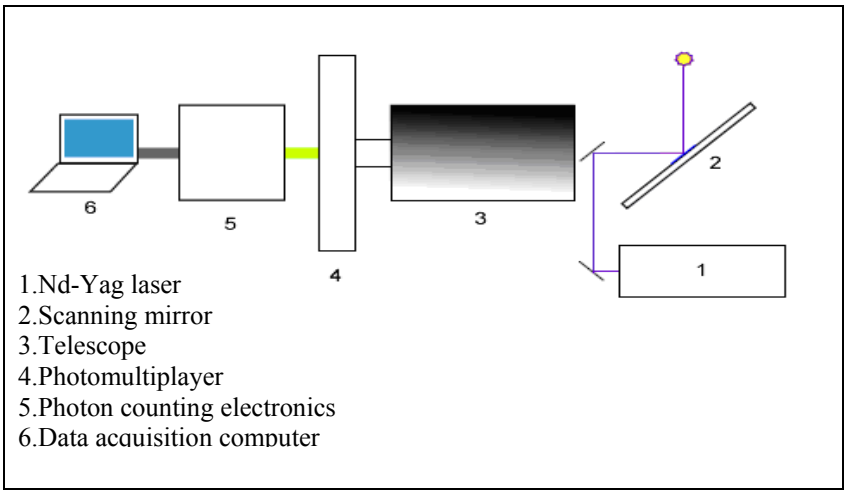


Figure 2- Schematic diagram of Raman Lidar.

5.Conclusion

Lidar technology has matured over the last decade to an extent that many applications are becoming routine. Meteorological parameters like the wind and temperature, as well as trace gas concentrations, can be measured in short time periods with high spatial resolution in three dimensions, often at large distances. Plumes containing air pollutants like hydrocarbons , VOCs and others can be measured by lidar at strategic sites.

The current joint R&D project focus on the development of real time acquisition system, together with automated multicomponent chemical analysis. Additionally fugitive emissions from oil processing and storage sites will be measured, together with the main greenhouse gases (CO_2, CH_4) and aerosols. Our first effort is to assess the potential chemical species coming out of an oil refinery site and to verify if Raman Lidar is efficient in detecting and quantifying the particular gas.

References

Intergovernmental Panel on Climate Change, 2001. Climate change 2001. In: Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., Van Der Linden, P.J., Dai, X., Maskell, K., Johnson, C.A. *The Scientistic Basis*. Cambridge University press, UK.

Intergovernmental Panel on Climate Change, 1996. Climate change 1996. The science climate chenge. Cambridge University press, New York.

M.W.Sigrist, 1994. Air monitoring by spectroscopic techniques, *Chemical analysis*, Vol, 127, New York.

J.M. Ondov and A.S.Wexler, 1998. Where do particulate toxins residue? An improved paradigm for the structure and dynamics of the urban mid-atlantic aerosol, *Environmental Science Technology*, Vol. 32, p. 2547-2555.

E.Frejafon, J. Kasparian, P. Rambaldi, B., Vezin, V, Boutou, J. Yu, M. Ulbricht, D. Weibauer, B. Ottobrini, E. de Saeger, B. Kramer, T. Leisner, P. Rairoux, L. Woste, J.P., Wolf. (1998) Laser applications for atmospheric pollution monitorin. *The European Physical Journal D*, Vol.4, p. 231-238.

C.R.Philbrick, (1994),Raman Lidar Measurements of Atmospheric Properties, *Atmospheric Propagation and Remote Sensing III*, SPIE Vol. 2222, p.922-931.

F. Balsiger, and C. R. Philbrick, (1996), Comparison of Lidar Water Vapor Measurements Using Raman Scatter at 266nm and 532 nm, *Applications of Lidar to Current Atmos. Topics*, SPIE Proc. Vol. 2833,p. 231-240.

F. Balsiger, P. A. T. Haris and C. R. Philbrick, (1996), Lower-tropospheric Temperature Measurements Using a Rotational Raman Lidar, *Optical Instruments for Weather Forecasting*, SPIE Proc. Vol. 2832, 53-60.

S. R. Esposito and C.R. Philbrick, (1998), Raman/DIAL Technique for Ozone Measurements, *Proceeding of Nineteenth International Laser Radar Conference*, NASA/CP-1998-207671/PT1, p.407-410, 1998.

C.R. Philbrick (1998), Raman Lidar Capability to Measure Tropospheric Properties,*Proceeding of Nineteenth International Laser Radar Conference*, NASA/CP-1998-207671/PT1, p. 289-292.

M.D.,O'Brien, T. D. Stevens and C. R. Philbrick, (1996), Optical Extinction from Raman Lidar Measurements, *Optical Instruments for Weather Forecasting*, SPIE Proceedings Vol. 2832, p. 45-52.

J. Seinfeld and S. Pandis, (1997), *Atmospheric Chemistry and Physics*, John Wiley and Sons, Inc.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Energy Efficiency in an Industrial Wet Cooling Tower through Improved Control

Celso A. X. Marques[a,1]; Cristiano H. Fontes[a,2]; Marcelo Embiruçu[a,3]; Ricardo A. Kalid[a,4].

Polytechnique School of Federal University of Bahia - Industrial Engineering Programme (PEI), Rua Aristides Novis, 2, Federação, Salvador, BA, Brazil. CEP-40210-630. TEL: (55) (71) 3283-9505 FAX: + (55) (71) 3283-9800.

Abstract

Cooling towers are too much used in process plants in order to allow heat removal from the process to the atmosphere. These pieces of equipment are designed for achieve the maximum performance at design conditions. However, under other conditions such as lower heat load and lower wet bulb temperature, which frequently occur during plant operation, there is an excess of cooling capacity which is not used. This paper presents an open-loop analysis of an industrial wet cooling tower and proposes a closed-loop strategy in order to control two indexes of performance (efficiency and effectiveness). The closed-loop results suggest the possibility of a sensible improvement in the eco-efficiency of this piece of equipment thorough an expressive reduction in its energy consumption, without dropping of performance regarding the satisfaction of the thermal requirements of the process.

Keywords: Cooling towers, eco-efficiency, dynamic simulation, split-range control.

1. Introduction

Cooling towers have been neglected as an important aspect regarding to the eco-efficiency in industrial processes. The analysis and implementation of optimization and/or control techniques in cooling towers represent a potential source of improvement of energy efficiency and reduction in water consumption. Some works consider the steady state thermal performance of cooling towers and also present some inconsistencies in the results predicted by the Merkel's model (Nahavandi and Oellinger, 1977; Sutherland, 1983; Khan and Zubair, 2001). Different approaches have been proposed in modeling, optimization and performance evaluation of cooling towers. Majumdar et al. (1983) presented a two-dimensional model based on mass, energy and momentum balances considering two case studies, namely, mechanical draft and natural draft. Dreyer and Erens (1996) proposed a one-dimensional model based on aerodynamic, hydrodynamic and mass and heat transfers for the evaluation and design of splash packing geometries. Dessouky et al. (1997) presented a modified expression for calculating the effectiveness and the NTU taking into account the resistance of heat transfer in the water film. Söylemez (2004) used the criterion of effectiveness for the thermo-hydraulic performance optimization and analysis of forced draft counter flow cooling towers. Hosoz et al. (2006) applied a neural network for the prediction of the

performance of a pilot tower in terms of the following parameters: heat loss rate, water evaporation rate, and outlet water temperature and outlet conditions of the air stream.

2. Dynamic modeling, simulation and parameter estimation

The phenomenological model adopted to describe the dynamic behavior of the counter flow wet cooling tower was formulated through energy and mass balances under unsteady state conditions considering the volume element of packing (Younis et al., 1987). The equations comprise a one-dimensional model with distributed parameters and the line method was used in the dynamic systems simulation. The derivatives of position were approximated through finite differences and the time integration was carried out using the Klopfenstein's method (Ramirez, 1997). The fill packing volume was divided into 10 sections, leading to a differential algebraic system with 30 ordinary differential equations and 10 algebraic equations. An additional equation related to the energy balance of the heat exchanger which represents the process heat load was also considered in the model. Normal operation data (open-loop and unsteady state conditions) collected from a commercial cooling tower operating in a cellulose production unit located in Brazil were used to estimate the model parameters. The design specifications of this equipment are presented in Table 1.

Table 1 – Design specifications of the commercial cooling tower

Thermal capacity	132 GJ/h
Tower dimension	22.3 m × 11.95 m × 6.7 m
Fill packing height	1.34 m
Water circulation flow rate	2100 m^3/h
Inlet water temperature	45.0 °C
Outlet water temperature	30.0 °C
Number of cells	2
Air flow rate (each fan)	17124 m^3/min
Nominal fan power	102 kW

2.1. Open-loop analysis

A step response analysis was carried out through a change from the initial steady state condition in each of the following inputs: water flow, air flow, wet bulb temperature and heat load. The water and air flows represent possible manipulated variables and the other variables are disturbances produced by the environment (wet bulb temperature) and by the process (thermal demand). Two index of performance are investigated, namely, efficiency (Fisenko and Petruchik, 2004) and effectiveness (Khan and Zubair, 2001). The efficiency (η) is associated with the distance of the outlet water temperature from the wet bulb temperature, according to the equation 1:

$$\eta = \frac{T_{w,1} - T_{w,2}}{T_{w,1} - T_{wb}} \tag{1}$$

where $T_{w,1}$ and $T_{w,2}$ are the water temperatures in the inlet and outlet streams, respectively, and T_{wb} is the wet bulb temperature. The effectiveness (ε) is related to the distance of the outlet air humidity from its saturation condition evaluated at inlet water temperature, according to the equation 2:

$$\varepsilon = \frac{i_2 - i_1}{i_{asw,2} - i_1} \quad (2)$$

where i_1 and i_2 are the specific enthalpies of the air in the inlet and outlet streams, respectively, and $i_{asw,2}$ is the specific enthalpy of the saturated air, in the outlet stream, evaluated at the inlet water temperature.

As a result of the step response (±10%) in the water flow rate, a first order behavior with some differences between the gains in positive and negative step responses was observed for the effectiveness. The same disturbance applied to the air flow rate showed an inverse response in the efficiency. The efficiency presented over-damped behavior for both process heat load and wet bulb temperature disturbances. The summary of the open-loop analysis is showed on Table 2 that presents the gain signals of each input-output pair. In all cases, efficiency and effectiveness have opposite signs for each input.

Table 2 – Gain signals: open-loop analysis

	Efficiency	Effectiveness	Outlet water temperature
Air flow	+	-	-
Water flow	-	+	+
Process heat load	+	-	+
Wet bulb temperature	+	-	+

3. Cooling tower control

Cooling towers are designed taking into consideration extreme conditions of operation, namely, the highest wet bulb temperature and the maximum heat load. In general, this results in oversized cooling capacity of the installations. Therefore, in the absence of any regulatory control it is impossible to achieve satisfactory performance and high energy consumption usually results. Despite the importance of appropriate control of cooling tower operation to save energy, few works have been reported in the literature. Dijk et al. (1985) considered the use of the water flow to keep the hot water temperature at its design value together with the air flow in order to control the cold water temperature. Moreover, the set-point of the cold water stream is established according to the wet bulb temperature trajectory. Fonstad (1988) considers that only the air flow can be frequently manipulated and states that a PID controller could be used to keep the hot water temperature at the desired set-point. Chen and Liang (2005) applied fuzzy control to provide a reduction in energy consumption in frigorific equipment's cooling system. According to Stout and Leach (2002) there is less potential for energy saving in cross flow rather than counter flow cooling towers through the manipulation of the tower fan. Fisenko and Petruchik (2004) showed that it is possible to reduce the energy consumption of the fan while both air humidity and local air temperature decrease.

Here, a split-range configuration is adopted and the control strategy comprises the control of a single process output by coordinating the actions of several manipulated variables. Based on the velocity algorithm of the digital PID controller, the controller output signal is:

$$\Delta c(k) = Kc \cdot \Delta e(k) + \frac{Ts \cdot Kc}{\tau_I} \cdot e(k) + \frac{Kc \cdot \tau_D}{Ts} \cdot \Delta^2 e(k) \quad (3)$$

where $\Delta c(k)$ is the variation in the controller output signal at instant k, T_S is the sample period, $\Delta e(k)$ and $\Delta^2 e(k)$ are the error variation and the variation of the error

variation, respectively. Parameters Kc , τ_I and τ_D are the tuning parameters (controller gain, integral time constant and derivative time constant). The current value of each manipulated variable can be obtained directly through the variation $(\Delta c(k))$ of the control output signal sent to each final control element. The signal sent to the water pump variable-frequency drive (VFD) is given by:

$$\Delta c_1(k) = -\Delta L \tag{4}$$

and the signal sent to the fan VFD is obtained by the following expression:

$$\Delta c_2(k) = -\Delta c_1(k) \cdot \frac{\Delta G_{\max}}{\Delta L_{\max}} \tag{5}$$

where $\Delta G_{\max}$ and $\Delta L_{\max}$ are the water flow and the air flow ranges, respectively.

Constraints must also be considered to establish the limits of the air and water flows. The limits of the water flow were established according to the design specifications of the equipment (Table 1) and according to the data used during the parameter estimation phase (L_{min} = 1560 m^3/h and L_{max} = 2600 m^3/h). The maximum value adopted for the air flow is based on the design conditions (G_{max} = 34248 m^3/min and 25% of this value (G_{min} = 8562 m^3/min) was assumed for the minimum air flow).

3.1. Closed-loop analysis

The plant was simulated starting from its steady-state and a sequence of disturbances in the process heat load and in the wet bulb temperature was applied. The disturbance sequence of the wet bulb temperature was defined based on data from a typical day of operation. The period was 0:00 AM to 6:00 PM with sampling time of two hours. The heat load disturbance sequence was defined through an energy balance using historical data of inlet and outlet temperatures and water flow. Thus, the wet bulb temperature in the period considered ranged from 23.5 to 25.5 °C and the heat load from 65 to 95 GJ/h. The effect of these disturbances on the efficiency, effectiveness and exit water temperature were used as a base case for comparison with the closed-loop results. The closed-loop strategy comprised an override/split-range combination in order to maximize the cooling tower performance through the use of two PID controllers considering simultaneously the control of efficiency and effectiveness, variables with opposite behaviours as shown earlier. The diagram of this override/split-range strategy is presented in Figure 1.

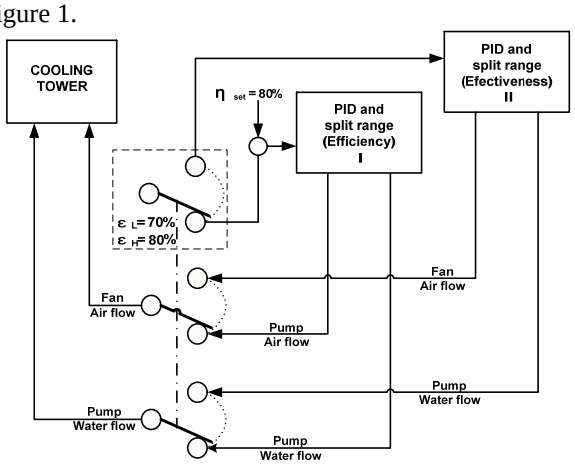


Figure 1 - Override/split-range strategy designed to control efficiency and effectiveness.

Here, whenever effectiveness drops below its minimum threshold value (low effectiveness, ε_L=70%), the control is switched to the PID controller associated to the maximization of the effectiveness. On the other hand, if the effectiveness value reaches its maximum threshold value (high effectiveness, ε_H=80%), the PID controller associated to the efficiency control is activated. Despite the fact that the air and water flows are manipulated simultaneously through a split-range strategy, the effect of the water flow is more expressive in the efficiency control and the same is valid for the air flow-effectiveness interaction. Figures 2a and 2b present the behavior of manipulated and controlled variables. Considering the infeasibility of keeping efficiency and effectiveness simultaneously at their set-points, the override/split-range strategy proposed is capable of monitoring both outputs and to assure the values are maintained within satisfactory levels.

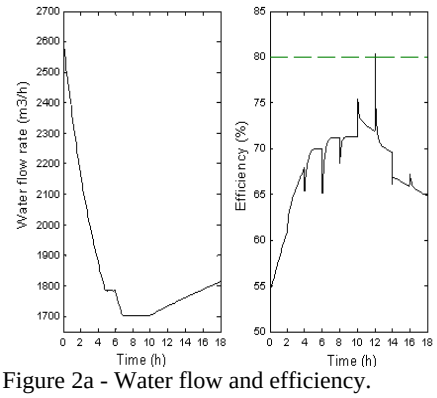


Figure 2a - Water flow and efficiency.

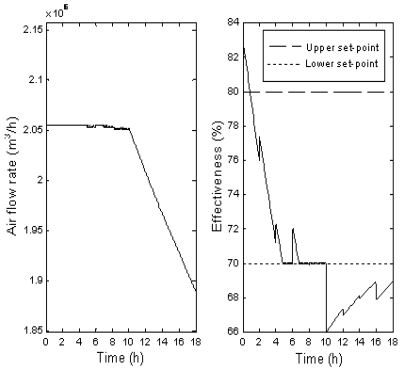


Figure 2b - Air flow and effectiveness.

The PID tuning parameters were selected through an optimal tuning method which involved the solution of an optimization problem to minimize the deviations between the controlled variable and its set-point (Alfano Neto and Embiruçu, 2000). In order to assure a workable profile for the manipulated variables based on the limits imposed by the physical components of pumps and fans, constraints were considered in the optimization problem, assuring that both water and air flow did not surpass them. A feasible performance for the manipulated variables can be verified and the control is able to cope with the disturbances to keep the controlled output close to its set-point. Moreover, it can be seen that the closed loop test presented in this work comprise a mixed control problem (regulatory and servo). The initial conditions of the controlled variables are not equal to the set-point established in each case. The potential to water and energy saving is summarized on Table 3.

Table 3 – Water and energy consumptions over 18 hours of operation

	Override/split-range strategy	Open-loop results
Water Consumption (m^3)	919.3	921.6
Energy consumption (pump) (kWh)	4332.0	11297.0
Energy consumption (fans) (kWh)	3469.8	3668.6
Total energy consumption (kWh)	9128.2	14966.0
Monetary savings (US$)	560	-

4. Conclusions

Open and closed-loop analysis of a counter flow wet cooling tower coupled with a shell and tube heat exchanger which represents the process heat load is presented. Apart from model parameter estimation using data from a commercial tower, the open loop analysis showed some particularities of the cooling tower efficiency, as first order behaviour, inverse response, nonlinearities, over-damped behaviour and opposing behaviours for the manipulated variables. The closed-loop analysis tested override/split-range strategy designed to control efficiency and effectiveness. The results demonstrate the good performance of the override/split-range strategy control in regulatory problems. Thus, potentials energy saving was estimated in US$ 250000/year. The results presented here show that significant savings can be achieved through appropriate control of certain process variables associated to the operation of a cooling tower without needing considerable investment.

References

A.N. Nahavandi, J.J. Oellinger, 1977, An improved model for the analysis of evaporative counterflow cooling towers, Nuclear Engineering and Design, v.40 (2), p.327-336.

J.W. Sutherland, 1983, Analysis of mechanical-draught counterflow air/water cooling towers, J Heat Transfer, v.105 , p.576–583.

J.R. Khan, S.M. Zubair, 2001, An improved design and rating analyses of counter flow wet cooling towers, Int. J. Heat Mass Transf, v.123, p.770-778.

A.K. Majumdar, A.K. Singhal, D.B. Spalding, 1983, Numerical modeling of wet cooling towers – Part I: Mathematical and physical models, Journal of Heat Transfer, v.105(4), p.736-743.

A.A. Dreyer, P.J. Erens, 1996, Modeling of cooling Tower splash pack, Int. J. Heat Mass Transf, v.39 (1), p.109-123.

H.T.A. El-Dessouky, A. Al-Haddad, F. Al-Juwayhel, 1997, A modified analysis of counter flow cooling towers, ASME J Heat Transf, v.119 (3), p.617–626

M.S. Söylemez, 2004, On the optimum performance of forced draft counter flow cooling towers, Energ Convers Manage,v. 45, p.2335–2341.

M. Hosoz, H.M. Ertunc, H. Bulgurcu, 2007, Performance prediction of a cooling tower using artificial neural network, Energy Conversion and Management,v. 48 (4), p.1349-1359.

M.A. Younis, M.A. Fahim, N. Wakao, 1987, Heat input-response in cooling tower-zeroth moments of temperature variations. J. Chem. Eng. Japan, v.20 (6), p.615-618.

W.F. Ramirez, 1997, Computational Methods for Process Simulation, Butterworths, 2nd edition.

S.P Fisenko, A.I. Petruchik, 2004, Towards to the control system of mechanical draft cooling tower of film type, Int. J. Heat and Mass Transfer, v.48, p.31-35.

J.C Van Dijk, C.G. Van Poll, P. Poldervaart, 1985, Investment in cooling tower control pays big dividends, Process Engineering, v.66 (10), p.57-60.

K.A. Fonstad, 1988, Energy saving through control of cooling tower air flow, Energy Engineering, v.85 (3), p.4-16.

M.-L. Chen, H.-Y. Liang, 2005, Application of Fuzzy Control in Cooling System's Save Energy Design, Int.Journal of Power and Energy Systems, v.465, p.340-344.

M.R. Stout Jr, J.W. Leach, 2003, Cooling tower fan control for energy efficiency, Energy Engineering, v.99 (1), p.7-31.

C. Alfano Neto, M. Embiruçu, 2000, Tuning of PID Controllers: an Optimization-Based Method, In Digital Control 2000: Past, Present and Future of PID Control, p.415-420.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Industrial Experience in the Deployment of Real Time Online Energy Management Systems

Diego Ruiz [a], Carlos A. Ruiz [b]

Soteica Europe, S.L.; Av. Parc nº 2; E-08940 Cornellà (Barcelona), Spain;
Tel: +34 93 375 3503; http://www.soteica.eu
[a] *diego.ruiz@soteica.com*
[b] *carlos.ruiz@soteica.com*

Abstract

This paper presents real industrial examples in which the whole utilities system of a production Site (i.e., steam, fuels, boiler feed water and electricity) is optimized with a real time online, industrially well established software.
Experiences gained during more than 20 years of industrial projects deployed worldwide are commented. Main project steps are explained and critical details to be taken into account to assure successful use and proper technology transfer are presented.
The optimization objective is the overall utilities system cost reduction and takes into account the constraints associated with the existing equipment, fuels and electricity pricing and contracts, including emissions limits, quotas and rights. The energy management system models are executed and optimized at a scheduled frequency, fed with online, real time data, flowing into and out the program using the standard OPC protocol.
Besides the optimization, Key Performance Indicators (KPIs) are also calculated and sent back to the Site Plant Information System or DCSs for Operations and Management use.
Application examples and results corresponding to projects implemented worldwide in refineries and chemical plants are presented and commented.

Keywords: energy costs reduction, real optimization, energy management

1. Energy Management Systems

Modern industrial facilities operate complex and inter-related power and utilities systems. Tighter and increasingly restrictive regulations related to emissions are also imposing constraints and adding complexity to their management. Deregulated electric and fuels markets with varying contracted prices (seasonal or daily) also introduce additional challenges.
Production Departments usually have the responsibility for the operation of the facility power system but, although Operators are instructed to minimize energy usage and usually tend to do it, the power system is seen as a subsidiary provider of the utilities needed to accomplish the production target, whatever it takes to generate it.
Industrial facilities like Refineries and Petrochemicals are becoming increasingly aware that power systems need to be optimally managed because any energy reduction that Operations accomplish in the producing Units could eventually be wasted if the overall power and utilities system cost is not properly managed.
In parallel, process engineers have always attempted to develop some kind of tool, many times spreadsheet based, to improve the way utilities systems are operated. The

evolution from plant information scattered throughout many islands of automation to unified and centralized Plant Information Systems was a clear breakthrough for such work. The long term, facility wide Plant Information System based historians constitute what is known as an *enabling technology*, because they became the cornerstone on top of which many other applications are built.

Process engineers used the internal optimizers or solvers provided with the spreadsheet software in an attempt to optimize their systems. One of the authors went through a very similar path when he was a young process engineer at a Petrochemical Complex. After many years of exposure to a lot of manufacturing sites worldwide, he found that almost all Process Engineering Departments had in use an internal spreadsheet with which several process engineers worked when on duty at the power house or utilities unit. The spreadsheets wildly evolved during the years and became extremely complex and "fragile" as hundreds or thousands of tags were added and, very often, without any documentation. Raw tag data becomes usually contaminated with unexpected problems and sometimes hard to identify and filter errors. The use of such a tools for real time optimization was seldom a real success.

1.1. Real time, Intelligent Systems

It was becoming more and more clear that a certain kind of *intelligence* should be added to those energy management tools in order to produce good results in a consistent way, dealing with real time information potential errors and maintained evergreen and usable for long periods, with minimal engineering effort.

The authors found that, for the Process Industry, the definition of an *intelligent system* is generic and not very well defined. The industry usually calls *intelligent* to any piece of software that helps to automate the decision making process, efficiently controls a complex process, is able to predict properties of products or process variables, alerts to prevent hazardous situations or, in last instance, optimize process or business economics. For the practical engineers, the definition of an intelligent system is factual, not methodological. The above mentioned systems comply, up to certain extent, with one of the classical definitions of intelligence (N. Wiener, 1948): the intelligent behavior is a consequence of certain feedback mechanisms, based on the acquisition and processing of information to accomplish with a certain objective.

A coherent engineering environment providing all the needed tools within a single shell was a real need. During the past 20 years, Visual MESA optimization software evolved from the earlier text based, offline application of the 1980's to an online, real time, graphical user interfaced, highly sophisticated *intelligent* system. It is today considered as the industry standard Energy Management System (EMS) real time online optimizer (D. Nelson et al., 2000). It has been widely implemented in the processing industry and it is applied routinely to reduce the cost of operating the energy systems at power, chemical, petrochemical, and refinery plants worldwide.

1.2. Energy Management System Description

A detailed model of the energy system (fuels, steam, electricity, boiler feed water and condensates network) is built within the EMS graphical user interface and it is continuously fed with validated, real time data. It includes all the actual constraints of the site and decision variables for their operation. Optimization is configured to minimize the total energy cost. The model writes back its results to the Real Time Data Base (Plant Information System). It also provides reliable data that helps to audit the energy productions and usages within the site energy system, and in that way wastes can be detected and eliminated. Greenhouse emissions are also taken into account.

1.3. Real Time, Online Systems

1.3.1. Manual, Open Loop versus Closed Loop Operation

Although the Operators still need to close the loop manually, EMS projects have proven to save substantial amounts of money with very fast pay-back. Even more economic benefits can be obtained if some of the manual optimization handles are automated under a closed loop scheme (D. Uztürk et al., 2006).

There are several and important additional advantages of using EMS's for online, real time, closed loop optimization because it increases the benefits already obtained in open loop, especially when fuels and power prices are market driven and highly variable. Several implementations of this kind have already been performed (Wellons et al., 1994 and D. Uztürk et al., 2006).

1.3.2. Online capabilities

The online capabilities are a relevant portion of the software structure and key to a successful closed loop implementation. A proper software tool should provide standard features right out of the box. Therefore, it should not require any special task or project activity to enable the software to easily interact and cope with real time online data. The EMS based models are created from scratch acquiring and relying on real time online data. A standard OPC based (OLE for process control) protocol interface has been provided to perform a smooth and easy communication with the appropriate data sources, such as a distributed control system (DCS), a plant information system, a historian or a real time database.

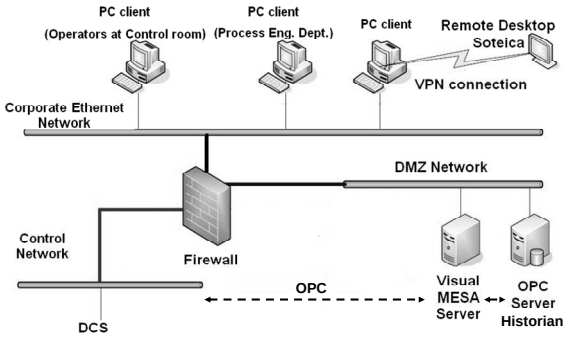


Figure 1: Installation Architecture for Closed Loop Implementation

Fig. 1 shows the typical installation architecture for closed loop real time optimization, including the proper network security layers and devices, for example firewalls and demilitarized zones (DMZ) domains.

1.4. Optimization Variables and Constraints Configuration for Closed Loop Optimization

Building a model that realistically represents the utilities and energy system topology, needs to include all the optimization variables and constraints and, at the same time, include all the system economic details, especially the fuels and electricity contractual complexity.

Optimization variables are those where some freedom exists regarding what value they might take. For example, the steam production rate at which a particular boiler operates is a free choice as long as the total steam production is satisfied, thus the most efficient boiler's production can be maximized. They can be continuous or discrete.

Constrained variables are those variables that cannot be freely chosen by the optimizer but must be limited for practical operation. Some of them are direct equipment

constraints, but a few can be defined as abstract constraints, where the variables are not directly measured in the system or they are not a related of a particular of equipment.

2. Project Activities

An Energy Management System (EMS) Implementation project is executed in 9 to 12 months. The main project steps are discussed below.

2.1. Required Information

After the Purchase Order is issued, a document would be submitted to the Site with all the informational requirements for the EMS project sent it to the project owner. By project owner we understand a Site engineer who, acting as a single interface, will provide the needed information and coordinate all the project steps.

2.2. Kick-Off Meeting

Prior to the Kick-Off Meeting, the provided information will be reviewed to have a better understanding of the Site facilities and process. Additional questions or clarifications would be sent to the Site regarding particular issues, as required. During the week of the on-site Kick-Off Meeting, all information would be reviewed with the Site staff, and additional information required for building the model would be requested, as needed. At that time, the optimization strategy would also be discussed.

2.3. EMS Software Installation

The software is then configured and licensed on the EMS server PC. It would also be connected to the OPC server. Remote access to the model would also need to be made available at this time and would need to be available throughout the rest of the project.

2.4. Functional Design Specification

With the information provided during the Kick-Off meeting, a Functional Design Specification document would be prepared, revised by both parties in concert, and then approved by the Site.

2.5. Visual MESA Model Building and Optimization Configuration

During this stage, the model and the report are built working remotely on the EMS server. The model grows with access to online real time data. The second trip to the facility would occur during this stage and would be used for mid-term review of the model and optimization. Upon model approval, a month-long testing period would commence. The EMS would run routinely, but optimization recommendations would still not be implemented by the operations staff. A base line could be obtained based on the cost reduction predicted by the optimizer during this period, in order to compare with the full implementation of the suggestions at the end of the project.

2.6. Optimization Startup

Site engineers would then train the operations staff to use Visual MESA and to implement the recommendations. Continuing in this period, operations staff would begin implementation of the optimization recommendations. Project developing staff would return to the Site facility a third time to review implementation of the optimization recommendations and make any final adjustments to the model, as required. Project documentation would be provided and a benefits report would be submitted.

3. Key Performance Indicators (KPI's)

In addition to the real time online optimization, during the EMS project appropriate energy performance metrics can also be identified and performance targets can be set.

Those metrics are usually known as Key Performance Indicators (KPI's) and can be related to:

- High level KPI's that monitor site performance and geared toward use by site and corporate management. For example: Total cost or the utilities system, predicted benefits, main steam headers imbalances, emissions, etc.
- Unit level KPI's that monitor individual unit performance and are geared toward use by unit management and technical specialists. For example: plant or area costs, boilers and heaters efficiencies, etc.
- Energy Influencing Variables (EIV's) that are geared towards use by operators. For example: Equipment specific operation parameters, like reflux rate, transfer line temperatures, cooling water temperature, etc.

They are calculated by the EMS software and written back to the Plant Information System.

4. Project Examples

The authors implemented several dozens of projects, all of them currently in service around the world. Some of them are described in D. Ruiz et al., 2005, 2006 and 2007, S. Benedicto et al., 2007, J. M. García Casas et al., 2007 and M. Kihn, et al 2008.
Two examples will be commented; the first one is an open loop implementation. The second belongs to a closed loop system.

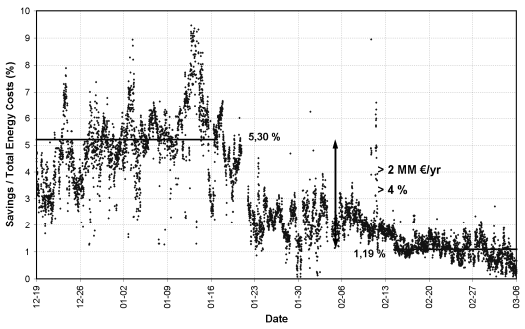


Figure 3. Energy cost reduction evolution by using an online energy management tool

4.1. *First Example*

The first example corresponds to the energy system of a Spanish refinery plus an olefins unit (D. Ruiz et al., 2006). In order to accurately evaluate the economic benefits obtained with the use of this tool, the following real time test has been done:

- *First month:* Base line, The EMS being executed online, predicting the potential benefits but no optimization actions are taken.
- *Second month:* Operators trained and optimization suggestions are gradually implemented.
- *Third month:* Optimization recommendations are followed on a daily basis.

Fig. 3 shows the results of this test. Over that period, in 2003, 4% of the energy bill of the Site was reduced, with estimated savings of more than 2 million €/year.

4.2. Second Example

The second example corresponds to a Dutch refinery where the EMS online optimization runs in closed loop, the so-called energy real time optimizer (D. Uztürk et al., 2006).

Typical optimization handles include letdowns, load boilers steam flow, gas turbine generators/steam turbine generators power, natural gas intake, gas turbine heat recovery, steam generators duct firing, extraction of dual outlet turbines, deaerator pressure, motor/turbine switches, etc.
Included constraints are the steam balances at each pressure level, boiler firing capacities, fuel network constraints, refinery emissions (SO_2, NO_x, etc.) and contract constraints (for both fuel and electric power sell/purchase contracts).
Benefits are reported to come from the continuous load allocation optimization between boilers, optimised extraction/condensing ratio of the dual outlet turbines, optimised mix of discretionary fuel sales/purchase, optimised gas turbine power as a function of fuel and electricity purchase contract complexities (trade off between fuel contract verses electricity contract penalties).

5. Future Perspectives

Although wide opportunities still exist for a growing number of real time online Energy Management Systems executed in open loop, an increased number of Closed Loop applications are expected in the near future. This evolution will bring additional economic benefits to the existing user base, especially when fuels and power prices are market-driven and highly variable. High frequency optimization opportunities that cannot be practically addressed by manual operating procedures would be captured and materialized. Of critical importance is having a robust and mature solver that reliably converges in a reasonable period of time in order to ensure buy-in from Operations for the continuous use of the system.
More focus on key process side operations, when tightly related with the Energy Network, will be also necessary. In addition to the Refining and Petrochemical industries, who were the early adopters of this kind of technology, other industries will take advantage of the real time energy management systems as well.

References

N. Wiener, 1948, Cybernetics, or Control and Communication in the Animal and the Machine, 2nd ed. of 1961, Cambridge, Mass.: MIT Press

M.C. Wellons, A.V. Sapre, A. I. Chang, T. L. Laird, 1994, On-line Power Plant Optimization Improves Texas Refiner's Bottom Line, Oil & Gas Journal, May 16

D. Nelson, G. Roseme, S. Delk, 2000, Using Visual MESA to Optimize Refinery Steam Systems, AIChE Spring Meeting, Session T9013, Georgia, USA

D. Ruiz, C. Ruiz, J. Mamprin, Depto. de Energías y Efluentes Petronor, 2005, Auditing and control of energy costs in a large refinery by using an on line tool, European Refining Technology Conference (ERTC) Asset Maximisation, Budapest, Hungary

D. Ruiz, C. Ruiz, D. Nelson, G. Roseme, M. Lázaro, M. Sartaguda, 2006, Reducing refinery energy costs, Petroleum Technology Quarterly (PTQ), Q1 2006, Pages 103-105

S. Benedicto, B. Garrote, D. Ruiz, J. Mamprin, C. Ruiz, 2007, Online energy management, Petroleum Technology Quarterly (PTQ), Q1 2007, Pages 131-138

J. M. García Casas, M. Kihn, D. Ruiz, C. Ruiz, 2007, The Use of an On-line model for Energy Site-Wide Costs Minimisation, , European Refining Technology Conference (ERTC) Asset Maximisation Conference, Rome, Italy

D. Uztürk, H. D. Franklin, J. M. Righi, A. T. Georgiou, 2006, Energy System Real Time Optimization, NPRA Plant Automation and Decision Support Conference, Phoenix, USA

D. Ruiz, C. Ruiz, D. Nelson, 2007, Online Energy Management, Hydrocarbon Engineering, September, Pages 60-68.

M. Kihn, D. Ruiz, C. Ruiz, A. García Nogales, 2008, Online Energy Costs Optimizer at Petrochemical Plant, Hydrocarbon Engineering, May, Pages 119-123

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimizing control action online using a neural model and the solver of an electronic worksheet

Tatiana L. Fujiki[a], Márcia R. G. Sanzovo, Manuela S. Leite, Flavio V. Silva, Ana M. F. Fileti[a]

[a] *Department of Chemical Systems Engineering, School of Chemical Engineering - University of Campinas - Av. Albert Einstein, n. 500, P. O. Box 6066, 13083-970, Campinas - SP, Brazil.*

Abstract

The worldwide market requirements for high standard products and safe and environmentally friendly processes enforce chemical industries to look for controllers that can afford nonlinearities and transient behavior of chemical processes. Intelligent systems based on Artificial Neural Network (ANN) have been studied as alternative solutions for ill-defined plants or nonlinear and transient systems. In this work, the algebraic equations of a neural model were implemented in a Microsoft Excel worksheet. As a case-study, this alternative controller was successfully implemented to maintain the bulk temperature of a bromelain precipitation tank at 5°C.

Keywords: artificial neural network, Fieldbus interface, digital process control, Microsoft Excel Solver.

1. Introduction

A 2-year industrial investigation into the application of ANNs in process control was summarized in 2001 by Lennox et al. Among their studies, the required accuracy in a polymer quality control system was successfully achieved, by integrating an ANN within a model based control algorithm. The controller was implemented online by interfacing an IBM personal computer with a 286 processor to the existing data acquisition and control system. Gadkar et al. (2005) proposed the application of neural networks for monitoring the state variables of an alcoholic fermentation tank. Through a weight updating algorithm, the neural model was implemented online and the comparative results showed that the error generated by the model with weights adaptation was smaller than that obtained by the model without weights adaptation. Effectiveness of neural networks was also investigated by Fileti et al. in 2006, during the online implementation of an inverse neural model to automatically control a basic oxygen steel-making (BOS) process. An increase in productivity was observed when applying the neural model to predict end-point temperature and carbon percentage in the steel at the end of the batch period. Gonzaga et al. (2009) used a feed forward ANN to develop a soft-sensor that estimated online the polyethylene terephthalate (PET) viscosity, for application in a control system. The integration between the developed soft-sensor and the supervisory system SETCIM (AspenTech) was performed using scientific programming languages, such as FORTRAN and C. As can be seen, ANNs have been largely employed for nonlinear process modeling because of their ease of construction and implementation (since the need for mechanistic modeling from first principles is suppressed). They have been proved effective in modeling and predicting

process variables. The use of a Microsoft Excel worksheet for online implementation instead of other commercial software or computational programming is proposed in this work due to its friendly interface and applicability.

2. Materials and Methods

2.1. Neural Model

Using the Neural Network Toolbox of MATLAB 7.0, a dynamic neural model was developed so that its optimized parameters could be determined. The input layer consisted of the current values (k) of the measured variables of the process. One sampling time was considered between the instants k and k+1 used in the model. In the hidden layer, an activation function was applied to twice the number of variables of the input layer. Another activation function was applied to the output layer, which predicts the one-step-ahead controlled variable, PV_{k+1}. The Levenberg-Marquardt algorithm was employed to train the multilayered feed forward network and the training procedure was performed using open-loop runs. The open-loop data set was obtained by gathering a wide range of values of the input layer variables, including the whole network action domain. The process dynamics was observed through an open-loop run with the manipulated variable at a fixed point. From there, the step disturbances in the manipulated variable, *MV*, were planned so that the controlled variable behavior could be monitored from several runs and this database was employed to train the neural network. The database was split in two sets: 75% and 25% for training and tests, respectively. Furthermore, closed-loop runs with a fuzzy controller were also used for the tests, since the neural model was expected to have good response in closed-loop, with the implementation of the optimizer for the process control. The neural model performance was assessed through dispersion plots of the testing runs, with a desirable result being represented by a slope coefficient of the linear fitting of the dispersion plots (network output *versus* target vector) close to the unity and an interception coefficient around zero. Following the offline tests, online validation was performed, in order to verify the prediction capacity of the model to be implemented.

2.2. Using an Electronic Worksheet as a Controller

The optimized weights and biases were inserted in an electronic worksheet (Microsoft Excel) to reproduce the algebraic equations of a neural model. The input vector with the process data was normalized, according to Equation 2.1, using the maximum and minimum values of the variables:

$$pn = 2\frac{(p - \min\ p)}{(\max\ p - \min\ p)} - 1 \qquad \textit{Equation 2.1}$$

The normalized signal from the input layer was multiplied by the weights between input and hidden neurons and each result was added to the correspondent bias. This number, called *n*, was then processed in the hidden layer according to the hyperbolic tangent activation function (Equation 2.2). Similarly, the output layer processed the signal from the hidden layer, *a*. In the output neuron a simple linear activation function was used.

$$a = \tan sig(n) = \frac{2}{(1+\exp(-2n))} - 1 \qquad \textit{Equation 2.2}$$

The computed output value was then denormalized, by isolating p in Equation 2.1, in order to be compared to its set-point. The quadratic error of the one-step-ahead controlled variable (PV_{k+1}) relative to its set-point was defined as the objective function to be minimized by the solver (Equation 2.3). The solution to be implemented in the plant was found by means of the *quasi*-Newton method of generalized reduced gradient, available in Excel, by changing the manipulated variable value, *MV*.

$$\min_{MV} \left(PV_{(k+1)} - PV_{sp}\right)^2 \qquad \textit{Equation 2.3}$$

Subject to the constraints of Equations 2.4, 2.5 and 2.6:

$$MV \geq MV_{\min} \qquad \textit{Equation 2.4}$$

$$MV \leq MV_{\max} \qquad \textit{Equation 2.5}$$

$$\left|MV_{solver} - MV_{k-1}\right| \leq \text{maximum step allowed} \qquad \textit{Equation 2.6}$$

2.3. Case-Study

Bromelain is an enzyme found in pineapple, even at its mature stages. It can be recovered through precipitation with ethanol. However, this solvent may cause protein denaturation, thus the precipitation should be carried out at low temperatures, reducing molecule flexibility and hence, solvent penetration (SCOPES, 1994). The experiments were carried out in a stirred fed-batch jacketed tank. The instrumentation diagram and the pilot-plant are presented in Figure 2.1.

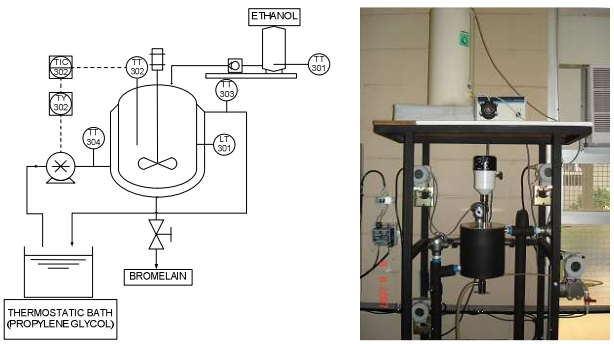


Figure 2.1: Instrumentation diagram and prototype of the bromelain precipitation process.

Intelligent instruments such as temperature transmitters were used for monitoring the process variables. The data acquisition and process control were performed by means of a digital Fieldbus network architecture (bus with spurs topology), configured by Syscon Software 7.0 (Smar). For the experimental tests, 150 mL samples of pineapple juice

were fed into the tank and the ethanol (99.5%) at room temperature was continuously fed by a micropump, until the liquid volume reached 750 mL. The coolant (propylene glycol) flow rate was manipulated through a variable speed pump. Syscon and Excel communicated with the supervisory software (Indusoft Web Studio) through Ole for Process Control (OPC) protocol.

3. Results and Discussion

3.1. Neural Model

The operating variables for the input layer, measured every 4 seconds (sample time), were chosen as follows: ethanol temperature ($T_{alc,k}$) has great influence in the precipitation temperature – the higher the ethanol temperature, the higher the overshoot obtained; coolant inlet ($T_{in,k}$) and outlet ($T_{out,k}$) temperatures provide information on the heat exchange in the tank jacket; since coolant pump speed (MV_k) is the manipulated variable, its value will determine the controlled variable response; the liquid level (L_k), represented by the liquid volume, provides the ANN with information on the run time, thus distinguishing equal input vectors that correspond to different output vectors; pump speed variation (ΔMV_k) indicates to the ANN which step disturbance in pump speed caused the given output; since bulk temperature ($T_{bulk,k}$) variation determines the manipulated variable action, it works as a reference for the ANN prediction of the one-step-ahead bulk temperature ($T_{bulk,k+1}$), that was chosen as the output variable of the ANN. In the hidden layer a hyperbolic tangent activation function was applied to 14 nodes. The structure of the neural model developed is shown in Figure 3.1.

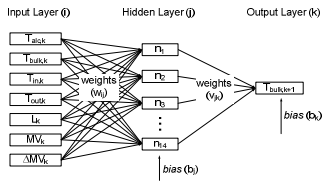
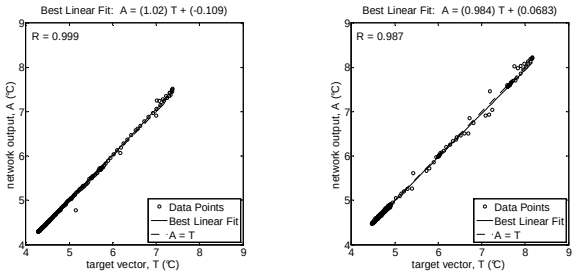


Figure 3.1: Topology of the neural model

In Figure 3.2 the results for the open-loop and the closed-loop offline tests of the neural model in MATLAB are presented.



(a) (b)

Figure 3.2: Dispersion plot of the network output and the target vector for an open-loop run (a) and for a closed-loop run (b), for the test of the neural model.

Offline tests with unseen data proved that the ANN successfully predicted the tank temperature, as can be seen by the agreement between real and predicted values of temperature. Both linear fits presented slope coefficients close to the unity and interception coefficients around zero, approaching to the ideal curve.

3.2. Online Validation and Closed-Loop Runs

At this stage, the bulk temperature was monitored during a closed-loop run with the PID controller ($K_c = 35\%/°C$, $\tau_I = 28s$ and $\tau_D = 7s$). Simultaneously, the neural model predicted the one-step-ahead bulk temperature. At the end of the experiment both temperatures were compared and the result is presented in Figure 3.2 (validation curves). The results showed that the ANN was capable of learning the inherent nonlinearities and also successfully predicted the bulk temperature in online mode, thus being considered suitable for the application with the solver. The neural model coupled with the Microsoft Excel Solver was used as a temperature controller and the experimental results are shown in Figure 3.2 (alternative controller curve). The manipulated variable, to be changed by the solver was subject to the following constraints: a range of 0 to 100% of speed variation was set, in function of its own actual limitation (Equations 2.4 and 2.5). Also, the manipulated variable action was smoothed by restricting the value of its step (the difference between the solver solution and the present control action) to 35%, based on the PID controller gain. Preliminary tests showed that these constraints were not sufficient to keep the bulk temperature around set-point: cooling outperformed heating after 200 s. To prevent the controlled variable from leaving the training range, an additional constraint was added: the pump was turned off when the bulk temperature would reach 4.9ºC.

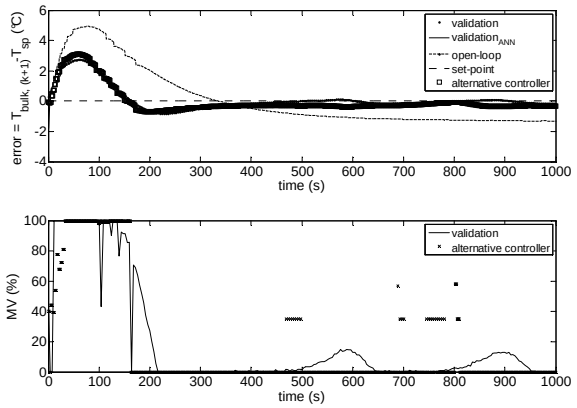


Figure 3.2: Results from the online validation and from the closed-loop run, in comparison with the open-loop run.

In Figure 3.2, it is shown that the developed alternative controller was able to maintain the controlled variable around set-point ($5°C$), with good temperature prediction. The overshoot observed in the first 200 seconds is due to the low liquid level in the tank,

which caused the heat exchange area to be at a minimum. It can also be explained by the dissolution heat produced during the ethanol addition.
In Table 3.1, a numerical comparison is presented:

Table 3.1: Comparison between PID and alternative controllers and open-loop run.

Controller	Overshoot (°C)	Rise time (s)	ITAE	Energy consumption (kJ)
PID	2.76	164	157843.7	143
Alternative	3.10	156	166716.0	121
Open-loop	4.93	334	597343.5	215

It can be noted that all the analyzed parameters for both PID and alternative controllers were smaller than the parameters obtained for the open-loop run (pump speed fixed at 40%), showing the need for process control. The low overshoot is important to avoid protein denaturation by the solvent. Smaller rise time values allow the process to be controlled faster and the lower energy consumption translates into financial savings.

4. Conclusions

This work regarded the development of an alternative controller for a semi-batch process, using artificial neural networks modeling coupled with a solver of an electronic worksheet. The neural model predicted the controlled variable with good accuracy and it was successfully used along with the solver to maintain the bulk temperature around set-point, with low energy consumption. The developed controller gathered the benefits of the artificial neural networks in affording nonlinearities and the accessibility of an electronic worksheet, making this methodology a promising new way to face complex process control problems, without spending efforts unnecessarily in mathematical modeling. Broadening the range of applicability of this method to a wide range of industrial processes is a very relevant topic of research, since industries of different fields can apply this solution to many of their problems.

References

A. M. F. Fileti, T. A. Pacianotto, A. P. Cunha, 2006, Neural modeling helps the BOS process to achieve aimed end-point conditions in liquid steel, Engineering Applications of Artificial Intelligence, v. 19, p. 9-17.

K. G. Gadkar, S. Mehra, J. Gomes, January 2005, On-line adaptation of neural networks for bioprocess control, Computers and Chemical Engineering, v.29, p. 1047-1057.

J. C. B. Gonzaga, L. A. C. Meleiro, C. Kiang, R. Maciel Filho, 2009, ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process, Computers and Chemical Engineering, v. 33, p. 43-49.

B. Lennox, G. A. Montague, A. M. Frith, C. Gent, V. Bevan, 2001, Industrial application of neural networks – an investigation, Journal of Process Control, v. 11, p. 497-507.

R. K. Scopes, 1994, Protein Purification: Principles and Practical, 2.ed, New York: Springer Verlag, 329 p.

Acknowledgments

The authors would like to thank National Counsel of Technological and Scientific Development (CNPq) for the financial support.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Ash Deposits Monitoring in a Convective Heat Transfer Section of a Kraft Recovery Boiler

Gustavo M. de Almeida[a], Song W. Park[b]

[a]*Federal University of Sao Joao del-Rei*
Rod. MG 443, Km 07, Fazenda do Cadete, 36.420-000, Ouro Branco, MG, Brazil
[b]*Department of Chemical Engineering, Polytechnic School, University of Sao Paulo*
Av. Prof. Luciano Gualberto, 380/Tv. 3, Butanta, 05.508-900, Sao Paulo, SP, Brazil

Abstract

A reliable monitoring of abnormal situations is still a challenge in industrial chemical processes. The earlier its detection and diagnosis, the greater the possibility of mitigating losses. An alternative approach to carry out these tasks is to make use of signal processing tools. In this direction, the hidden Markov model (HMM) method is used to monitor the accumulation of ash deposits along a convective heat transfer section of a boiler belonging to a Kraft pulping mill in Brazil. The identified model was able to provide the current state of the process working as a warning in case of a coming abnormal event.

Keywords: Chemical process monitoring, Signal processing tool, Hidden Markov model, Kraft pulping mill, Industrial data analysis

1. Introduction

Process monitoring actions in chemical industries are to guarantee operating, economic, safety and/or environmental aspects. The main idea, after detecting and diagnosing an abnormal situation, is to take actions to recover the process to a normal operating condition mitigating this way the effect of potential losses [1-2]. Since abnormal situations are in general incipient, deviations from normal conditions are smaller at the beginning, and hence reaching early detection without the support of a computer-based system is practically unfeasible. Most of the applications are normally based on residue metrics, and an alternative approach for accomplishing monitoring tasks is to make use of signal processing tools [3]. This work applies the hidden Markov model (HMM) technique to monitor the operation of a heat transfer section of a boiler. Therefore, this data-driven technique, which belongs to the signal processing field, constitutes an other way to develop chemical process monitoring systems. The case study is based on a chemical recovery boiler from a Kraft pulping mill in Brazil. Studies concerning the application of HMMs in the process monitoring area from computer-simulated cases are those by [4-13].

2. Hidden Markov Model (HMM)

Every chemical process is under random influences due to an inherent variability, and hence measurements of process variables may be seen as realizations of an underlying stochastic process. Then, normal operating conditions can be described by particular probability distributions, which fail in case of changes in process conditions [14]. This is the motivation for applying the hidden Markov model (HMM), a signal processing

tool, to the chemical process monitoring area, once it is capable of identifying changes of statistical nature in signals (composed by measurements of process variables in this case) over time. Such nature is useful in order to deal with both noise and disturbances. The successful applications of HMMs are in the speech processing field, including both speech recognition and speaker verification [15].

2.1. Fault Detection Tasks with HMMs

The goal of hidden Markov models is to model sequential data. Fig. 1 shows its input-output relation, in which the input is an observation sequence of T symbols (discrete case) or real vectors of same size (continuous case) ($O = \{o_1,o_2,\ldots,o_t,\ldots,o_T\}$), and the output is a likelihood value ($-log[P(O|\lambda)]$), a measure of the capacity of the model (λ) in generating O [15]. For example, a HMM representing an operating state in particular can be used to detect changes from it over time, which is characterized by more and more lower likelihood values.

$O=\{o_1,o_2,\ldots,o_t,\ldots,o_T\}$ ⟶ HMM (λ) ⟶ $-log[P(O|\lambda)]$

Fig. 1. Input-output relation for HMMs: observation sequence (O) and likelihood value ($-log[P(O|\lambda)]$), respectively.

2.2. Mathematical Formulation

Hidden Markov models are a doubly stochastic process, in which the former is responsible for the state-transitions ($P(q_t|q_{t-1})$), which obeys the Markov property, whereas the latter concerns the observation-emissions ($P(o_t|q_t)$). The difference between Markov- and hidden Markov-models refers to this second process, that is deterministic to the first. This explains the *hidden* term in HMMs once the underlying Markov chain is not directly observable. Table 1 shows the three parameters of discrete HMMs, namely π, A, and B, where M_D is the number of distinct observation symbols of the emission probability distributions (one per state), and N is the size of the discrete state space. (A compact notation is given by λ, i.e. $\lambda = (\pi,A,B)$.) For the continuous case (used in this work), the B matrix is replaced by probability density functions, whose usual representation is a finite mixture of Gaussians, where o_t is the observation vector at time t, M_C is the number of mixture components per state, and c_{jk} is the mixture component (subjected to stochastic constraints), μ_{jk} is the mean vector, and Σ_{jk} is the covariance matrix, for the kth mixture component in the state j (cf. Table 1). The parameters π and A are the same as in the discrete case [15].

Table 1. Elements of discrete and continuous HMMs.

Description	Parameter	Formulation
State-transition probability distribution	$\boldsymbol{A} = \{a_{ij}\}$	$a_{ij} = P(q_{t+1} = j \mid q_t = i)\ ,\ 1 \le i, j \le N$
Observation-emission probability distribution	Discrete case (DHMM) ($\boldsymbol{B} = \{b_j(k)\}$)	$b_j(k) = P(\boldsymbol{o}_1 = \boldsymbol{v}_k \mid q_t = j)\ ,\ 1 \le k \le M_D\ ,\ 1 \le j \le N$
	Continuous case (CHMM)	$b_j(\boldsymbol{o}_t) = \sum_{k=1}^{M_C} c_{jk} \boldsymbol{N}(\boldsymbol{o}_t, \boldsymbol{\mu}_{jk}, \boldsymbol{\Sigma}_{jk})\ ,\ 1 \le j \le N$
Initial state probability distribution	$\boldsymbol{\pi} = \{\pi_i\}$	$\boldsymbol{\pi}_i = P(q_1 = i)\ ,\ 1 \le j \le N$

3. Case Study

This study uses a chemical recovery boiler from a Kraft pulping mill in Brazil. One of its goals is to produce steam for electric power generation and heat transfer operations. Its fuel is the residual liquor originating from a previous stage. This equipment has two regions: a furnace, where the liquor combustion and the recovery of specific inorganic compounds occur, and a heat transfer section, as in power boilers, co-responsible for the steam production. This region contains three heat exchangers in series, namely super-heater (SH), convector (Conv), and economizer (Eco), as in Fig. 2. A risk situation refers to the accumulation of ash deposits, composed for sulfur- and sodium-based salts mostly sodium sulfate (Na_2SO_4), over tubes of the heat exchangers. This is mainly caused by sodium-based salts, vapor sodium, fuel droplet residues, and unburned carbon particles, which are carried in the fuel gas line. The consequences are: reduction of the efficiency in generating steam, lower disponibility of the boiler due to the blockage of the fuel gas path, and greater maintenance costs. The established practice in mills to manage it is the use of sootblowers [16-17]. Then, this study applies the hidden Markov modeling to monitor the increase of ash deposits over tubes of heat exchangers of a heat transfer section of a boiler. The considered system also includes the eletrostatic precipitator (cf. Fig. 2), an equipment used to recover the sodium and sulfur in the fuel gas line. The monitored variable, which is strongly correlated to the increase of ash deposits, is the pressure drop (ΔP) of the gas along the system.

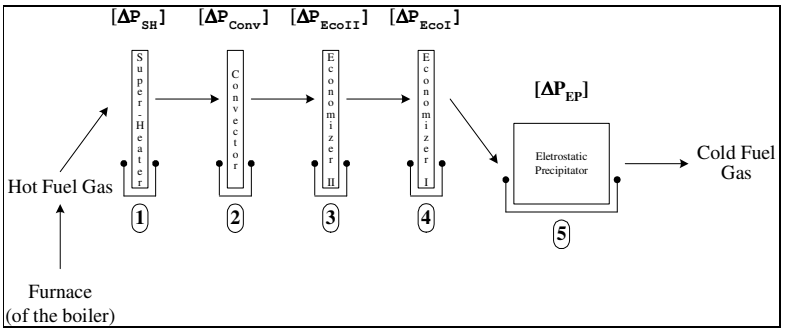


Fig. 2. Scheme for the convective heat transfer section and the eletrostatic precipitator of the boiler, with the five measurements of the pressure drop of the fuel gas.

3.1. Data Set

The data set comprehends ten months of operation, with a sampling interval of one minute. Table 2 presents the collected measurements of the pressure drops of the fuel gas along the heat transfer section until the eletrostatic precipitator (cf. Fig. 2). Other variables were collected in order to verify the operating conditions during this period.

Table 2. Measurements collected in the chemical recovery boiler.

Description	Code	Unit
Pressure drop of the fuel gas through the …		
1. Super-heater	ΔP_{SH}	Pa
2. Convector	ΔP_{Conv}	Pa
3. Economizer II	ΔP_{EcoII}	Pa
4. Economizer I	ΔP_{EcoI}	Pa
5. Eletrostatic Precipitator	ΔP_{EP}	kPa

4. Methodology

4.1. Model Identification Step

The goal is to identify a representative HMM for the boiler under a critical operating condition concerning the pressure drop of the fuel gas along the heat transfer section, which is given by ΔP_{EcoII} (the pressure drop of the fuel gas through the second economizer) ≥ 1000.0 Pa. A continuous HMM is used in this work (cf. Table 1) and each observation sequence is composed of 5 real vectors of size 1, i.e. $O = \{\Delta P_{SH}, \Delta P_{Conv}, \Delta P_{EcoII}, \Delta P_{EcoI}, \Delta P_{EP}\}$ (cf. Fig. 1). Every 1 minute (the sampling interval), a probability of a sequence of being generated by the reference model ($-log[P(O|\lambda)]$) is plotted in the trend plot of the likelihood function, providing information about the current state of the system. Initially, using a training subset, plenty of models are generated by varying the number of mixture components per state (M_C) from 1 up to 3, the number of states (N) between 2 and 3, and the model topology between ergodic and left-to-right. The model selection is based on the likelihood function being selected the one with the highest mean value calculated onto the validation subset. This steps makes use of the classical Baum-Welch algorithm, based on the Maximum Likelihood Estimation (MLE) principle [18].

4.2. Control Limits Construction Step

The calculus of the lower control limits (LCL) is based on the pressure drop of the fuel gas through the second economizer (ΔP_{EcoII}). This is accomplished by splitting its range of pressure drop into five intervals, namely [600 700[, [700 800[, [800 850[, [850 900[, and [900 950[, in Pa units. For each interval, with a set of representative observation sequences, the lowest likelihood value of an observation sequence in particular becomes the lower control limit to such interval. The model used to this determination is that previously identified. The classical Forward algorithm is used to calculate the model outputs [15].

4.3. Test Step

Finally, in order to verify the performance of the model, it is fed with another subset of observation sequences, simulating an on-line monitoring action. The idea is to get the behavior of the model when subjected to real operating conditions. As the boiler operates in a continually, a trend plot for the likelihood function (the model output) arises over time, which constitutes a source of information about the state of the equipment. As before, this step uses the Forward algorithm [15].

5. Results and Discussion

5.1. HMM Identification Step

The identification subset contains 2880 observation sequences being 1920 (≈ 66.7%) for training and 960 for validation. The selected HMM, characteristic of an operation under a critical pressure drop of the fuel gas along the heat transfer section of the boiler, among the 12 candidate models, has 1 mixture component per state (M_C), 3 states (N), and an ergodic topology.

5.2. Control Limits Construction Step

Table 3 shows the lower control limits (LCL) for all five intervals (cf. Subsection 4.2), embracing the entire range of pressure drops of the fuel gas through the second economizer (ΔP_{EcoII}). For example, a likelihood value for an observation sequence equal to –36.1 means that it is already possible to have ΔP_{EcoII} = 850.0 Pa. It can be observed that higher the pressure drop measurement greater the likelihood value.

Table 3. Lower Control Limits (LCL).

Data Set	Lower Control Limits: A likelihood value ($-log[P(O\|\lambda)]$) equal to …	… means a possibility of already having a value for ΔP_{EcoII} equal to ...	Number of Observation Sequences
1	–3.2	950.0	377
2	–19.6	900.0	529
3	–36.1	850.0	1529
4	–51.8	800.0	1095
5	–110.3	700.0	389
6	–160.3	600.0	185

5.3. Test Step

The test subset contains 4712 observation sequences. First of all, Fig. 3(a) shows the strong correlation ($\rho = 0.985$) between the likelihood values calculated by the model onto the test subset and the measurements of pressure drop of the fuel gas through the 2nd economizer (ΔP_{EcoII}). The correlation between ΔP_{EcoII} and the measurements of pressure drop of the fuel gas through both the convector (ΔP_{Conv}) and the 1st economizer (ΔP_{EcoI}), which are the most critical heat exchangers with regard to the accumulation of ash deposits [17], is also close to 1. Hence, higher the likelihood value greater the measurement of the pressure drop of the fuel gas. (The pressure drop through the superheater (ΔP_{SH}) is negligible.) Fig. 3(b) shows the result of brushing the likelihood values for all observation sequences for which $\Delta P_{EcoII} \geq 700.0$ Pa. It can be verified that the lowest value (–77.8) is greater than the respective lower control limit (LCL) equal to −110.3, which means that if the model output exceeds such limit it is already possible to have a pressure drop of the fuel gas through the second economizer (ΔP_{EcoII}) above 700.0 Pa. This information works as a warning about the state of the system concerning the accumulation of ash deposits. The results (not shown) are the same to all intervals for ΔP_{EcoII} (cf. Table 3), which attests that the monitoring of the accumulation of ash deposits along heat transfer sections of chemical recovery boilers using a HMM–based monitoring system is very promising. (For implementation, the sensibility concerning the correlation between likelihood and pressure drop values needs to be adjusted.)

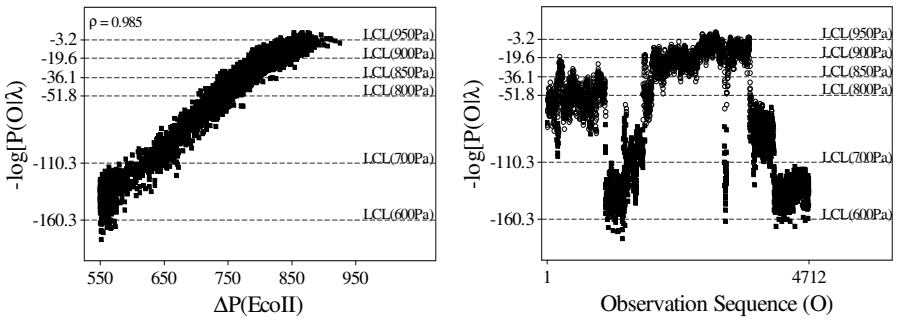


Fig. 3. (a) Linear correlation between the likelihood values calculated by the model (λ) onto the test subset ($-log[P(O|\lambda)]$) and the measurements of pressure drop of the fuel gas through the second economizer (ΔP_{EcoII}), and (b) resulting trend plot for the likelihood function, where the values associated with observation sequences (O) for which $\Delta P_{EcoII} \geq 700.0$ Pa are highlighted by circles (○).

6. Final Considerations

A hidden Markov model (HMM), representative of high measurements of pressure drop of a fuel gas, is used to monitor the accumulation of ash deposits along a heat transfer section of a boiler. It was verified that the resulting trend plot for the likelihood function (the output of the model), which is highly correlated with the measurements of pressure drop, works as a source of information since it is able to provide the current state, and hence the tendency of the operations. In brief, HMM–based systems may work as decision support tools for monitoring actions in chemical processes.

Acknowledgments

The authors thank the pulp mill, for the cession of the data base, and the Fundacao de Amparo a Pesquisa do Estado de Sao Paulo (FAPESP), under grant 03/00405-9, for the financial support.

References

[1] L.H. Chiang, E.L. Russel, R.D. Braatz, 2001, Fault detection and diagnosis in industrial systems, Springer, London.
[2] I. Nimmo, 1995, Adequately address abnormal operations, Chem. Eng. Prog., 91, 9, 36-45.
[3] R.J. Patton, J. Korbicz, S. Lesecq, 2006, Preface, Control Eng. Practice, 14, 6, 575-576.
[4] G.M. Almeida, S.W. Park, 2008, Fault detection and diagnosis in the DAMADICS benchmark actuator system – A hidden Markov model approach, Proc. of the 17th World Cong. of The Int. Fed. of Automatic Control (IFAC), Korea, 12419-12424.
[5] G.M. Almeida, S.W. Park, 2008, Process monitoring in chemical industries – A hidden Markov model approach, Proc. of the 18th European Symp. on Comp. Aided Proc. Eng., France, 25, 1-6.
[6] G.M. Almeida, S.W. Park, 2005, Fault detection in a sugar evaporation process using hidden Markov models, Proc. of the Int. Symp. on Advanced Control of Ind. Proc., Korea, 309-313.
[7] A. Bakhtazad, A. Palazoglu, J.A. Romagnoli, 2000, Detection and classification of abnormal process situations using multidimensional wavelet domain hidden Markov trees, Comp. and Chem. Eng., 24, 2-7, 769-775.
[8] J. Chen, W. Chang, 2005, Applying wavelet–based hidden Markov tree to enhancing performance of process monitoring, Chem. Eng. Sc., 60, 18, 5129-5143.
[9] W. Sun, A. Palazoglu, J.A. Romagnoli, 2003, Detecting abnormal process trends by wavelet-domain hidden Markov models, AIChE Journal, 49, 1, 140-150.
[10] F. Tokatli, A. Cinar, 2004, Fault detection and diagnosis in a food pasteurization process with hidden Markov models, The Can. J. of Chem. Eng., 82, 6, 1252-1262.
[11] J.C. Wong, K.A. McDonald, A. Palazoglu, 2001, Classification of abnormal plant operation using multiple process variable trends, J. of Proc. Control, 11, 4, 409-418.
[12] J.C. Wong, K.A. McDonald, A. Palazoglu, 1998, Classification of process trends based on fuzzified symbolic representation and hidden Markov models, J. of Proc. Control, 8, 5-6, 395-408.
[13] X. Yangsheng, G. Ming, 2004, Hidden Markov model-based process monitoring system, J. of Int. Man., 15, 3, 337-350.
[14] V. Venkatasubramanian, R. Rengaswamy, S.N. Kavuri, S.N., K. Yin, 2003, A review of process fault diagnosis – Part III: Process history based methods, Comp. and Chem. Eng., 27, 3, 327-346.
[15] L. Rabiner, 1989, A tutorial on hidden Markov models and selected applications in speech recognition, Proc. of the IEEE, 77, 2, 257-286.
[16] T.N. Adams, W.J. Frederick, T.M. Grace, M. Hupa, K. Iisa, A.K. Jones, H. Tran, 1997, Kraft Recovery Boilers, Tappi Press, Atlanta.
[17] E.K. Vakkilainen, 2005, Kraft recovery boilers: Principles and practice, Valopaino Oy, Helsink.
[18] L.E. Baum, T. Petrie, G. Soules, N. Weiss, 1970, A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains, Ann. Math. Stat., 41, 1, 164-171.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Evaluation of Control Structures for a Debutanizer Column

Lilian R. Canabarro,[a] Mario C. M. M. Campos,[b] Enrique L. Lima[a]

[a]*Programa de Engenharia Química/COPPE, Universidade Federal do Rio de Janeiro, Cidade Universitária, CP 68502, Rio de Janeiro, RJ, Brazil 21945-970*
[b]*PETROBRAS S.A. Av. Horacio Macedo, 950, Rio de Janeiro RJ, Brazil 21941-915*

Abstract

In this work is described the choice of the control structure for a specific industrial distillation column presenting a significant discrepancy between the inventories of the top and bottom sections. The studied distillation column showed to be an excellent research field, but looking for the balance between scientific development and industrial application it is always important to take into consideration the variable "time invested". In this context, and searching for a satisfactory control structure, static analysis tools, as RGA and SVD, had been used in the characterization of a set of alternatives structures based on the primary composition control variables L, V, D and B, as well as some relations between them. With the reduction of the "time invested" kept in mind, it was used a rigorous dynamic simulator capable to deal with thermodynamic calculations for complex multicomponent mixtures, as those found in the processing units of crude oil and its fractions. Results have demonstrated the importance of rigorous dynamic simulation to solve design control problems in distillation columns which frequently presents unique characteristics and should be treated accordingly.

Keywords: Distillation control, control structure selection, dynamic simulation

1. Introduction

The interest on distillation control systems started in the decade of the 50s, significantly increasing along the time, with special emphasis in the 80s. One conclusion that can be drawn from the huge amount of technical and scientific literature produced in this period is that distillation is a difficult system to understand. The scientific research is continuously improving our knowledge on the subject, but the practice is still confronted with serious difficulties, because each distillation column seems to be different from other ones and, in such a way, requires a personalized treatment. Recently Skogestad (2007) wrote "Many books...and papers, including several of my own..., have been written on the merits of the various configurations, but it is probably safe to say that the importance of the choice of configuration (level control scheme) has been overemphasized". These words are a good example of the knowledge (scientific or general) evolution process, as they clearly express the permanent and necessary substitution of old "truths" by new ones. Certainly this subject deserves much more deep thoughts, but the idea here is only to show that in complex cases, as certainly can be the distillation columns, many of the "general truths" that had been enunciated had not been confirmed in practice, requiring more specific approaches. In his contribution Skogestad (2007) proposed a distillation column control methodology with reasonable general application, bringing great contribution for the practical side of the problem. The focus of this work is on the control structure design for a specific debutanizer

column using simple and practical tools, complemented with rigorous dynamic simulation.

The Problem

The studied system is a 26 trays naphtha stabilizer column of a crude atmospheric distillation unit. The feed, coming from the top of the atmospheric tower and introduced in tray 13th (from the top), contains 30 components, including 2% of light components (methane, ethane, propane, butanes, pentanes, etc.), water and 20 pseudo-components. The bottom product of the column is used to preheat this feed. Because it is almost totally in the liquid phase, vapor and liquid flows in the stripping section are larger than in the rectifying section, and as a consequence the thermal load to the reboiler is larger than to the condenser. There are three products in the system: combustible gas and liquefied petroleum gas (LPG) in the vapor and liquid streams that leave a partial condenser; and stabilized naphtha in the bottom stream. The control objective is to obtain a top product with the highest purity, guaranteeing the LPG quality, without reducing the quality of the stabilized naphtha. As this is produced in a large amount, significant quality variations are not expected, so the top composition control was considered a priority.

Technical-scientific literature shows that the Relative Gain Array (RGA) introduced by Bristol (1966) is still an important analysis and design tool for multivariate systems. Initially developed to determine static interaction in systems, allowing appropriate pairing between input and output variables in a control structure, its area of applicability was significantly extended later, both for static and dynamic conditions (Grossdidier et al., 1985).

Considering the stationary case and using the system gain matrix K, the RGA matrix Λ can be obtained as follows,

$$\Lambda = \left[\lambda_{ij}\right] = \left[\underline{\underline{K}}_{ij} \cdot \left[\left(\underline{\underline{K}}^{-1}\right)^{T}\right]_{j}\right]$$

where λ_{ij} is the ratio of the static gain of the output variable i related to the input variable j, with all the other pairs in open loop, to the static gain of the same pair (i,j), with all the other pairs under perfect control. A RGA matrix close to identity indicates small interaction, determining the appropriated pairing. Another tool, sometimes complementary to the RGA, results from Singular Value Decomposition, SVD (Grossdidier et al., l985). Also considering the static case and starting from the gain matrix K, the SVD is defined as

$$K = U\Sigma V^{T}$$

where U and V are unitary matrices, and Σ is a diagonal matrix with the singular values in descending order of magnitude. The ratio of the largest to the smallest singular value defines the Condition Number, γ, a measure of the difficulty to control the associated system. As this number is dependent on variables units (scaling) it can take very different values for the same system. The correct way to take advantage of the information supplied by the condition number is to use the scaling that produces the minimum value, γ^* (scaling doesn't change system's characteristics). Research literature clearly shows limitations when evaluating a system operating in dynamic

condition using static information. However, it is important to notice that in many practical cases it is difficult to get a reasonable dynamic description of the system, limiting the use of complex elaborated theoretical methods.
An approach for calculating the gain matrices of different alternative control structures from the knowledge of only one aroused great interest in the end of the 80s and early 90s. The method introduced by Häggblom and Waller (1988) was based on transformations and relations of consistency derived from global and component mass balances. This approach seemed to be quite useful for simplifying the calculation of RGA and SDV of different control structures.
Considering only the top and bottom distillation products quality control, there are 4 primary manipulated variables available: reflux (L), distillate (D), bottom (B) and vapor (V) flowrates. From these primary variables the literature reports a large number of combinations that generates secondary variables, allowing the design of a large number of control structures. Disregarding any one of them could be a difficult task, requiring significant testing work. This is facilitated by combining the use of appropriate indices, such as RGA and γ (there are many others, with varying complexity), system knowledge and computer simulation.

2. Implementation

A set of control structures was chosen for evaluation disregarding some literature recommendations (sometimes divergent), because they are commonly based on results obtained from very simplified systems. Due to the large bottom flowrate the popular LB structure was not considered. The structures chosen were: LV, DV and (L/D, V/B). A fourth structure was also chosen where, as suggested by Skogestad (2007), a fast tray temperature loop was include in the stripping section. In this structure, identified as LV-cascade, the top composition is controlled by a cascade control system, manipulating a pump's rotor speed, where the internal loop controls the reflux. The bottom composition is controlled by another cascade control system, manipulating V, where the internal loop controls the 19th tray temperature. Table 1 shows the values of K, RGA, γ e γ^* for each control structure.

Table 1 Static gains, RGA and conditional numbers

Structure	K	Λ	γ	γ^*
LV	$\begin{bmatrix} 1.023 & -0.5806 \\ 3.88e^{-6} & -1.12e^{-3} \end{bmatrix}$	$\begin{bmatrix} 1.0018 & -0.0018 \\ -0.0018 & 1.0018 \end{bmatrix}$	1259.5	1.0892
DV	$\begin{bmatrix} -2.12 & -0.6014 \\ -6.86e^{-6} & -1.12e^{-3} \end{bmatrix}$	$\begin{bmatrix} 1.0017 & -0.0017 \\ -0.0017 & 1.0017 \end{bmatrix}$	2048.7	1.087
L/D V/B	$\begin{bmatrix} 3.70e^{-2} & -5.6202 \\ 9.45e^{-8} & -1.08e^{-2} \end{bmatrix}$	$\begin{bmatrix} 1.0013 & -0.0013 \\ -0.0013 & 1.0013 \end{bmatrix}$	79155	1.0757
LV-Cascade	$\begin{bmatrix} 2.1790 & -1.11e^{-2} \\ 8.37e^{-6} & -2.68e^{-6} \end{bmatrix}$	$\begin{bmatrix} 1.0162 & -0.0162 \\ -0.0162 & 1.0162 \end{bmatrix}$	826230	1.2887

The RGA and γ^* values indicate that the different structures should present similar behavior, in terms of insignificant interaction and good conditioning (although system scaling is weak). From this information it is possible to choose good variable pairing

and to have a feeling of the control viability of the different structures, but it is not enough to define which structure is the best. To this end dynamic simulation (gPROMS®, Multiflash®) was used, comparing the behavior of the different structures when the system is perturbed by the most common disturbances in distillation columns: changes in feed flowrate, temperature and composition. Before presenting the simulation results, it is important to shortly comment about the experience we have with the transformations method (Häggblom e Waller, 1988). Using the gain matrix of the LV structure the method was used in order to calculate the gain matrices of the other ones. As can be seen in Table 2, there was no consistence in the obtained results.

Table 2 RGA by transformations from LV

Structure	**Λ**
L/D V/B	$\begin{bmatrix} -2.4442 & 3.4442 \\ 3.4442 & -2.4442 \end{bmatrix}$
DV	$\begin{bmatrix} 1.0002 & -0.0002 \\ -0.0002 & 1.0002 \end{bmatrix}$

Although the result obtained for DV was similar to the one obtained from data specifically acquired for this structure (Table 1), the result obtained for (L/D,V/B) was completely different. This could be an indication that the assumptions on which the method is based are not satisfied in this case. As in practice this disagreement can be quite common, it may justify the little impact of the method observed in research literature after some initial popularity. The papers studying the transformation method were certainly of great value to improve the knowledge on distillation columns, but their use in practical problems requires a careful preliminary analysis. Some old "truths" must probably be re-evaluated to be put in their real dimension.

Searching for the best control structure for the studied column the behavior of each structure was simulated against feed flowrate (+/- 10%), temperature (+/- 1%; the system is highly sensitive to this disturbance) and composition (+/- 10%) disturbances. Controllers were tuned in such way as to get the best performance for each structure. Results in Figure 1 were obtained for the top of the column against the positive flowrate and composition disturbances, and the negative temperature disturbance. In addition to the controlled variables responses, the behavior of primary manipulated variables is also shown. Figure 2 displays similar results for the bottom of the column.

Considering the top (Figure 1), where composition control was more rigorous, it can be notice that there are close relations between the disturbances effects on the composition. The effect of 10% variation in feed flowrate was almost equivalent to -1% variation in feed temperature, while the effect of +10% variation in feed composition was less significant and in the opposite direction. The differences observed in the magnitude of the initial reflux flowrate have been attributed to the larger vapor amount required by the addition of the temperature control loop in the LV-cascade structure. But the fact that after 1% reduction in the feed temperature all the structures lead to the same reflux flowrate value still requires justification. There are large performance differences between the best (LV-cascade and [L/D,V/B]) and the worse (DV and LV) structures. Observing the light top mass fraction responses it can be verified that the best structures succeed in keeping the setpoint, with a minor advantage for the LV-cascade structure.

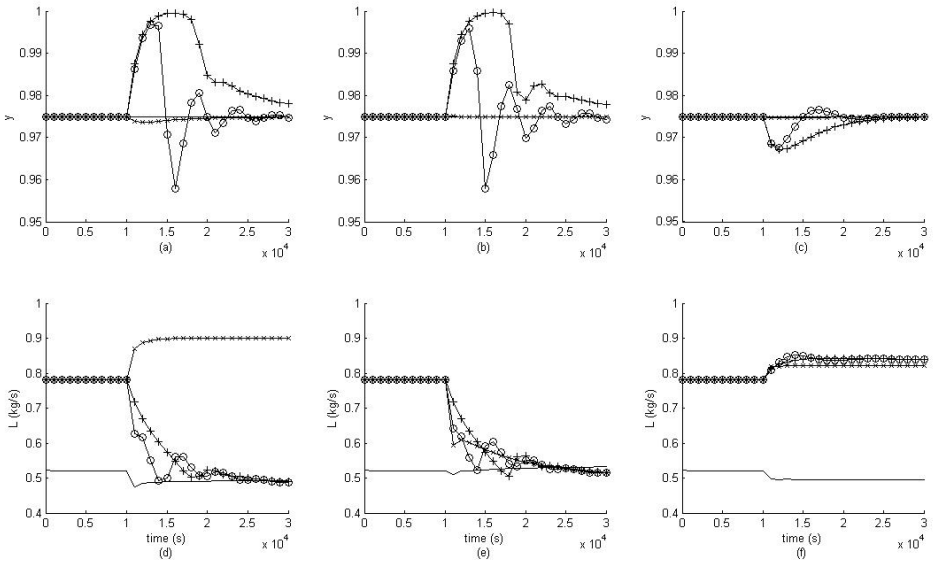


Figure 1. LV (+), DV (o), [L/D,V/B] (x) and LV-cascade (-) responses for feed flowrate (a, d), temperature (b, e) and composition (c, f) disturbances; sum of distillate light components mass fractions (a, b, c), reflux mass flowrate (d, e, f).

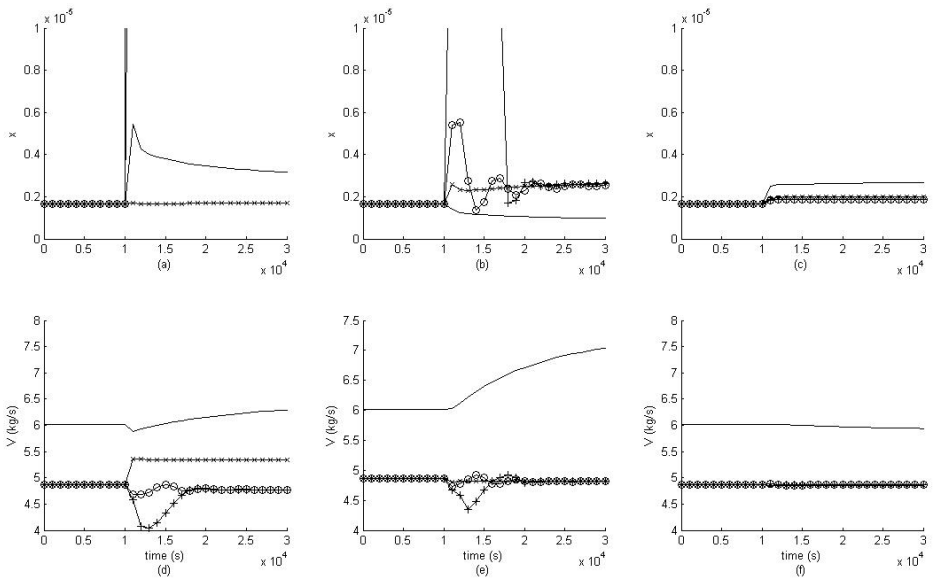


Figure 2. LV (+), DV(o), [L/D,V/B] (x) and LV-cascade (-) responses for feed flowrate (a, d), temperature (b, e) and composition (c, f) disturbances; sum of bottom's light components mass fractions (zoom) (a, b, c) and vapor mass flowrate coming from the reboiler (d, e, f).

The control efforts showed by changes in manipulate variables are almost equal, with a notable exception in the case of the (L/D, V/B) structure submitted to a disturbance in F, were the manipulated variable presents an inverse behavior. There is no simple explanation to this fact, but it is important to remind that the complex multicomponent feed, with small amounts of light component, could produce complex thermodynamic effects. Considering the bottom, where the tuning was less rigid, it can still be observed the best performance of LV-cascade and (L/D,V/B), but now it is not possible to distinguish any clear difference between them. It should be note that in this case disturbances effects are very small due to non-significant light components concentration.

3. Conclusions

The information obtained from simple static indexes was not sufficient to determine a best control structure. RGA results agreed with Shinskey (1984), apparently showing that the (L/D, V/B) structure has better decoupling characteristics than DV, and this than LV. Although there is not enough room for a detailed presentation, it is worth mentioning that setpoint top composition changes produced larger interaction effects in the structure (L/D, V/B) than in DV, in opposition to the previous statement. Moreover, the best performance during the simulation tests was obtained for the LV-cascade structure, whose index γ^* was the worst. It is interesting to note that, for the studied case, the positive characteristics of the "general" structure proposed by Skogestad (2007) were confirmed. The main conclusion that can be extracted form this work is that every effort should be made to have accurate dynamic simulators to aid the design of distillation columns control structures.

Acknowledgment

The authors are grateful to CNPq and Petrobras for their support.

References

E. H. Bristol, 1966, On a New Measure of Interaction for Multivariável Process Control, IEEE Trans. Auto. Con., AC-11, pp.133-134.

P. Grossdidier, M. Morari, B. R. Holt, 1985, Closed-Loop Properties from Steady-State Gain Information, Ind. Eng. Chem. Fundam., 24, pp. 221-235.

K. E. Häggblom, K. V. Waller, 1988, Transformations and Consistency Relations for Distillation Control Structures, AIChE Journal, 34, 1634-1648.

F. G. Shinskey, 1984, Distillation Control, McGraw-Hill, New York.

S. Skogestad, 2007, The Dos and Don´ts of Distillation Column Control, Trans IChemE, Part A, 85(A1), pp. 13-23.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Development of multivariate statistical-based tools for monitoring of sour water unit

Douglas F. B. Lima[a], Fernando A. Zanella[a], Alexandre C. Teixeira[a], Luiz F. L. Luz Jr.[b], Carlos A. U. Gontarski[b], Enrico M. Gomes[b], Saulo H. Chiquitto[b], Marcelo K. Lenzi[b]

[a]*Refinaria Presidente Getúlio Vargas REPAR/PETROBRAS – Setor de Otimização Rodovia do Xisto, BR476, Km 16, Araucária – PR, 83700-970, Brazil*
e-mail: dfalleiros@petrobras.com.br
[b]*Departamento de Engenharia Química, Universidade Federal do Paraná Centro Politécnico, Jardim das Américas, Caixa Postal 19011, Curitiba – PR, 81531-980, Brazil*

Abstract

High production combined with high product quality and low environmental damage and energy consumption represents one of the main challenges of any chemical plant. In petrochemical plants, sour water treatment unit plays a key role regarding environmental issues, due to the necessity of ammonia and sulphur-based compounds removal. Although fundamental-based models can be successfully used for monitoring tasks, faster models based on multivariate statistics theory have been widely introduced in industrial sites. They present attractive features such as the capacity of handling a high amount of data, contain the process history and can be used for on-line and real-time process monitoring. In this study, multivariate statistical based methods were used to develop an alternative tool for an actual operating sour water treatment unit monitoring. Input variables collected from the process history were used to not only adequately predict the behavior of a given response variable, but also another potential use of this is to gain understanding of how and to which extent different process variables affect a specific response variable. Some issues like the length of predictive horizon, data structure, data selection and treatment during the development of the tool are also discussed. The obtained results indicate that multivariate techniques were able to monitor the unit variables behavior, representing a key alternative tool for sour water unit monitoring.

Keywords: sour water, multivariate statistics, process monitoring.

1. Introduction

High production combined with high product quality and low environmental damage and energy consumption represents one of the main challenges of any chemical plant. In petrochemical plants, sour water treatment unit plays a key role regarding environmental issues, due to the necessity of ammonia and sulphur-based compounds removal (MANDAL, 2007).

Ammonia and sulphur-based compounds, with minor amounts of other chemicals such as cyanides, are the main constituents of sour gas streams. The treatment

of sour gas represents a challenging issue (GAI *et al.*, 2008). One of the most common design strategies for sour water purification units is the use of two stripping columns. The first is used for sulphur-based (mainly H_2S) removal. The second column is used for ammonia removal (HARDISON, 1988). Further details regarding process design peculiarities, operating conditions and optimization issues can be found elsewhere (van HOORN, 2002).

Although fundamental-based models can be successfully used for monitoring tasks, faster models based on multivariate statistics theory have been widely introduced in industrial sites (HARRIS *et al.*, 1999). They present attractive features such as the capacity of handling a high amount of data, contain the process history and can be used for on-line and real-time process monitoring. According to MILETIC *et al.* (2004) multivariate statistics have been used for a wide range of industrial applications such as a sulfite pulp digester plant, desulfurization process, among other.

Several multivariate statistical methods are reported in literature, for example, principal component analysis (PCA), partial least squares (PLS). Further details concerning the methods, such as advantages, formulation, and application, can be found elsewhere. BRERETON (2003), OTTO (1999), JOHNSON & WICHERN (2002).

In this study, multivariate statistical methods were used to develop an alternative tool for an actual operating sour water treatment unit monitoring. Input variables collected from the process history were used to not only adequately predict the behavior of a given response variable, but also another potential use of this is to gain understanding of how and to which extent different process variables affect a specific response variable. Some issues like the length of predictive horizon, data structure, data selection and treatment during the development of the tool are also discussed. The obtained results indicate that multivariate techniques were able to monitor the unit variables behavior, representing a key alternative tool for sour water unit monitoring.

2. Methodology

Data from REPAR process plant information (PI) of three complete years were collected and analyzed. The analysis comprised statistical tests already reported in the literature (HIMMEMLBLAU, 1970) to obtain data sets which represent process steady state operation. After a steady state operation period was found, the values as they are were used in the analysis and no outlier test was conducted. Linear regression analysis was used in order to obtain different empirical models in order to predict important process variables. Further details of the linear regression analysis and parameter estimation methods can be found in the literature. (JOHNSON & WICHERN (2002)). The models have a basic structure given by:

$$\textbf{(dependent variable)} = \textbf{constant} + (\textbf{coefficient} \cdot \textbf{independent variable})_1 + \ldots + (\textbf{coefficient} \cdot \textbf{independent variable})_N.$$

Temperature measurements were obtained with a YTA-110 YOKOGAWA temperature transmitter equipped with a thermocouple of type K. Pressure measurements were obtained with EJA-430A YOKOGAWA pressure transmitter. Finally flow rates were obtained with 3051 EMERSON device. This work focused the first stripping column of the sour water unit of REPAR. For each case studied, the model which generated the predicted values is presented. Different independent variables were tested and the best obtained results are reported here. Plots of predicted values of the desired variable was against the real (observed) values are also presented. These plots were used to obtain the presented correlation coefficient. Results are

normalized and units are omitted for presentation. It must be stressed that these models are used to check the obtained value of the process sensor. Data horizon of 3 years was used. Data were treated using moving average data filter in order to eliminate mean less information. After filtering, operation steady states were identified/recognized by exploring mean and variance of collected data in a given horizon, which could range from one to five hours, using techniques reported in the literature (MaCAULEY, 2003). The linear least square technique used is a straightforward technique, representing an important tool for seeking correlation among process variables.

3. Results and Discussion

The first model can be used for prediction of the pressure at the top of the stripping column (dependent variable). The model equation is given by a linear combination of the variables presented in Table 01.

Table 01 – Variables of model 01

Independent Variable	Estimated Coefficient
Steam temperature at the exit of the column reboiler.	0.01436
Temperature at the top of the column	0.002021
Column Pressure Drop	-0.18614
Ambient Temperature	0.008012
Constant	6.209758

Figure 01 presents a plot of predicted versus real values. The correlation coefficient between predicted and real values of the pressure is 0.85. It can be seen from the picture that model predictions are reasonable for values over 0.5 up to 1. The value of the correlation bellow 0.85 probably occurs due to the experimental below 0.5 which were not omitted in the analysis.

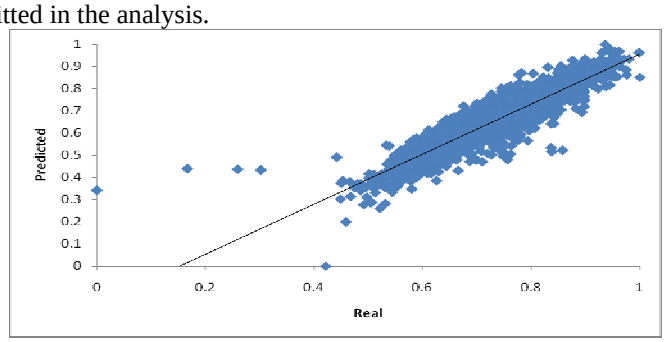


Figure 01 – Predicted versus real values of the dependent variable – model 01

It must be emphasized that not outlier test was conducted and better prediction can be obtained by either increasing the number of model parameters or changing the independent variables. Although linear, the used model did not took into account variables interactions as they would only increase the amount of parameters and did not significantly improve the correlation coefficient, showing that from a statistical point of view they are not relevant.

As a consequence, a second model can be used for prediction of the pressure at the top of the stripping column, leading to better predictions when compared to the first model. Different independent variables were chosen for this second model. The model equation is given by a linear combination of the variables presented in Table 02. The obtained correlation coefficient was of 0.85. Although the correlation coefficient was

roughly the same, this model is better because the range from 0 to 1 of predicted values lie over the straight line, differently from the previous model that values ranging from 0 to 0.5 cannot be adequately predicted.

Table 02 – Variables of model 02

Independent Variable	Estimated Coefficient
Condenser pressure	0.280782
Ambient Temperature	0.007457
Column feed stream temperature	0.004443
Temperature of cold inlet heat exchanger fluid	0.002483
Constant	5.60264

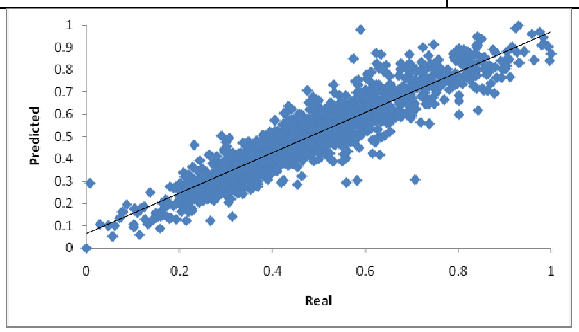


Figure 02 – Predicted versus real values of the dependent variable – model 02

Another important variable is the amount of live steam need for column operation. Our model presented a correlation coefficient of 0.83, which is a reasonable value. Table 03 presents the variables used in model. Predicted against real values are shown in Figure 03. The model present the data tendency as predicted values increase as does observed values. On the other hand, the chart presents three cluster data regions. This probably happened due to either a nonlinear relation among these variables or a strong interaction among the variables.

Table 03 – Variables of model 03

Independent Variable	Estimated Coefficient
Pressure at the top of the column	1.997464
Column feed rate	0.083669
Temperature at the top of the column	0.009882
Column Pressure Drop	2.186367
Constant	-17.412

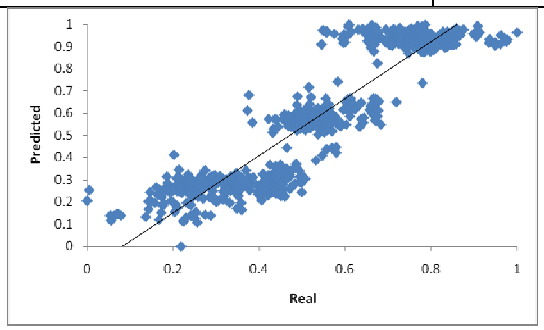


Figure 03 – Predicted versus real values of the dependent variable – model 04

The models previously presented are used to predict variables which are continuously measured. However, one of the aims of this work is the development of models to accurately predicted variables which are not continuously measured such as the NH_3 concentration at the bottom of the column. Table 04 presents the independent variables used for model derivation.

Table 04 – Variables of model 04

Independent Variable	Estimated Coefficient
NH_3 concentration in the column feed stream	0.927587
Column feed rate	0.083669
(Column feed rate)·(0.001·(NH_3 concentration in the column feed stream)) / (column bottom exit rate)	-200.267
Constant	634.2305

A poor correlation coefficient has been achieved 0.75. However it should be stressed that outlier data was not removed in order to keep all possible real situation observed during plat steady state operation. Besides, only a linear model on the parameters was worked. The use of nonlinear models on the parameters may improve the correlation coefficient. On the other hand, it can be seen that the model predicted values present the same tendency of the real data. Figure 04 presents predicted values against real data.

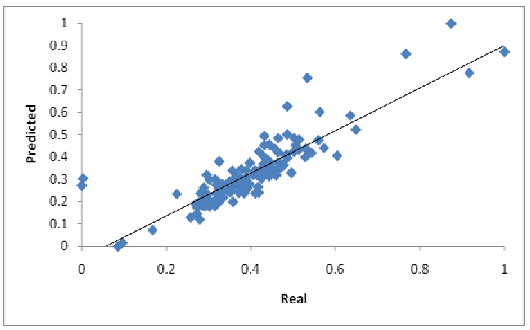


Figure 04 – Predicted versus real values of the dependent variable – model 04

Finally, the column H_2S recovery could be successfully modeled by a linear model. Table 05 presents the independent variables and the estimated coefficients.

Table 05 – Variables of model 05

Independent Variable	Estimated Coefficient
NH_3 concentration at the top of the column	33.3646
(NH_3 concentration at the top of the column)^2	-4.16765
Temperature at the top of the column	-0.15596
Temperature variation at the column reboiler	-0.36381
Temperature variation at the column	-0.1567
Constant	-38.5063

This model could successfully describe the column H_2S recovery as a correlation coefficient of roughly 0.94 has been achieved. Besides, the plotted values are randomly distributed on both sides of the line indicating that no bias is present. Also, the data is not clustered in two or more regions which might also lead to a good correlation coefficient and all range from 0 to 1 can be predicted.

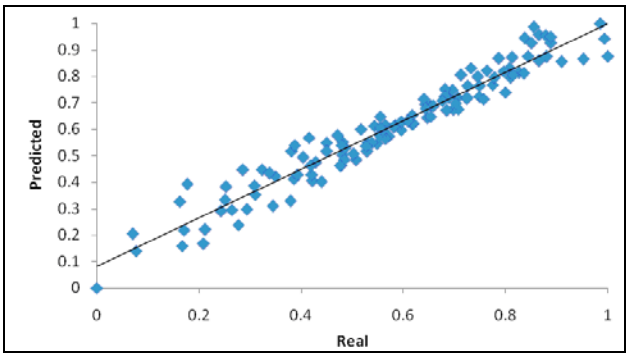


Figure 05 – Predicted versus real values of the dependent variable – model 05

4. Conclusions

The models were able to successfully describe different process steady states. These models represent and alternative tool for process monitoring in case of sensor failure. Although linear, the models can adequately describe the actual process values, being an important tool mainly for prediction of intermittent data which are not continuously measured such as chemical concentrations, or reduce lab analysis. It can also be used for sensor measurement checking.

5. Acknowledgements

The authors thank REPAR/PETROBRAS, Fundação Araucaria and CNPq for financial support.

References

R.G. BRERETON, 2003, Chemometrics: Data Analysis for the Laboratory and Chemical Plant. John Wiley & Sons: London.

H. GAI; JIANG,Y.; QUIAN,Y.; KRASLAWSKI,A., 2008, Conceptual Design and Retrofitting of the Coal-Gasification Waste Water Treatment Process. Chemical Engineering Journal. 138, 84-94.

L.C. HARDISON, 1988, Removal of Hydrogen Sulfide From Sour Water. US. PATENT 4.784.775.

T.J. HARRIS; SEPPALA,C.T.; DESBOROUGH,L.D., 1999, A Review of Performance Monitoring and Assessment Techniques for Univariate and Multivariate Control Systems. Journal of Process Control. 9, 01-17.

D.M. HIMMELBLAU, 1970, Process Analysis by Statistical Methods. 1ST Ed. John Wiley & Sons: New York.

R.A. JOHNSON; WICHERN,D.W., 2002, Applied Multivariate Statistical Analysis. 5TH Ed. Prentice-Hall: Upper Saddle River.

K.K. MANDALL, 2007, Optimize Your Gas Concentration Unit Operations. Hydrocarbon Processing. 86, 51-56.

K.B. MacCAULEY, 2003, Analysis of Process Data. Internal Report. Queen's University.

I. MILETIC; QUINN,S.; DUDZIC,M.; VACULIK,V.; CHAMPAGNE,M., 2004, An Industrial Perspective on Implementing On-Line Applications of Multivariate Statistics. Journal of Process Control. 14, 821-836.

M.OTTO, 1999, Chemometrics: Statistics and Computer Application In Analytical Chemistry. 1ST Ed. Wiley-VHC: Weinheim.

E. Van HOORN, 2002, Basics of Sour Water Stripping. Amine Experts, 411-444.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Operational Strategy for Water Supply in a Petrochemical Plant. Steady-State and Dynamic Approaches

Rita M. B. Alves, Antonio E. Bresciani, William S. Maejima, Claudio A. O. Nascimento

LSCP/CESQ - Department of Chemical Engineering, Polytechnic School - University of São Paulo, Av. Prof. Luciano Gualberto, n. 380, tv. 3, CEP 05508-900 São Paulo, SP, Brazil.

Abstract

The aim of this work is to present a mathematical model developed in order to simulate several operating scenarios involving water supply to a vapor generation unit in a petrochemical plant. The operational strategy suggested involves steady state and dynamic considerations. The case study is a petrochemical plant that uses water from two distinct sources with different qualities and prices. Thus, two case studies were carried out: 1. definition of the optimal proportion between both sources of water in case of supply stability: a model based on the minimum total cost per cubic meter of useful water was developed; 2. definition of the best operational strategy in case of failure in the main water source: a dynamic model was developed representing the flow rates as a time-dependent mathematical function in order to evaluate possible disturbances and all the operating conditions of interests, including the risk of total lack of water for vapor generation, which would cause the whole plant to shut down. The results show that it is possible to preserve operational continuity and stability, with lower water consumption and wastewater generation.

Keywords: Water supply, Optimization, Steady-state model, Dynamic Modeling.

1. Introduction

Water is an essential resource in industrial activities. In most process industries water is vital for many operations and is utilized for different purposes. In many regions with high concentration of industrial activities, or those close to urban centers, the operation of industrial plants is limited by the reduction in the capacity of water sources to supply demand, by increasing restrictions in the discharge of effluents into the environment and by competition with urban consumption. This situation has led to high costs of water supply and wastewater treatment (Goldblatt et al., 1993), resulting in a powerful economic motivation to rationalize water use, and an increasing the need for adequate water management and wastewater minimization, especially in the petrochemical industry, due to the large rates of consumption involved. Water and wastewater minimization techniques have been widely researched, developed and applied to process industries. Bagajewicz (2000) presented a review of recent design procedures for water networks in refineries and process plants. The basic concept of water minimization is to maximize water reuse and to identify regeneration opportunities (Wang and Smith, 1994). In general, there are four approaches to water minimization (Zver and Glavic, 2005): process changes, reuse, regeneration reuse and regeneration recycling. On the

other hand, process engineers seek simpler alternatives for reducing water demand in their plants. One strategy consists of evaluating individual process and utility units to reduce their inherent water requirements. Examples include studying operational strategies in order to guarantee water optimization in steady-state and transient scenarios.

Industrial water management and optimization require, in many cases, tailor made solutions. Thus, the present work presents a mathematical model developed for simulating several operating scenarios involving the water supply to a vapor generation unit in a petrochemical plant. The operational strategy suggested involves steady-state and dynamic considerations. The results show that it is possible to preserve the operational continuity and stability, with lower water consumption and wastewater generation at reduced cost.

2. Problem Statement

The case study is a petrochemical plant located in a region with a high concentration of industrial activities and close to an urban agglomeration. This industry uses water from two distinct sources: A and B, tap water and industrial water, respectively. Water from source A presents good quality, however at a high price and with low supply reliability. On the other hand, water from source B has, conversely, low quality, low price, and high supply reliability. Part of the water fed to the plant is treated in a demineralization unit in order to remove dissolved mineral matter that causes the water to be hard, thus making it proper to feed the vapor generation unit.

Figure 1 shows the studied scheme of water supply to the demineralization process, which consists of three sets of cationic and anionic resin beds, M_1, M_2 and M_3. Eventually, the resin beds become saturated, and at this point they stop softening the water. It is then time to regenerate it. It is considered as a semi-batch operation. Two batteries operate simultaneously while another one is being regenerated. Thus, it is possible to provide the required water flow rate to the vapor generation unit. The operating cycle length is proportional to water hardness.

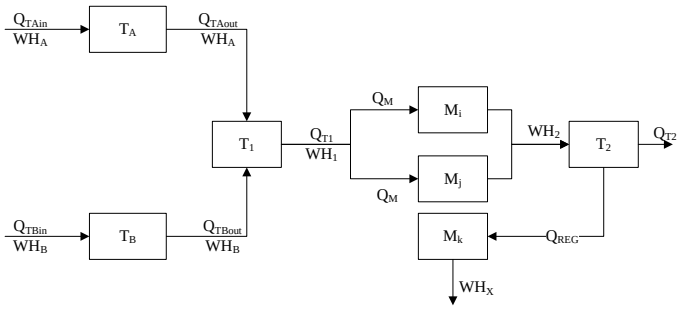


Figure 1 - Demineralization unit

The level of the tank that contains water from source A, T_A, is controlled at 90% and the tank with water from source B, T_B, is aligned to provide water to the system. In order to reduce costs, the studied plant used to feed the demineralization unit with a mixture from both sources in a pre-defined proportion; in general, 15-20% from source B. Assuming a fixed regeneration time and fixed flow rate of treated water used for regeneration of the resin beds, a greater operating cycle length is expected whenever the unit uses more volume of water from source A, which presents lower hardness than

water from source B. Thus, two main case studies were carried out: 1. definition of the optimal proportion between both sources of water in case of supply stability – the steady-state model; 2. definition of the best operational strategy in case of failure in the supply of water from source A. Matlab version 6.5 was used to develop the models.

3. Modeling

Figure 2 shows the algorithm developed in order to model the studied system.

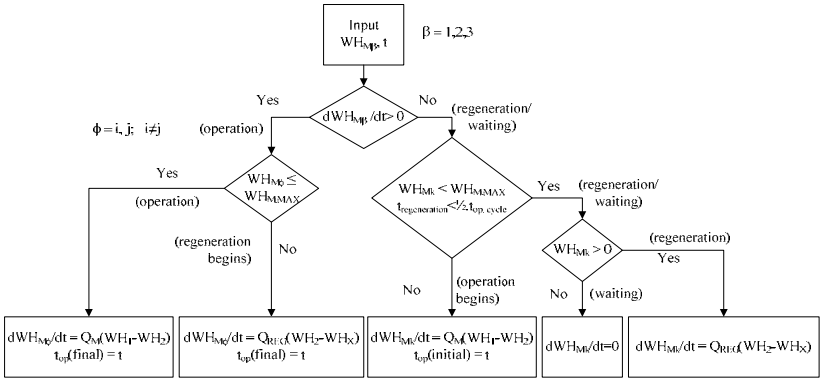


Figure 2 - Algorithm developed to model the system

3.1. Steady-State Model

The steady-state model aims to establish the optimal proportion between both sources of water in case of supply stability, taking into consideration economical aspects. The developed model was based on the minimum total cost per cubic meter of useful water. In order to carry out the calculations, the following conditions were assumed:

- the hardness of water from sources A and B are WH_A = 30 ppm and WH_B = 100 ppm, respectively;
- the maximum hardness removed by the demineralization unit, WH_{M_max}, is 48600 g/operating cycle. This value was evaluated through the Eq. (1) by assuming the usual proportion of water from source B (x_B = 0.15) and 1200 m^3/operating cycle;

$$WH_{M,max} = V_{Total} \cdot \left[\left(x_A . WH_A \right) + \left(x_B . WH_B \right) \right] \tag{1}$$

- the total flow rate of water treated in the demineralization unit, V_{Total} (m^3/operating cycle), was evaluated by Eq. (2);

$$V_{Total} = 48600 \Big/ \left\{ \left(x_A . 30 \right) + \left[\left(1 - x_A \right) . 100 \right] \right\} \qquad x_A = \left(0 \ldots 1 \right) \tag{2}$$

- The number of operating cycles per day, Ncycles, that the demineralization unit is able to run was calculated by Eq. (3);

$$Ncycles = \frac{Q_{T2} \cdot 24}{V_{Total} - Q_{REG}} \qquad \textit{(operating cycles/day)} \tag{3}$$

where Q_{T2} (≥110m^3/h) corresponds to the treated water flow rate used to feed the vapor generation unit and Q_{REG} (240 m^3/operating cycle) is the treated water flow rate used to regenerate the resin bed.

Eq. (4) was used to evaluate the total cost per hour of the process, *Cost 1*, including the costs for water acquisition, C_A and C_B, and the operational costs, C_{op}.

$$Costs\ 1\ (US\$/h) = Q_{TA} \cdot C_A + Q_{TB} \cdot C_B + C_{op} \cdot f \tag{4}$$

where C_A = US$ 3.00/m^3 ; C_B = US$0.70/m3; C_{op} = US$ 1000.00/operating cycle *and f= (Ncycles/24)*. Eq. (5) shows the total costs per cubic meter of treated water.

$$Costs\ 2\ (US\$/m^3) = \frac{Costs\ 1}{(V_{Total} - Q_{REG}).f} \tag{5}$$

Table 1 shows the total cost as a function of the proportion of water from source A.

Table 1 – Results of the steady-state model

Source A (%)	0	10	20	30	40	50	60	70	80	85	90	100
Ncycles	10.73	9.34	8.12	7.04	6.07	5.2	4.42	3.7	3.05	2.75	2.46	1.91
Costs (US$/m3)	5.45	5.26	5.09	4.94	4.81	4.69	4.59	4.49	4.40	4.36	4.32	4.25

As the proportion of water from source A increases, the number of operating cycles per day and the total cost decreases. This result shows that the costs of water acquisition are compensated by the operational costs, i.e., the better the water quality, the higher the length of the operating cycle and the lower the treatment costs. Thus, the optimal operating condition entails using water solely from source A, despite its higher price. Operation with 100% of water from source A, compared to the normal operation, would lead to the following advantages: reduction in the total useful water cost, equivalent to about US$ 110 thousand per year; and a reduction of 73365 m^3/year in the total water consumption and wastewater generated in the process.

3.2. Dynamic Model

A model was developed in order to dynamically simulate failure scenarios in the supply of water from source A, a somewhat frequent occurrence. The objective was to evaluate the process in the hypothesis of reduced flow rate or total lack of water from source A, allowing fluctuation in the level of tanks T_A and T_2, but keeping a constant water flow rate to the vapor generation unit in order to guarantee operational continuity and stability. The flow rates are represented as a time-dependent mathematical function in order to evaluate possible disturbances and all the operational conditions of interest.

In order to dynamically represent the system, a set of material balances was carried out:

$$\frac{dV_{TA}}{dt} = Q_{TAin} - Q_{TAout} \tag{6}$$

$$\frac{dV_{TB}}{dt} = Q_{TBin} - Q_{TBout} \tag{7}$$

$$\frac{dV_{T1}}{dt} = Q_{Taout} + Q_{TBout} - 2Q_M \tag{8}$$

$$\frac{dWH_1}{dt} = \frac{Q_{TAout} \cdot WH_A}{V_{T1}} + \frac{Q_{TBout} \cdot WH_B}{V_{T1}} - \frac{2Q_M \cdot WH_1}{V_{T1}} \tag{9}$$

$$\frac{dWH_{M\phi}}{dt} = Q_M \cdot (WH_1 - WH_2) \qquad \phi = i \text{ and } j \qquad \text{operating cycle} \tag{10}$$

$$\frac{dWH_{Mk}}{dt} = Q_{REG} \cdot (WH_2 - WH_X) \qquad \text{resin regeneration} \tag{11}$$

$$\frac{dV_{T2}}{dt} = 2Q_M - Q_{REG} - Q_{T2} \tag{12}$$

The dynamic study was carried out assuming the optimal operating condition as defined in the steady-state study, i.e., only water from source A was used instead of a mixture of both sources, considered by this industry as the standard operating condition. The optimal operating conditions were evaluated by minimization of an objective function defined in terms of total cost according to Eqs. (4) and (5).

4. Results and Discussion

Different scenarios involving the failure of water supply from source A were evaluated: flow rate reductions of 25%, 50%, 75% and 100%. The minimum cost was obtained at tank T_A level equal to 10% for all scenarios studied. The observed differences in the costs calculated as a function of the tank level were not significant.

For the dynamic simulation, the following conditions were assumed:

- after total failure of the water supply from source A ($Q_{TAin} = 0$), the level of the tank T_A decreases until it reaches a minimum of 10%.
- to avoid entire plant shutdowns, some action should be taken before the level of the tank T_2 reaches a minimum of 30%.
- fixed regeneration time and fixed flow rate of treated water used for regeneration;
- the regeneration time, t_{REG} cannot be greater than the operating cycle time, $t_{op\text{-}cycle}$;
- water hardness of the regeneration process output stream, WH_X, is constant and equal to 1000ppm.

4.1. Case Study 1 :

In this case, the cause of the failure was fixed and the flow rate of A, $Q_{TAin,}$ came back to its original value 11.4 hours after the failure takes place. Water from source B was not used. Figure 3 shows the results. Just after the water supply is normalized, the three demineralization unit batteries start to run simultaneously. Since the water flow rate to each bed set is lower than that fed during normal operation, the beds take longer to become saturated and the regeneration process carries out. However, it is interrupted and the three beds are required to operate in order to guarantee the treated water supply to the vapor generation unit. Over time, the whole operation tends to achieve the steady-state regime.

4.2. Case Study 2:

In this scenario, water from source B was required to feed the system. Water from source A was not used. Figure 4a shows the results. The system behavior is similar to the case study 1. A higher number of operating cycles is observed since water from source B presents greater hardness than water from source A and the process takes longer to achieve a steady-state operation. The system presents a great oscillation which can cause lost of stability in the entire plant or reduction in the quality of treated water. This means that a quick action can avoid operational issues. As an example, tank T_B

was aligned to provide the required water flow rate to stabilize the system and restore the original level of tank T_2. Figure 4b shows this result.

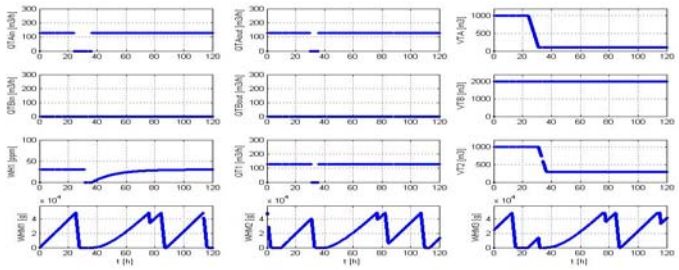
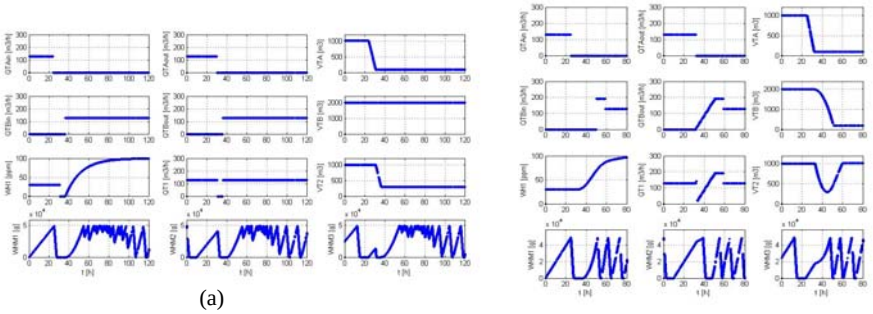


Figure 3 - Case study 1 simulation results



(a) (b)

Figure 4 - Case study 2 simulation results

5. Conclusion

The simulation of steady-state scenario shows that is possible to optimize the system with reduction in the total useful water cost as well in the water consumption and wastewater generated in the process.

The developed model was able to represent the behavior of the transient regime due to disturbances in the water supply. In emergency situations, considering safer operating conditions rather than minimum cost as a deciding factor, in order to take appropriate actions, is recommended. Thus the system should be able to return to the steady-state condition keeping the quality of treated water and avoiding the plant shutdowns.

Acknowledgments

The authors gratefully acknowledge FAPESP and FUSP for the financial support.

References

M. E. Goldblatt, K. S. Eble and J. E. Feathers (1993) Zero Discharge: What, Why and How, Chemical Engineering Progress, 89(4) 22-27.

M. Bagajewicz (2000) A review of recent design procedures for water networks in refineries and process plants, Computers and Chemical Engineering, 24(9-10): 2093-2113

Y. P. Wang and R. Smith (1994) Wastewater minimisation, Chemical Engineering Science, 49(7): 981-1006

L. Žbontar Zver and P. Glavic (2005) Water minimization in process industries: case study in beet sugar plant, Resources, Conservation and Recycling, 43(2): 133-145

10th International Symposium on Process Systems Engineering–PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento,
and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Best feature selection for texture classification

Daeyoun Kim[a], J. Jay Liu[b], Chonghun Han[a]

aSchool of Chemical and Biological Engineering, Seoul National University, San 56-1, Shillim-dong, Kwanak-gu, Seoul 151-742, South Korea
bSamsung Electronics, Myeongam-ri 200, Tangjeong-myeon, Asan, Chungchengnam-do 336-840, South Korea

Abstract

Texture analysis techniques enable to determine the quality of product surfaces measured by image sensors. In previous works, wavelet texture analysis based on the conventional wavelet transform and wavelet packets have been recognized as the most successful technique for classifying steel quality. In this work, we propose a texture classification strategy based on a best feature selection method, which improves classification accuracy. Our proposed methodology has been applied and validated in the classification of the surface quality of rolled steel sheets.

Keywords: Wavelet texture analysis, texture classification, feature selection, product surfaces, best features selection.

1. Introduction

The visual quality of a product surface must be controlled or maintained for most product manufacturers.(Liu *et al.*, 2007) The surface quality of steel sheets is important to automakers since it can significantly affect the coating quality. Some visual characteristics of product surfaces can be measured by digital image sensors. Texture analysis enables to determine the quality of product surfaces by using image sensors for measurement. Texture can be defined as an attribute representing the spatial arrangement of the gray levels of the pixels in a region of a digital image.(Geraci, 1990) Texture analysis has been successively applied to determine quality of product surfaces.(Chang *et al.*, 1993) Previous work(Bharati *et al.*, 2004) found that wavelet texture analysis offered the best performance in texture classification. Liu *et al.*(Liu, *et al.*, 2007) proposed an advanced methodology based on wavelet packets for characterizing steel surfaces.(Bharati, *et al.*, 2004) As wavelet packets use a large number of texture signatures, it has greater potential to include discriminative features for classification. These over-complete wavelet packets also impart increased computational costs and unnecessary features in the classification process. In order to use wavelet packets as a feature extraction method, the significant features for classification from full wavelet packets need to be selected.

This work proposes a texture classification methodology that selects the best feature from full wavelet packets. The proposed methodology was implemented and validated in a steel surface case. The results of the proposed methodology show performance improvements as compared to conventional wavelet transform and the full wavelet packets.

2. Theory for Wavelet Texture Analysis

2.1. Conventional Wavelet Transform

For multi-resolution representations, using wavelets is well established and successfully applied in signal and image processing.(Livens *et al.*, 1997) The principle is essentially filtering a signal through a pair of filters—a low-pass filter (*h*) and a high-pass filter (*g*)—and down-sampling the filtered signals by two (i.e., dropping every other sample), respectively.(Chang *et al.*, 1992) A signal $f(x)$ is decomposed into its components through shifting and dilating of a prototype function $\psi(x)$.(Rajpoot, 2002) In this algorithm, a detail section is extracted at each decomposition level from the approximation section of the previous level, and the number of detail sections is reduced to half the number of the previous level. This is termed a conventional wavelet transform.(Chang, *et al.*, 1992; Liu, *et al.*, 2007)

2.2. Wavelet Packets

The conventional wavelet transform has a set of frequency channels comprising narrower bandwidths in the lower frequency region. In many practical cases, information on natural textures is not found in the lower frequency region, but rather in the middle or higher frequency region. These are termed quasi-periodic signals.(Chang, *et al.*, 1993; Liu, *et al.*, 2007) To analyze quasi-periodic signals, a generalized form of wavelet termed a wavelet packet is needed.(Rajpoot, 2002) A conventional wavelet transform is implemented through an iterative decomposition of approximation coefficients, using a two-channel filter bank. On the other hand, wavelet packets use all the coefficients to decompose, resulting in an equal frequency bandwidth.(Chang, *et al.*, 1993; Liu, *et al.*, 2007)

2.3. Wavelet Texture Analysis

Wavelet texture analysis is a very powerful method as compared to other texture analysis methods.(Rajpoot, 2002; Liu, *et al.*, 2007) To reflect scale-dependent characteristics, textural features can be extracted from each sub-image.(Livens, *et al.*, 1997) Each sub-image from wavelet decomposition is represented as D(J, I), where J is the depths of level of decomposition and I is the number of packets in the depths. The energy of the sub-image can be defined as follows:

$$E_{D(J,I)} = \left\| D(J,I) \right\|_F^2 \tag{1}$$

The energy divided by the number of pixels equals the averaged power (normalized energy). The most popular wavelet textural feature, the wavelet energy signature, is a vector organized by the energies of all the sub-images.(Kim *et al.;* Kim *et al.*, 2007; Liu, *et al.*, 2007)

3. Texture Classification Strategy Based on Best Feature Selection

The proposed texture classification methodology based on the best feature selection has four steps: feature extraction, feature selection, feature reduction, and classification.

3.1. Feature Extraction Step

The extraction of important textural information from images is essential for texture analysis. Bharati *et al.*(Bharati, *et al.*, 2004) compared these methods and concluded that the transform-based method—particularly the wavelet texture analysis—offered the best performance in terms of classifying steel textures. For texture classification, we selected a wavelet texture analysis based on wavelet packet bases, since this has been shown to perform better than other methods.(Liu, *et al.*, 2007)

3.2. Feature Selection Step

3.2.1. Criterion for feature selection

A criterion is needed to select the best features from wavelet packets. For this study, an index that represents class separation or distances among classes can be used.(Saito *et al.*, 1995) Fisher's criterion provides reliable class separation among the different bases. Let the between-class scatter matrix be S_B, and the within-class scatter matrix be S_W. Then, the optimal projection W_{opt} is chosen as the matrix with orthonormal columns, which maximizes the ratio of the between-class scatter to the within-class scatter.(Duda *et al.*, 1973)

$$J(W_{opt}) = \frac{\left|W^T S_B W\right|}{\left|W^T S_W W\right|}. \tag{2}$$

The objective value, denoted as $J(W_{opt})$, is termed Fisher's criterion. Fisher's criterion obtained from Eq. (2) can be used as an approximated classification performance measure for different datasets. Among the different feature vectors, a well-classifiable one will have a high Fisher's criterion value. We utilized these characteristics to measure a discriminative basis extracted from wavelet packets for the texture of images.(Rajpoot, 2002; Liu, *et al.*, 2007)

3.2.2. Simple searching method

We used a top-down searching method for feature selection. Feature selection was performed using a simple rule, which has already been used for other pattern recognition problems.(Kumar *et al.*, 2001) This simple agglomerative rule does not require recursive computation for sorting discriminations of features.

C_{opt}: best features consist of only selected wavelet energy signatures

C_{JI}: wavelet energy signatures at level J; the number of packets is I

$\overline{C_{JI}}$: decomposed wavelet energy signatures of C_{JI}

$\overline{\overline{C_{JI}}}$: decomposed wavelet energy signatures of $\overline{C_{JI}}$

As one moves from the upper level to the lower level—and from lower packet numbers to higher packet numbers in each level—the best features are updated by using simple selection rules as follows:

if $J(C_{JI} \cup C_{opt}) > J(\overline{C_{JI}} \cup C_{opt})$, then C_{opt} includes undecomposed signatures C_{JI}, and

if $J(C_{JI} \cup C_{opt}) \leq J(\overline{C_{JI}} \cup C_{opt})$, then C_{opt} includes decomposed signatures $\overline{C_{JI}}$,

where J is Fisher's criterion. Feature selection is performed from the upper level (level 1) to the lower level. Only the decomposed bases have higher criteria than the present levels, to be included in the best features. Otherwise, undecomposed bases were included.

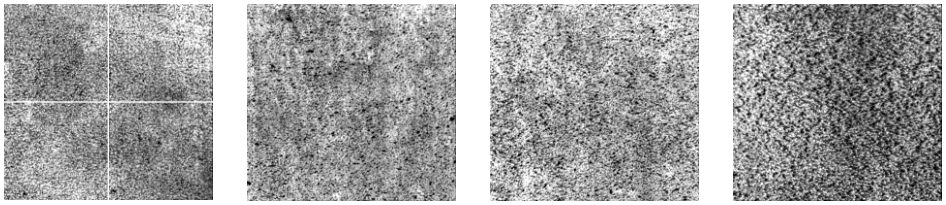
Top-down searching requires an assumption for pruning appropriately.

If $J(C_{JI} \cup C_{opt}) > J(\overline{C_{JI}} \cup C_{opt})$, then $J(C_{JI} \cup C_{opt}) > J(\overline{\overline{C_{JI}}} \cup C_{opt})$.

3.3. Feature Reduction Step

Further dimension reduction can be achieved by using projection methods. In this study, we selected Fisher's discriminant projection for maximizing discrimination among textural classes.

Figure 1. Steel Surface Sheets(Kim, *et al.;* Liu, *et al.*, 2007)



(a) Excellent (b) Good (c) Medium (d) Poor

3.4. Texture Classification Step

Classification algorithms can be differentiated according to their features. In this study, we used K-nearest neighbors, a simple supervised classification algorithm, to classify the features.

4. Application to Classification of Steel Surfaces

4.1. Steel Surfaces

To validate our proposed methodology, we compared proposed best features with previous conventional and full-wavelet packets in the classification of steel-surface quality. The quality of each steel surface was determined as excellent, good, medium, or poor, as shown in Fig. 1. For this study, 35 images of steel surfaces were used. The example of bad surface quality contains deep pits that have joined to form craters. The number of samples within each class was 8 (excellent), 9 (good), 6 (medium), and 12 (poor). To obtain greater consistency, each original image was divided into four sub-images (see Fig. 1a). Therefore, the new image set contained 140 images.

4.2. Wavelet Texture Analysis

We applied a two-dimensional wavelet packet transform to all 140 images, using order-one Coiflet wavelet filters up to level 4. The maximum decomposition level used for all the three methods (best features, conventional wavelet transform, and wavelet packets) was four. The maximum decomposition level used for the wavelet transforms was selected in accordance with certain guidelines.(Chang, *et al.*, 1993) For feature extraction, each sub-image was converted to a one-scalar value.

4.3. Best Feature Selection

Conventional wavelet transform and full wavelet packets do not require feature selection and thus proceed to the next feature-reduction step. As explained earlier, best features were selected according to a top-down searching rule.

4.4. Quality Classification

Fisher's discriminant analysis maximizes discrimination in each class. Following Fisher's projection, we equally applied a three-nearest neighbor classification for all bases. The misclassification rate was estimated by using a "leave one out" cross-validation.

5. Results of Case Study

Figure 2 shows the feature reduction by Fisher's discriminant plot. In Fig. 2, level 4 wavelet packets and conventional wavelet transform have some linearly inseparable points. The good classes and excellent classes are not perfectly separable in wavelet packets. On the other hand, Fig. 1(c) shows that each class in the best features is perfectly separated.

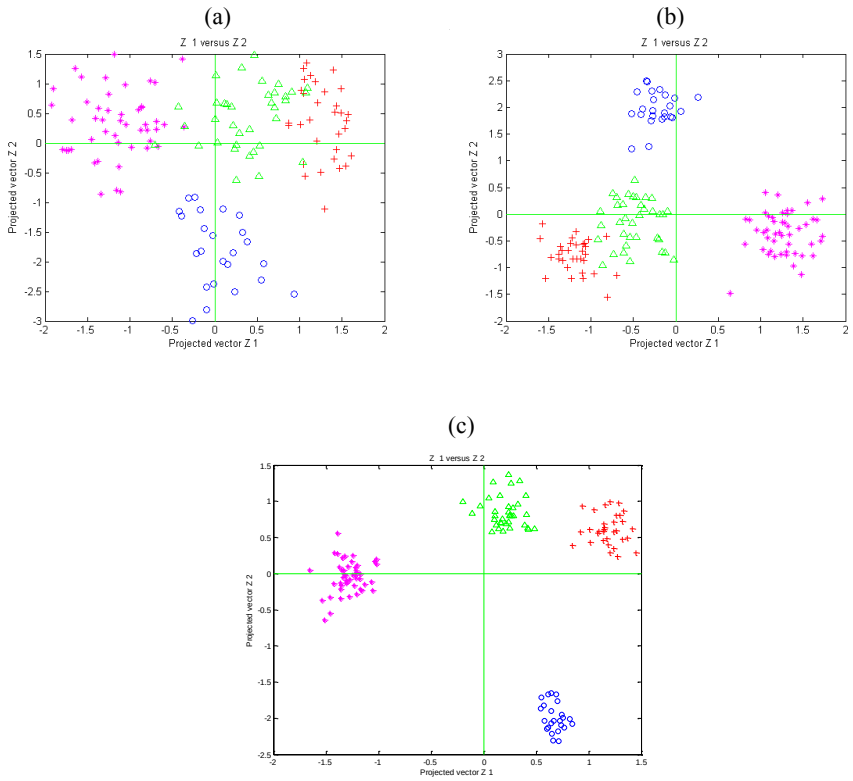


Figure 2. Fisher's Discriminant Plots.

(a) Conventional Wavelet Transform; (b) Full-Wavelet Packets; (c) Proposed Best Features.
Class labels are +: excellent; △: good; ○: medium; *: poor.

Table 1 shows the same results as Fig. 2 in a quantitative manner. The proposed approach achieves the best classification performance among the three methodologies (wavelet packets, conventional wavelet transform, and proposed best features). Figure 2 and Table 1 show that selected features from top-down searching suitably selected and improved classification performance.

Table 1. Classification Errors of Steel Surfaces

	Conventional Wavelet Transform	Wavelet Packets	Proposed
Excellent	5.9%	0.9%	0%
Good	24.1%	2.9%	0%
Medium	0%	0%	0%
Poor	1.1%	0%	0%

6. Concluding Remarks

A previous work employed wavelet texture analysis based on full wavelet packets or conventional wavelet transform for texture classification. These two approaches have some disadvantages with regard to such classification. Wavelet packets include insignificant packets, which lead to misclassification. Conventional wavelet transform does not provide sufficient packets to represent whole information. In this study, a texture classification methodology, based on best feature selection, was proposed and validated using steel-surface sheets. We compared the performance of the proposed method with that of full wavelet packets and conventional wavelet transform. This work proved that the proposed method is more appropriate for characterizing steel surface quality than both the conventional wavelet transform and full wavelet packets.

References

M. H. Bharati, J. J. Liu, and J. F. MacGregor, 2004, Image texture analysis: methods and comparisons: Chemometrics and Intelligent Laboratory Systems, 72, 57-71.

T. Chang, and C. C. J. Kuo, 1992, A wavelet transform approach for hierarchical texture segmentation, 812, 816-820.

T. Chang, and C. C. J. Kuo, 1993, Texture analysis and classification with tree-structured wavelet transform: Image Processing, IEEE Transactions on, 2, 429-441.

R. O. Duda, and P. E. Hart, 1973, Pattern Classification and Scene Analysis: Wiley New York.

A. K. Geraci, 1990, IEEE Standard Glossary of Image Processing and Pattern Recognition Terminology: Institute of Electrical and Electronics Engineers (IEEE) Std, 610.614-1990.

D. Kim, J. Liu, and C. Han, 2007, Comparison of features in classifying steel surface quality, ESCAPE 18proceedings.

D. Kim, J. Liu, and C. Han, 2007, Quality monitoring of steel surface using wavelet packet transfrom, ICCAS2007 proceedings, 861-864.

S. Kumar, J. Ghosh, and M. M. Crawford, 2001, Best-bases feature extraction algorithms for classification ofhyperspectral data: Geoscience and Remote Sensing, IEEE Transactions on, 39, 1368-1379.

J. J. Liu, D. Kim, and C. Han, 2007, Use of Wavelet Packet Transform in Characterization of Surface Quality: Ind. Eng. Chem. Res., 46, 5152-5158.

S. Livens, P. Scheunders, G. van de Wouwer, and D. Van Dyck, Year, Wavelets for texture analysis, an overview: Image Processing and Its Applications, 1997., Sixth International Conference on, 581-585 vol.582.

N. M. Rajpoot, Year, Texture classification using discriminant wavelet packet subbands, III-300-III-303 vol.303.

N. Saito, and R. R. Coifman, 1995, Local discriminant bases and their applications: Journal of Mathematical Imaging and Vision, 5, 337-358.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Algorithm For Integrated Production And Utility System Scheduling Of Batch Plants

Mujtaba H. Agha[a], Raphaele Thery[a], Gilles Hétreux[a], Alain Hait[b], Jean M. Le Lann[a]

[a]*Laboratoire de Génie Chimique (UMR 5503), 5 rue Paulin Talabot, 31106 Toulouse.*
[b]*Institut Supérieur de l'Aéronautique et de l'Espace, 10 av. E. Belin, 31055 Toulouse*

Abstract

Onsite utility systems feature in many industrial units. Due to technological advances in utility generation technologies, even energy intensive batch plants are attracted towards using onsite utility system. However in batch plants emphasis is placed solely on production process and utility aspects are not incorporated into the general scheduling model. This article presents an extended resource task network (ERTN) representation that can be used to develop a universal scheduling model that would concurrently undertake scheduling of production unit as well as an onsite utility system.

Keywords: Short-term scheduling, Utility system, Network representation, MILP.

1. Introduction

Increased competition and shrinking profit margins has forced chemical processing and manufacturing related industries to look for ways of improving productivity and reducing operational costs. To achieve these objectives, development of scheduling and production planning models has received increased attention. Scheduling can be defined as allocation of limited resources to produce one or more products. This allocation answers three primary questions – when (at what time), where (in which equipment) and how much (quantity) of material needs to be processed to produce desired products. Mathematical programming, especially Mixed Integer Linear Programming (MILP), because of its rigorousness, flexibility and extensive modeling capability is widely used for scheduling problems [Floudas & Lin, 2005].

Multiproduct / multipurpose batch plants are inherently flexible and their scheduling is dependent on the nature of product demand. When reliable forecasting can not be established then production is only driven by the available orders leading to short-term scheduling problem [Mendez *et al.*, 2006].

State Task Network (STN) proposed by Kondili *et al.* [1993] was a major step towards developing a universal short-term scheduling model for a batch plant. STN proposed a generalized framework for presenting the production process which was then used to develop a mathematical formulation applicable to all types of batch plants. Pantelides [1994] proposed Resource Task Network (RTN) which is extended framework for presenting the production process. RTN contains more information about processing equipment and their utilization, an aspect which is not explicitly shown in STN.

2. Problem Statement

Although STN and RTN are equally good at handling scheduling problems, but in both representations utility is simply considered as an external resource which allows

processing equipment to transform material from one form to another. The limited availability of utility is incorporated either as a scheduling resource constraint or as a term in objective function. This is not sufficient as unlike the resources classically considered in the scheduling problems (machinery or work force), utility has special characteristics which must be taken into account. Utility is present in various forms (electricity, steam at different pressure levels, hot/cold water, etc). It is also a resource which must be consumed immediately after its generation as generally it can not be stored in its ultimate useful form. Moreover, in case of an onsite utility system, the utility availability at a given time is dependent on the intrinsic functioning of the utility system. In consequence, the scheduling of the utility system is intertwined with the scheduling of the production process. Thery *et al.* [2008] demonstrated that significant gains in operational costs can be achieved by using an integrated production unit and utility system scheduling model. However, their formulation was too specific and lacking on the aspect of universality that is present in STN and RTN representations. Hence, there is a need to develop a broader universal framework that would incorporate the operational planning of utility system considerations into scheduling problem.
This work aims to attain this objective by developing an extended resource task network (ERTN) representation which can be used to carry out the scheduling of production unit and utility system simultaneously in a single universal scheduling model.

3. Extended Resource Task Network (ERTN) Representation

The main feature of ERTN is unified representation of production unit and onsite utility system. In ERTN various nodes and arcs are connected in such a manner that physical illustration of industrial process is attained. These nodes and arcs are described below:

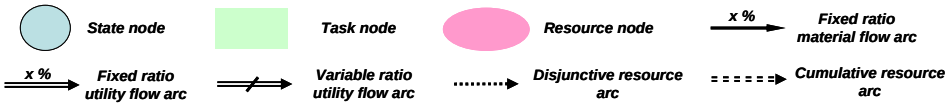


Figure 1: Nodes and arcs used in ERTN representation

3.1. Nodes used in ERTN

ERTN uses the same types of nodes as employed by its predecessor STN and RTN representations. The state node (circle) represents materials (feeds, intermediate and finished products) and the utilities (water, steam, electricity, fuel, etc). It is important to point out that the same state node can be considered as material resource (for onsite utility system) and as a utility resource (for the production system). The task node (rectangle) represents a process operation which transforms material or utility from input state(s) into output state(s). The resource node (oval) represents the processing equipment in which a particular processing operation, i.e., a task can be undertaken.

3.2. Arcs used in ERTN

Arcs are used to provide the necessary connection between the state, task and resource nodes of ERTN. Contrary to the RTN representation, an explicit distinction is made between matter flow (materials and utilities) and resource flow. Before discussing the respective arcs it is essential to outline the concept of principal and support utility that is used by the ERTN framework. The ‘principal utilities’ comprises of water and its derivatives (steam at various pressures, hot and cold water) while all other utilities (electricity, fossil fuel, etc) are considered as ‘support utilities’.
In the ERTN framework the flow of material between task node and state node is represented by material arc. Similarly the flow of principal utility between task node

and state node is represented by utility arc. The material and principal utility flows can be either in fixed ratio of task batchsize or in variable ratio of task batchsize. The cumulative resource arc principally represents flow of different utilities which are generated by utility system and then consumed by production unit. On the other hand, the disjunctive resource arc just identifies the resources (in our case processing equipments) where a particular task can be undertaken.

3.3. Demonstration of ERTN representation through an example

To illustrate how ERTN can be constructed, this article uses example of an industrial unit comprising of a batch production unit and an onsite utility system (figure 2).

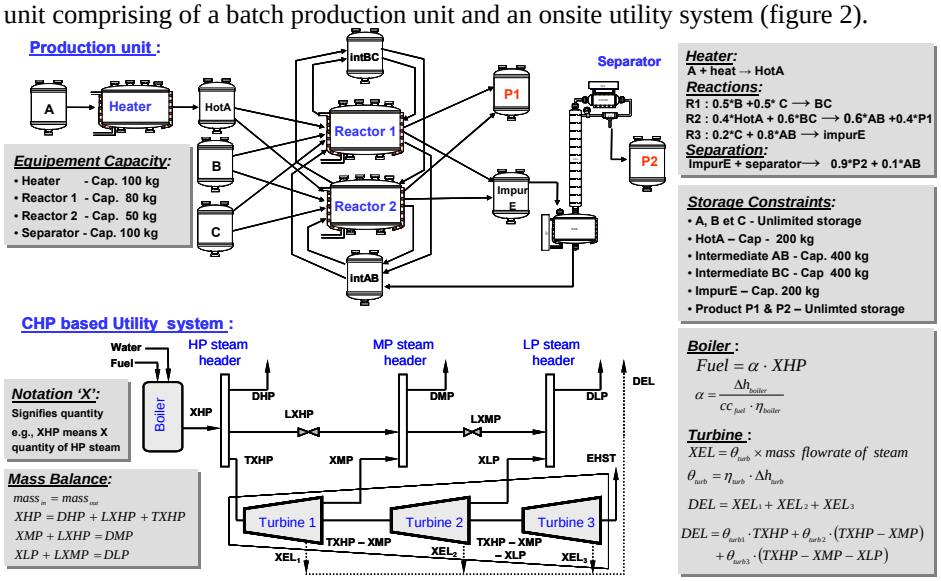


Figure 2: A batch production unit with a CHP based onsite utility system

The production unit converts three feeds (*A*, *B* & *C*) into four intermediate products (*HotA*, *intBC*, *intBC* & *ImpurE*) and two finished products (*P1* & *P2*). To achieve this, four process equipments (*Heater*, *Reactor 1*, *Reactor 2* & *Separator*) are used which consume following utilities – electricity (*DEL*), steam at high pressure (*DHP*), medium pressure (*DMP*) and low pressure (*DLP*). Three reactions (*R1*, *R2* & *R3*) can take place either in *Reactor 1* or in *Reactor 2*. To produce these utilities an onsite Combined Heat & Power (CHP) based utility system is used. The boiler consumes fuel to transform water into HP steam. HP steam is converted into MP and LP steam by using multistage turbine or by using pressure release valves. The advantage of using turbine is that it not only reduces pressure but also simultaneously generates electricity. The stages of turbines are represented as *turbine 1*, *turbine 2 & turbine 3*.

The ERTN representation of above mentioned industrial unit is shown in figure 3. The upper half depicts the production unit while the lower half depicts utility system. The link between the two units is provided by flow of utilities *DEL*, *DHP*, *DMP* and *DLP*.

It is important to clarify the need of a cumulative resource arc. The utility arc simply illustrates the flow and transformation of principal utility from one phase to another within the utility system. However during this transformation process a task in utility system might not only consume support utility (e.g., task *T9* consuming *fuel* to convert water into *HP* steam) but it can also generate support utility (e.g., task *T11* not only converts *HP* steam into *MP* steam but it also generates *electricity*). As a result, the notion of cumulative resource arc is introduced which would incorporate the use of

support utilities without violating utility mass balance. This concept is further extended to include flow of utilities from utility system to production unit. The utility states providing utilities are considered as additional resources to the production unit and flow from them is represented as cumulative resource arc.

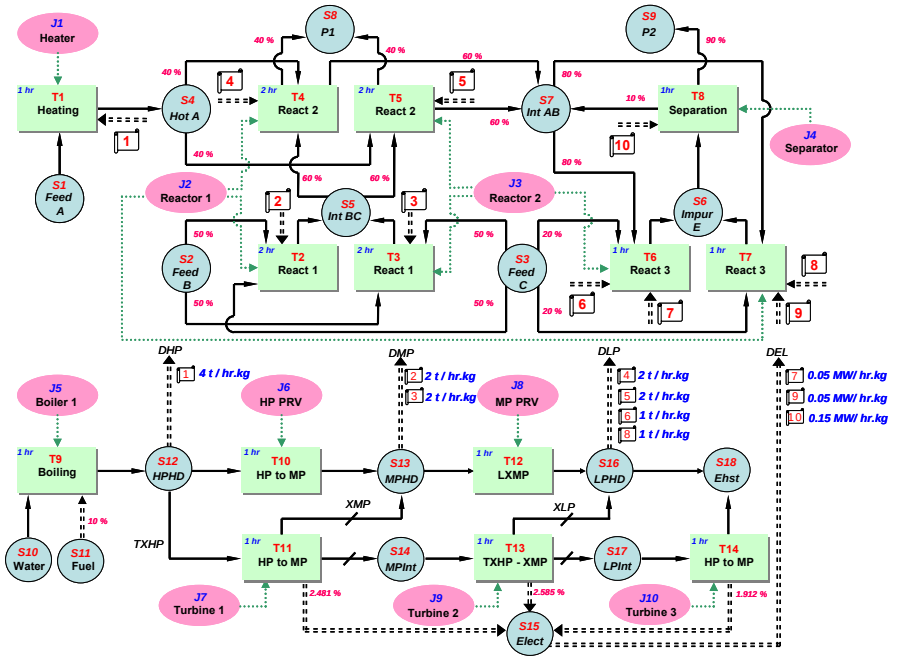


Figure 3: Extended resource task network (ERTN) representation of industrial unit

3.4. Mathematical formulation

To formulate the scheduling model, discrete time based Mixed Integer Linear Programming (MILP) is used. This model is established on following hypotheses:

1. Processing times for each task are known and are independent of batch size.
2. Once a task has been started it can not be stopped before its completion.
3. A disjunctive resource can be shared by several tasks. However, during the same time period only one task can be undertaken in a resource.
4. A task consumes one or more type of utilities dependent on the batch size undertaken.
5. A state can have multiple output and input streams. However, all the input streams must be of same quality. If different quality streams need to be mixed then this would constitute as a separate task.
6. A state can be considered as a storage station where material / utility can be stored. However, only primary energy (e.g., fuel) can be stored in utility state. Other utilities (steam, electricity, etc) can not be stored in their ultimate useful form.
7. The material inputs and outputs from a task in production units are fixed and known proportion of its batchsize. However, the utility inputs and outputs from a task in utility system are dependent on the utility demands.

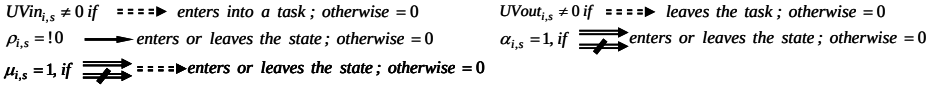


Figure 4: Proportions signified by each arc

Equation (1) represents allocation constraints, that a resource *j* can only initiate task *i* at a given time *t*. It also illustrates that the resource *j* is not available to perform another task during the periods $t'= t- p_i+1$ till $t'= t + p_i -1$ (i.e., duration of the task). Equations (2) and (3) show production capacity and storage limitation constraints respectively. Equations (4) accounts for the material and utility mass balance across all the states in ERTN. Equations (5) and (6) represent the utility mass balance across the tasks in utility system. Equation (7) illustrates the utility demands from utility system by tasks in production unit. Equation (8) accounts for the support utilities generated by the tasks in utility system. Finally, Eq. (9) corresponds to objective function that is minimized to achieve the short-term schedule. The objective function is a weighted sum of electricity purchase cost and storage parameters for tardiness starting date.

$$\sum_{i\in I_j}\sum_{t'=t-p_i+1}^{t} W_{i,t} \le 1 \qquad \forall j \in J,\ \forall t \in 1,..,T \tag{1}$$

$$W_{i,t}V_i^{\min} \le B_{i,t} \le W_{i,t}V_i^{\max} \qquad \forall i \in I,\ \forall t \in 1,..,T \tag{2}$$

$$0 \le S_{s,t} \le C_s^{\max} \qquad \forall s \in S, \forall t \in 1,..,T \tag{3}$$

$$S_{s,t} = S_{s,t-1} + \sum_{i\in I}\rho_{i,s}^{prod} B_{i,t-p_i} + \sum_{i\in I}\mu_{i,s}^{prod} BUout_{s,i,t} - \sum_{i\in I}\rho_{i,s}^{cons} B_{i,t} - \sum_{i\in I}\mu_{i,s}^{cons} BUin_{s,i,t} + imt_{s,t} - ext_{s,t} \qquad \forall s \in S, \forall t \in 1,..,T \tag{4}$$

$$B_{i,t} = \sum_{s\in S_{util}} \alpha_{i,s}^{prod} BUout_{s,i,t} \qquad \forall i \in I_{util}, \forall t \in 1,..,T \tag{5}$$

$$B_{i,t} = \sum_{s\in S_{util}} \alpha_{i,s}^{cons} \cdot BUin_{s,i,t} \qquad \forall i \in I_{util}, \forall t \in 1,..,T \tag{6}$$

$$BUin_{s,i,t} = uf_{i,s}W_{i,t} + UVin_{i,s}\sum_{t'=t-p_i+1}^{t} B_{i,t'} \quad \forall s \in S_{util}, \forall i \in I, \forall t \in 1..T \tag{7}$$

$$BUout_{s,i,t} = uf_{i,s}W_{i,t} + UVout_{i,s}\sum_{t'=t-p_i+1}^{t} B_{i,t'} \quad \forall s \in S_{util}, \forall i \in I, \forall t \in 1..T \tag{8}$$

$$Funobj = \sum_{t}^{T}\sum_{s\in S_{elec}} electricity\ price \cdot imp_{s,t} + \sum_{t}^{T}\sum_{s\in S} h_s \cdot ext_{s,t} + \sum_{t}^{T}\sum_{s\in S}\frac{h_s \cdot S_{s,t}}{1000} \tag{9}$$

3.5. Results and concluding remarks

To judge effectiveness of ERTN the above mentioned example was formulated and resolved using software XPRESS-MP. The problem was solved on an Intel(R) Core(TM) 2 Duo CPU @ 2.00 GHz and 1.00 GB of RAM. A 24 hour planning horizon is considered with the time intervals divided into 24 one hour periods. The demand of the finished product P1 & P2 is 300 and 250 kilograms respectively. The short-term schedule attained through ERTN representation is presented in figure 4. The schedule gives nature, time and batchsize of tasks undertaken in each resource of industrial unit.

This article presents Extended Resource Task Network (ERTN) representation which can be used to solve the solving the short-term scheduling of the batch plant. The ERTN extends its predecessors network representation by incorporating the operational planning aspects of an onsite utility system. The use of ERTN representation results in a universal scheduling model which simultaneously performs scheduling of production unit and that of onsite utility system.

This study is a part of research being conducted at CNRS, whose objective is energy management solutions for mono and multi-sites. In future the discrete time model will be replaced by a continuous time formulation.

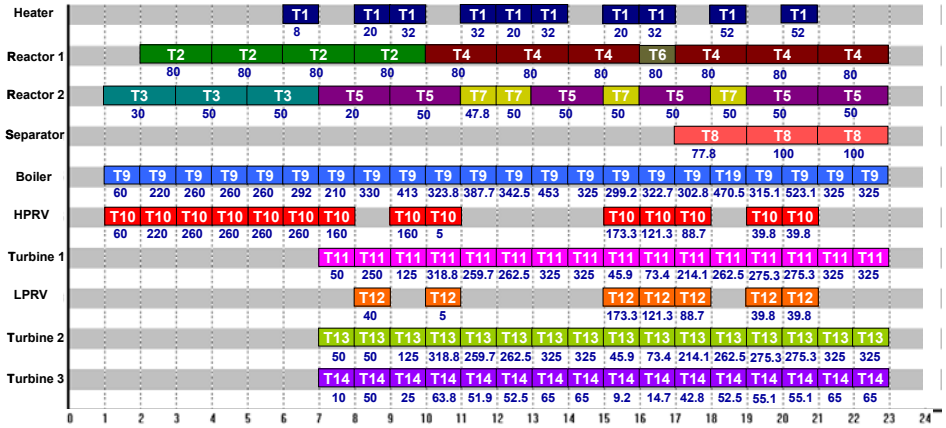


Figure 5: Scheduling of the industrial unit based on ERTN representation

Nomenclature

Indices *i*: tasks, *s*: states, *j*: resources, *boil*: boiler, *elec*: electricity, *mat*: material, *util*: utility

Sets I: Set of tasks, I_{util}: Set of tasks in utility unit, J: Set of resources, S: Set of states, S_{mat}: Set of material states, S_{util}: Set of utility states

Variables

$B_{i,t}$	Amount of material at time t being undertaken by task i
$imt_{s,t}$	Import into a state from external source
$emt_{s,t}$	Export to a state from external source
$BUin_{s,i,j,t}$	Utility input to a task
$BUout_{s,i,j,t}$	Utility output to a task
$S_{s,t}$	Storage of state s at time t
$W_{i,t}$	Binary variable, with value '1' if task i is undertaken at time t, otherwise value is '0'.

Parameters

$\rho_{i,s}^{cons}, \rho_{i,s}^{prod}$	proportion of material of state s consumed or produced by task i
$\alpha_{i,s}^{cons}, \alpha_{i,s}^{prod}$	proportion of utility of state s consumed or produced by task i
$\mu_{i,s}^{cons}, \mu_{i,s}^{prod}$	proportion of utility directed into or from state s of utility system to task i in production unit
C_s^{max}	maximum storage capacity of state s
h_s	Coefficient for tardiness starting date
$uf_{i,s}$	Fixed utility demand of a task
$UVin_{i,s}$	Utility demand of a production task
$UVout_{i,s}$	Support utility generated by utility task

References

C.A. Floudas & X. Lin, X, 2005. Mixed Integer Linear Programming in Process Scheduling: Modeling, Algorithms, and Applications, Annals of Operations Research, 139(1), 131-162.

E. Kondili, C.C Pantelides, W.H. Sargent, 1993, A general algorithm for short-term scheduling of batch operations–I. MILP formulation, Computers and Chemical Engineering, 17(2), 211-227.

C.A. Méndez, J. Cerda, I.E. Crossmann, I. Harjunkoski and M. Fahl, 2006, State-of-the-art review of optimization methods for short-term scheduling of batch processes, Computers and Chemical Engineering, 30, 913-946.

C.C. Pantelides, 1994, Unified frameworks for the optimal process planning and scheduling. In: CAChE Publications. Proceedings of the Second Conference on Foundations of Computer Aided Operations, 253-274.

R.Thery, G. Hétreux and M. Agha, 2008, Modèle intégré pour l'ordonnancement d'ateliers batch et la planification de centrales de cogénération. Proceedings of 7eme Conférence Internationale de Modélisation et Simulation (MOSIM 08).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Operation and Composition Control of a New Pressure Swing Batch Distillation System

Arpad Kopasz, Gabor Modla, Peter Lang

BUTE Department of Building Services and Process Engineering, Muegyetem rkp. 3- 5. H-1521 Budapest,Hungary, E-mail: lang@mail.bme.hu

Abstract

The pressure swing separation of a binary minimum azeotrope (n-pentane-acetone) in a double column batch stripper is studied by rigorous simulation. For controlling the product compositions a simple scheme is presented. On the basis of temperatures of bottoms product PID controllers manipulate their flow rates varying the reboil ratios. The controllers are tuned. The influence of the most important operational parameter (division ratio of liquid leaving the common top vessel) is investigated. For rigorous simulation calculations a professional dynamic flowsheet simulator is applied.

Keywords: batch stripping, pressure swing, minimum azeotrope, control.

1. Introduction

Binary pressure sensitive azeotropes can be separated by pressure swing distillation (PSD). Continuous PSD was first applied in the industry in 1928. Phimister and Seider (2000) studied first the batch (stripping) and semicontinuous application of PSD by simulation. First Repke et al. (2007) investigated experimentally the batch PSD (PSBD, pilot-plant experiments for the separation of a minimum azeotrope in a batch rectifier (BR) and stripper (BS)). Modla and Lang (2008) studied different batch configurations (BR, BS, combination of BR and BS and middle vessel column(MVC)) by feasibility studies and rigorous simulation for the separation binary (max. and min.) homoazeotropes. By modifying the MVC, which has not been proven suitable for the PSBD, they suggested two new double column batch configurations: rectifier (DCBR, *Fig. 1a*) and stripper (DCBS, *Fig. 1b*). They compared the different configurations for a given set of operational parameters without optimisation and control. For min. azeotropes the best results (minimal specific energy consumption for the same quality products) were obtained with the DCBS and for maximum azeotropes with the DCBR, respectively. The columns of these configurations can be operated practically in steady state. Modla et al. (2009) studied the feasibility of batch PSD separation of most frequent types of ternary homoazeotropic mixtures.

When operating these new configurations the liquid composition of the common vessel of the two columns must be kept between the two azeotropic compositions. The ratio of two product flow rates of a DCBS can be changed by varying the /reboil ratios and/or the ratio of division of the liquid flow leaving the common vessel.

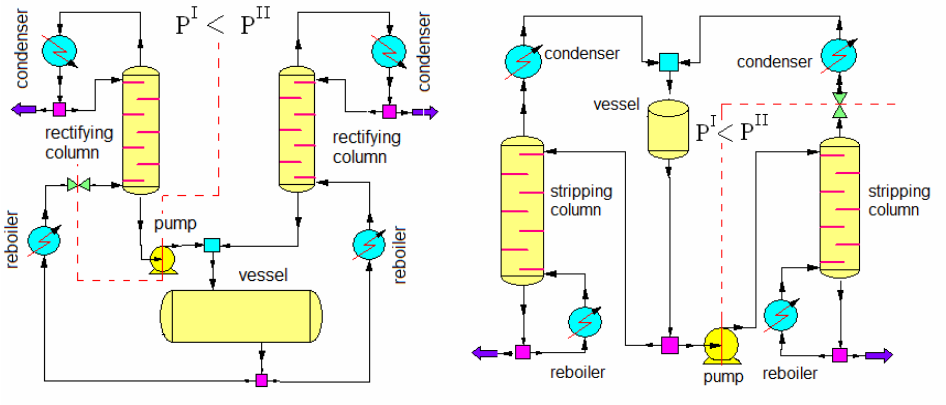
The goals of this paper are:

-to investigate the operation of the DCBS for the separation of a minimum azeotrope,

-to study a simple scheme for the control of product compositions (temperatures of bottoms product are controlled and their flow rates are manipulated) ,

-to investigate the effects and to determine the optimal value of the liquid division ratio.

The calculations were made for the mixture n-pentane-acetone by using a professional dynamic simulator (CCDCOLUMN).



a. DCBR *b. DCBS*

Figure 1. The scheme of a DCBR and DCBS

The temperature-composition (*T-x,y*) diagrams and azeotropic data of the mixture studied are shown for the two different pressures in *Fig. 2* and *Table 1*, respectively.

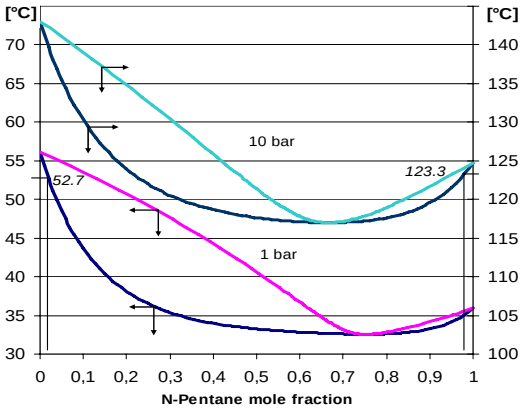


Figure 2.T-x,y diagrams of n-pentane-acetone

Component		pentane(*A*)	acetone(*B*)
P	[bar]	1.01	10
x_{az}		0.754	0.668
T_{az}	[°C]	32.5	116.9
$T_{BP,A}$	[°C]	36.0	124.7
$T_{BP,B}$	[°C]	56.2	142.9

Table 1. Azeotropic data (UNIQUAC parameters: 571.98 and 95.033 cal/mol.)

2. Simulation method

The following simplifying assumptions were applied
- theoretical stages,
- negligible vapour hold-up,
- constant volumetric liquid plate hold-up.

The model equations to be solved are well known:
a. Non-linear differential equations (material balances, heat balances),
b. Algebraic equations (vapour-liquid equilibrium (VLE) relationships, summation equations, hold-up equivalence, physical property models).

For solving the above model equations we used the CCDCOLUMN dynamic flow-sheet simulator (ChemCad 6.0). For the simulation of columns simultaneous correction method was applied.
The following modules were used:
-DYNCOLUMN (column sections),
-DYNAMIC VESSEL (top vessel and product tanks),
-HEAT EXCHANGER, PUMP, VALVE,
-MIXER, DIVIDER,
-CONTROLLER, CONTROL VALVE.
The ChemCad model of the double column batch stripper with control of product compositions is shown in *Fig. 3*.

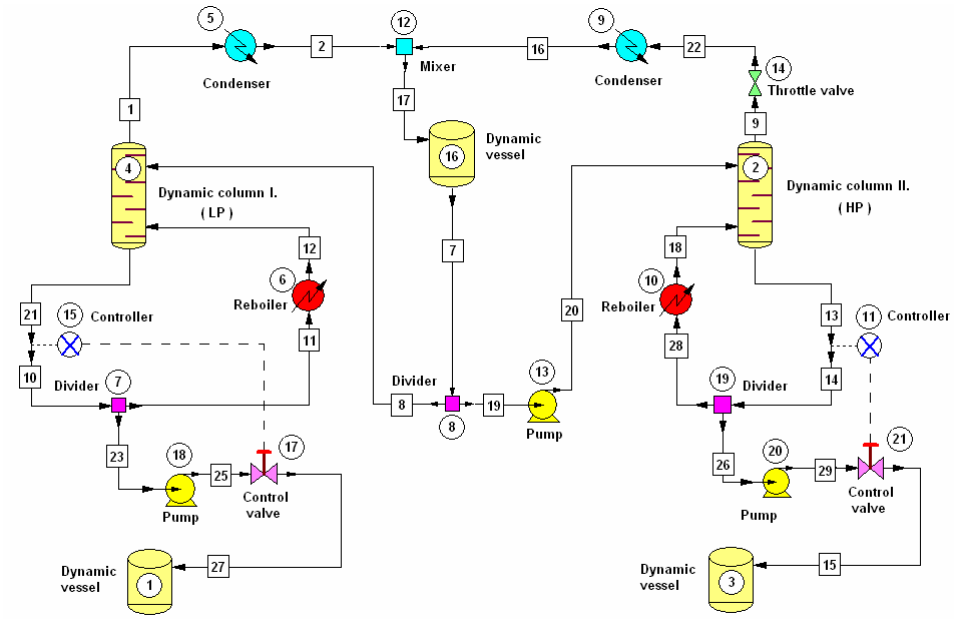


Figure 3. ChemCad model of the double column batch stripper with control loops.

3. Results

The number of theoretical stages for each column sections is 40. (The total condenser and total reboiler do not provide a theoretical stage.) The liquid hold-up is 2 dm3/plate. the pressure of the columns: P^{LP}=1.013 bar and P^{HP}=10 bar. At the start of the distillation plates of the columns are wet (they are filled with charge at its boiling point at the given pressure). The total flow rate of liquid leaving the common vessel: $L_{0,total}$ = L_0^{LP}+ L_0^{HP} = 6 m^3/h. The quantity of charge containing 30 mol% pentane is 4.022 kmol (0.471 m3). The prescribed purity is 98 mol% for both products. The reboil ratios R_s^{LP} and R_s^{HP} are changed by PID controllers manipulating (with linear control valves) the product flow rates W^{LP} and W^{HP}, respectively. The whole process is finished when the amount of liquid in the vessel decreases to 12.5 % of the charge.
First, the parameters of the two PID controllers (A_P, T_I and T_D) providing stable, good quality control of the product compositions in the whole region of liquid division ratio

($\phi=L_0^{LP}/L_{0,total}$) studied are determined. Then, the influence of this operational parameter on the performance of the PSBS is studied and its optimum value yielding the minimal overall specific energy consumption $((SQ^{LP}+SQ^{HP})/(SW^{LP}+SW^{HP})$ is determined.

3.1. Tuning of PID controllers

Our aim is to determine a set of parameters of the PID controllers which provide good quality control of product compositions by taking into consideration the usual criterions (maximal overshoot, control time, number of oscillations).

The quality of control is determined by the evolution of not only the controlled variables (temperature of the two bottoms products) but also that of the position of the two control valves (varying the flow rate of the two bottoms product). The following criteria of quality of control are given concerning the two control valves:

-maximal overshoot: 33 %,

-maximum number of oscillations during the settling time T_S (within an error band of +-5 %): 3.

In *Fig. 4* the evolution of the position of the control valve and bottoms composition of Column I for an inappropriate set of controller parameters is shown. (Both the position of the control valve and the controlled variable show oscillations without damping.)

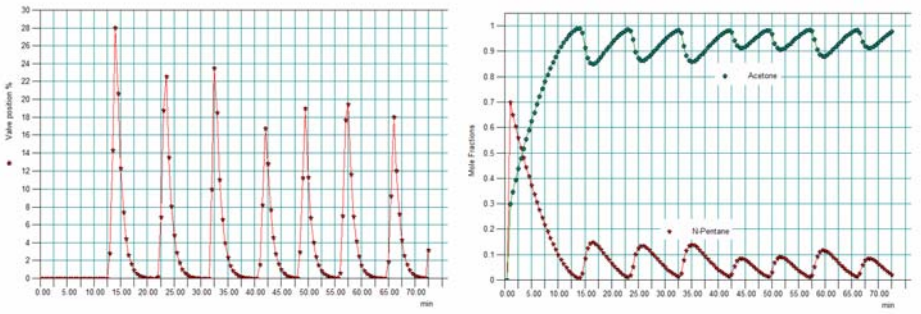


Figure 4. The evolution of the position of the control valve and bottoms composition for an inappropriate setting of PID parameters (Column I, PB=10 %, T_I= 1 min, T_D= 0)

Table 2. Parameters and quality data of control for an appropriate tuning

a. PID parameters:

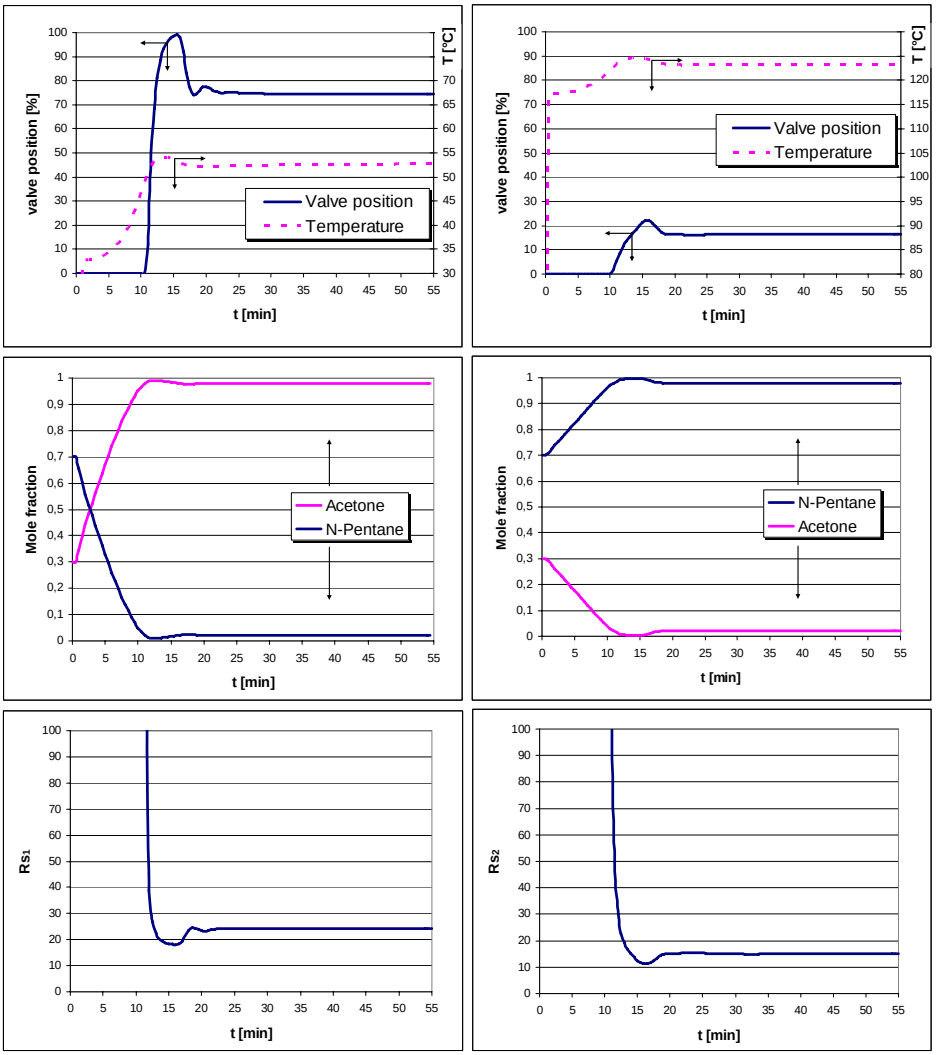
	PB, %	$T_{I,}$ min	$T_{D,}$ min	Set point,°C
Column I	45	13	0.50	52.7
Column II	120	3	1	123.3

b. Valve flow coefficients: K_v^I= 0.13 K_v^{II}=2

c. Control quality data:

	Column I:	Column II:
Maximal overshoot:	(99.13-74.45)/74.45=**0.33**	(21.97-16.64)/)/16.64=**0.32**
Settling time, min:	17.25-10.5=**6.75**	17.75-10=**7.75**
No. of oscillations within T_S:	1	1

For the controller parameters selected (*Table 2*) the evolution of the position of the control valves, bottoms compositions and reboil ratios is shown in *Fig. 5*. (*Table 2* contains the control quality data, as well.)



a. Column I *b. Column II*

Fig. 5. The evolution of control valve positions and bottom temperatures (a), bottoms compositions (b), reboil ratios (c) (ϕ=0.55)

3.2. Influence of the liquid division ratio

The liquid division ratio is varied in the region 0.3-0.9. The specific energy consumption is minimal at ϕ=0.55 (*Fig. 6*). Prescribed purity products are obtained with reasonable recovery (*Table 3*). This table contains also the most important results for the process, such as the total and specific energy consumptions of the production. It must be

still noted that the recoveries could be still increased by reducing the quantity of residue in the common top vessel. In the case studied we were able to practically empty the vessel while maintaining the prescribed purities in the product tanks. However under a certain amount of residue (12 % of the charge) the operation of the control loops became unstable, therefore we stopped the process.

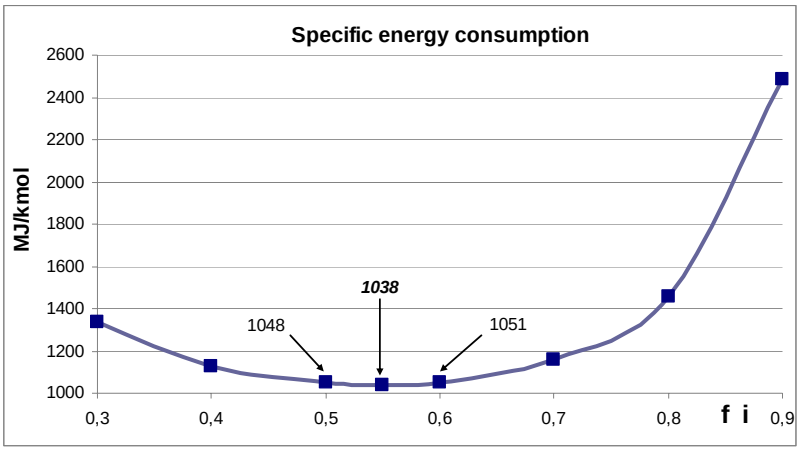


Fig. 6. The influence of the liquid division ratio on the specific energy consumption

Table 3. Most important results of the production for the optimal liquid division ratio

N-pentane recovery	%	75.32
Acetone recovery	%	67.54
N-pentane purity	mol %	98.20
Acetone purity	mol %	98.03
Total energy (SQ)	MJ	3106
Specific energy: $SQ/(SW_A + SW_B)$	MJ/mol	1038
Production time	min	54

Acknowledgement

This paper was supported by the Janos Bolyai Research Scholarship of the Hungarian Academy of Sciences and the Hungarian Research Funds (OTKA) (No: T-049184).

References

Modla G., A. Kopasz A. and P. Lang (2009). Feasibility of separation of ternary mixtures by pressure swing batch distillation, *PSE2009*.

Modla G. and Lang P. (2008). Feasibility of new pressure swing batch distillation methods, *Chem. Eng. Sci.*, 63, 2856-2874.

Phimister, J.R.; Seider, W.D. (2000). Semicontinuous pressure swing distillation, *Ind. Eng.Chem. Res.*, 39, 122-130.

Repke .J. U., Klein A., Bogle D., Wozny G.,(2007). Pressure Swing Batch Distillation for Homogenous Azeotropic Separation", *Chem. Eng. Res. Des.*, 85, 492-501.

10th International Symposium on Process Systems Engineering - PSE2009
Elisangela Orlandi de Sousa, Sandra Lúcia da Cruz and João Alexandre Ferreira da Rocha Pereira. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Monitoring Pipelines Through Acoustic Method.

Elisangela O. Sousa, Sandra L. Cruz, João A. F. R. Pereira

Department of Chemical Systems Engineering. School of Chemical Engineering
University of Campinas, UNICAMP. P.O. Box 6066, 13083-970, Campinas-SP, Brazil
E-mail address: eorlandi@desq.feq.unicamp.br

Abstract

Pipeline networks are complex systems of ducts used for liquid and gas transportation through long distances. They frequently cross highly populated regions, water supplies or natural reserves. Even small leaks in pipelines can lead to great losses of products and serious damages to the environment before it could be detected. With the purpose to track these leaks, a methodology is proposed for detection of leaks in pipelines based on acoustic method and on analysis of pressure transients generated by leak occurrence. Pressure transients and the sound noise generated by leakage are detected and analyzed in a pipeline operating with continuous flow of gas (air) under various operation conditions. The experimental results showed that it is possible to detect leaks in pipelines based on hydraulic transients and on acoustic methods. The acoustic method was decisive for leak detection; since the changes in pressure observed through the pressure transducer was not significant depending on the gas flow rate. The analysis of the signal amplitude for different frequencies shows that the leakage noise signal changes with both leak magnitude and pipeline pressure
Keywords: leak detection, acoustic method, pipeline networks.

1. Introduction

Pipeline networks are frequently used for transportation and distribution of liquid an gas products. Transport pipelines can cover large geographical areas and can be several thousand of kilometers long. They operate at relatively high pressures. Compressors at the beginning of the line provide the energy to move the gas through the pipeline and compressor stations are required at a number of points along the line to maintain the required pressure. In general, these pipelines are made of steel pipe and are buried below ground surface. The individual sections of pipes are joined by welding, and the pipe is externally coated to protect against corrosion. Pipe sizes can be as large as 1.5 m in diameter.

Shimanskii et al (2005) based on high temperature microphones, which respond rapidly to the appearance of the acoustic noise due to a leak. The main problem of the subsequent investigations was to develop a method for detecting a small leak at an early stage of its development. A leak detection method using microphones allows contact-free monitoring, which makes it possible to decrease several fold the number of sensors required, and the microphones can be placed in freely accessible zones around the piping, which makes it easier to service the microphones themselves and the piping.

Jing-pin et al (2006) analyzed the propagation characteristics of guided waves in acoustic leak location in pipelines. Time frequency methods were used in the analysis of acoustic leak signals. The mode components of the acoustic leak signal were obtained based on the relation of time-frequency distribution of acoustic leak signal and the

dispersion curves of guided waves. The research can provide a guideline for the mode selection in pipeline leak location, and help improve the accuracy of leak location.

Morozov et al (2007) described the first domestic multi-channel automated acoustic system for monitoring leaks, called SAKT. It was based on the generation of high frequency stress waves on the surface of pipes during outflow of a liquid which is under high pressure and high temperature through a rupture in the form of a crack.

Verde et al (2007) presented a method for the identification of two leaks in a pressurized single pipeline where both transient and static behavior of the fluid in leaks conditions are used to identify the parameters associated to the leaks without requirements of valve perturbation. They used an iron galvanized pipeline with 132.56 m long with a diameter of 0.105 m, thickness of 4.7 mm. Two sonic sensors had been installed extreme to measure flows of the static and transient state of the liquid. The measures of pressure had been carried through a pressure transmitter and the signals had been measured using a system of acquisition of data with a frequency of 100 hertz. The results had shown that to detect multiple leaks in conditions of static state of the liquid the test with a dynamic model is necessary complementary to reduce the time interval expense ploughs the identification of the leaks.

2. Experimental Work

The rigid pipeline consisted of a bunch of ½" in diameter, 60 m long galvanized iron tube. Figure 1 show the experimental assembly used to simulate the occurrence of leaks in the rigid pipeline.

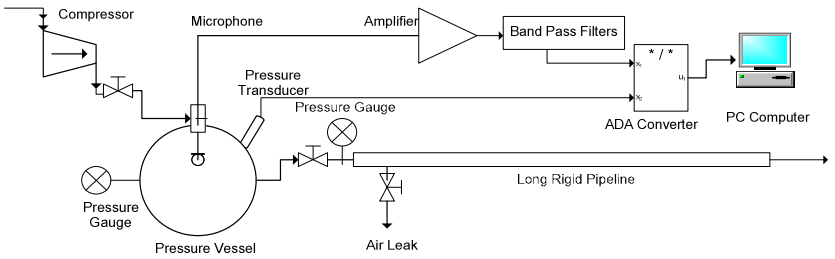


Figure 1. Rigid gas pipeline experimental assembly

The flexible pipeline consisted of a rolled up 50 m long transparent rubber helical wired. Figure 2 show the experimental assembly used to simulate the occurrence of leaks in the flexible pipeline.

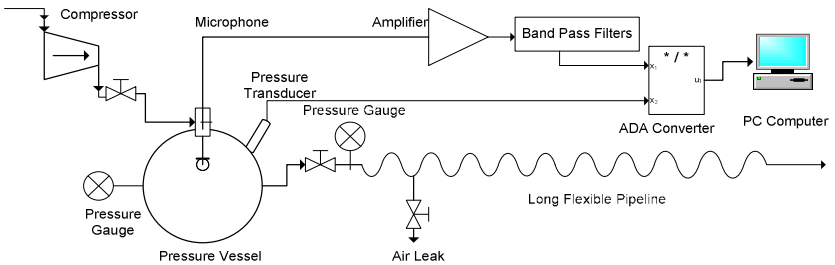


Figure 2. Flexible gas pipeline experimental assembly

Leaks were provoked manually through a rapid aperture valve installed in a side outlet at the entrance of the pipeline. The leak magnitude was controlled through orifice installed in the leak line. The orifice size varied from 0.1 mm to 5 mm in diameter.

The gas (air) used in the experiments was provided by the compressor which feed the air line of faculty laboratories and was fed to the pipeline through a 37.5 liters pressure vessel. Compressed air pressure was measured by manometers installed in the pressure vessel and in the pipeline. The air pressure (P0) varied from 1.0 to 7.0 kgf/cm^2.

In the leakage experiments with steady state gas flow a 0.8 mm orifice was installed at the end of the pipeline to keep it pressurized.

Pipeline monitoring was made through a microphone and also a pressure transducer installed in the pipeline, both connected to a PC computer through an ADA converter.

The microphone was developed in the laboratory to catch the sound noise generated by leak occurrence. On the first stage the signal was amplified and on the second one, the signal was filtered through three band pass filters, centered in 1 kHz, 5 kHz and 9 kHz each one, generating three continuous signals in different frequencies. A piezoelectrical transducer. The transducer was connected to the computer through an ADA converter.

The PC computer used for the data acquisition was a Pentium 233 MHz, 500MB and 16 MB RAM. The data acquisition software, written in C language, was developed to read and filter all the pressure transducers signals and to display the pressure transient and sound noise amplitude profiles plots.

3. Results and Analysis

3.1. Leak Detection in Rigid Pipeline

Figures 3 to 6 show the pressure transient and sound noise amplitude profiles caused by gas leak through orifices of 1.0 mm and 3.0 mm, when the rigid pipeline operated under constant pressure of 4.0 kgf/cm² and 6.0 kgf/cm².

When the leak occurred the pressure dropped, suddenly as the leak magnitude (orifice size) increased, while the sound noise amplitude increased rapidly, both remaining practically constant during leak occurrence. In Figure 1 the sound amplitude was about 1 V and 2 V in Figure 2. With the orifice of 3.0 mm the sound amplitude was about 3.5 V when the pressure was 4.0 kgf/cm² and reached 4.75 V when the pressure was 6.0 kgf/cm².

The dominant frequency was 5 kHz, independently of the orifice size, indicating that leak occurrence generated a medium sound noise. The second frequency was a high frequency, 9 kHz, with the 1.0 mm orifice and low frequency, 1 kHz, with the 3.0 mm orifice, indicating a more serious sound.

The obtained results also show that leaks were readily detected through the sound noise amplitude changes, but that not always happened through pressure transient profiles.

The initial work pressure has no influence on the leak; the pressure drop is almost constant. This is considered as choked flow where de fluid at a certain pressure and temperature through a restriction, such as the hole, or orifice, or a valve in pipe into a lower pressure, limiting the air speed.

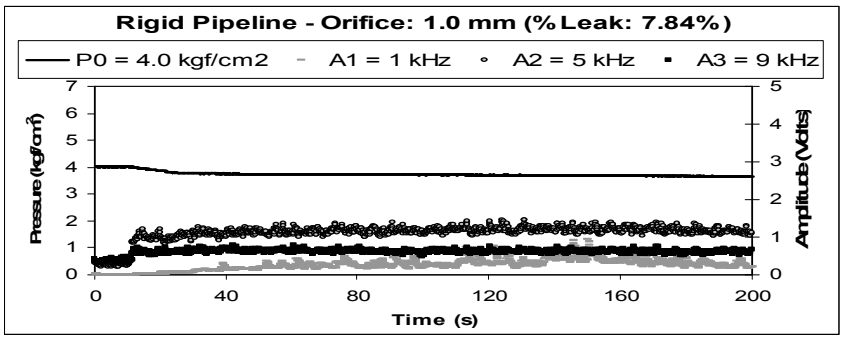


Figure 3. Pressure transients and sound amplitude. (Orifice 1.0mm). P0 = 4 kgf/cm^2.

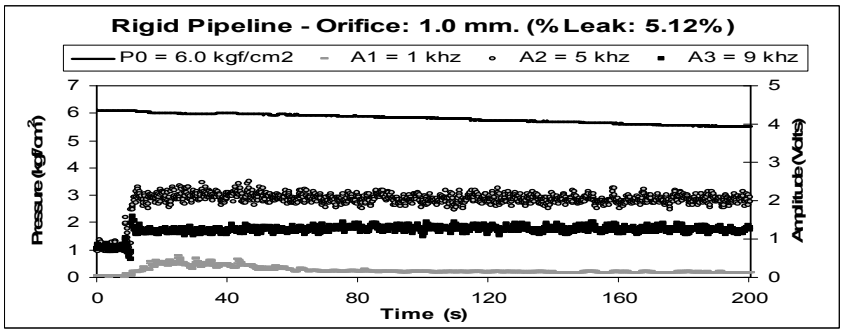


Figure 4. Pressure transients and sound amplitude. (Orifice 1.0mm). P0 = 6 kgf/cm^2.

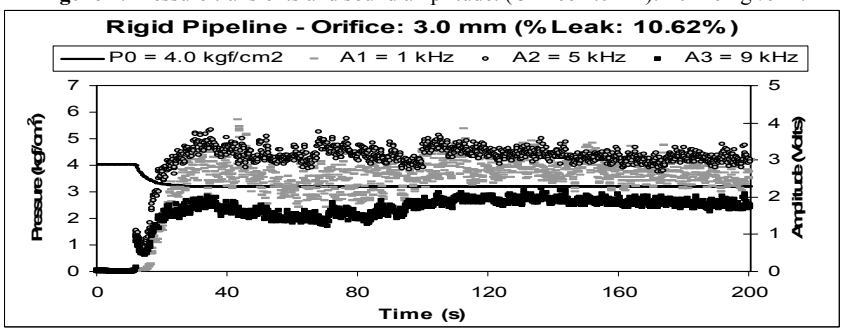


Figure 5. Pressure transients and sound amplitude. (orifice 3.0mm). P0 = 4.0 kgf/cm^2.

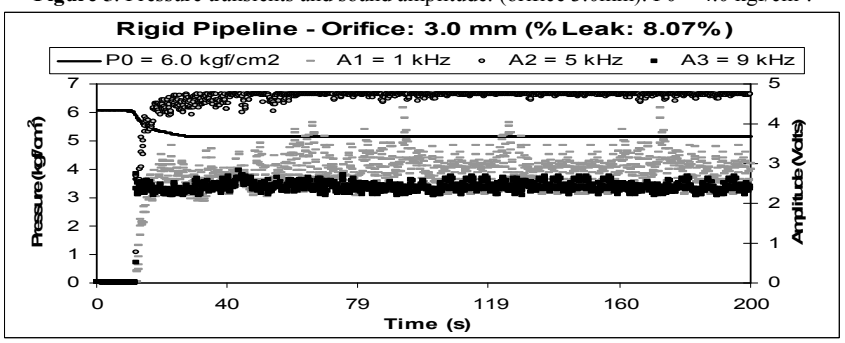


Figure 6. Pressure transients and sound amplitude. (orifice 3.0mm). P0 = 6.0 kgf/cm^2.

3.2. Leak Detection in Flexible Pipeline

Figures 7 to 10 show the pressure transient and sound noise amplitude when leak occurred through orifices of 1,0 mm and 3.0 mm and the flexible pipeline operated under constant pressure of 4.0 kgf/cm² and 6.0 kgf/cm².

The obtained results show that the pressure transient profiles presented the same behavior as that in the rigid pipeline. From the point of view leak detection it was not also be possible to detect the leak by using the pressure transient data for the 1.0 mm orifice.

The leak percentage calculated for the experiments showed that the leak percentage in flexible pipeline was lower when compared with the rigid pipeline, proving that when the leak is provoked in flexible pipeline, the pressure speeds of the wave has a damping. The graphics show that the sound noise amplitude was a few increasing and was not always a sudden change. The characteristics of sound noises were the same as in the rigid pipeline.

Figures 7 to 10 show that sound noise amplitude increase with both pressure and leak magnitude. With the orifice of 1.0 mm in diameter the sound amplitude was about 1 V when the pressure in the pipeline was 4.0 kgf/cm² and 1.5 V when the pressure was 6.0 kgf/cm². With the orifice of 3.0 mm the sound amplitude reached 2.5 V when the pressure was 4.0 kgf/cm² and reached 3.8 V when the pressure was 6.0 kgf/cm². These results show that the sound wave was attenuated in flexible pipeline.

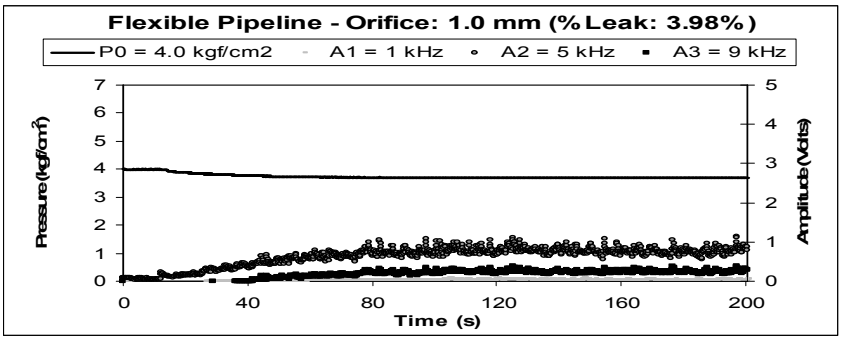


Figure 7. Pressure transients and sound amplitude. (orifice 1.0mm). P0 = 4.0 kgf/cm^2

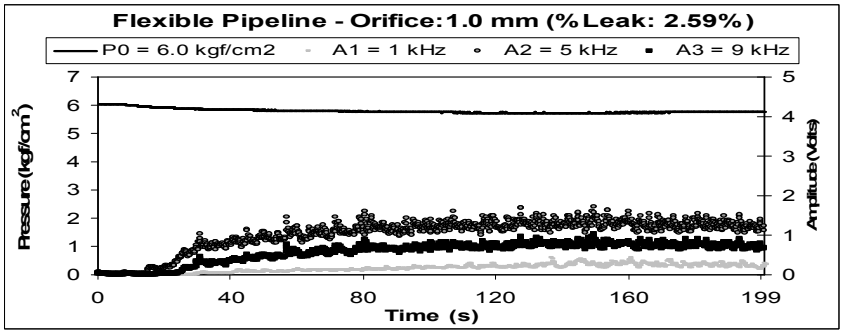


Figure 8. Pressure transients and sound amplitude. (orifice 1.0mm). P0 = 6.0 kgf/cm^2

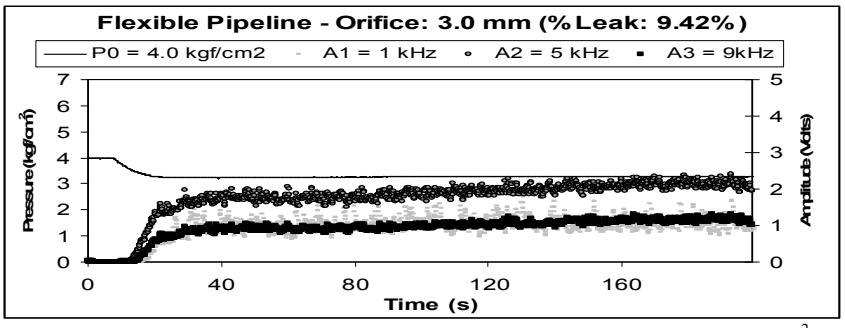


Figure 9. Pressure transients and sound amplitude. (orifice 3.0mm). P0 = 4.0 kgf/cm^2

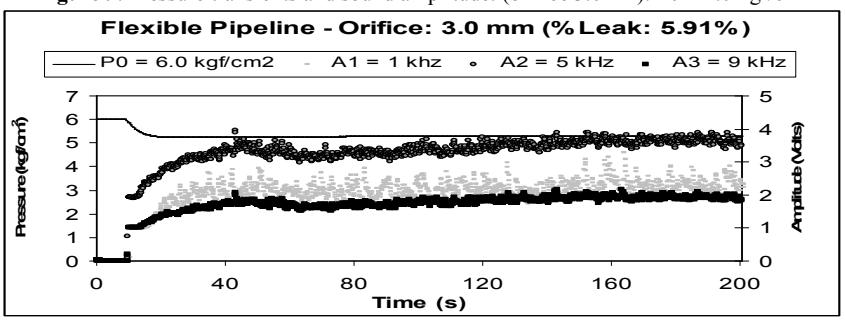


Figure 10. Pressure transients and sound amplitude. (orifice 3.0mm). P0 = 6.0 kgf/cm^2

4. Conclusions

A leak detection technique based on the sound noise generated by leak occurrence has been developed and tested to detect gas leaks in a long pipeline. The results showed that it is possible to detect leaks in pipelines based on acoustic method. In leak detection the presence of the microphone was decisive, since the changes in pressure measured through the pressure transducer may be negligible depending on the gas flow rate.

The analysis of the sound amplitude showed that the characteristics of leak sound noise changes with both leak magnitude and pipeline pressure. The present work is being continued so that the developed technique could become a tool for pipeline supervision.

References

Da Silva, H.V. Morooka, C.K. Guilherme, I.R. Da Fonseca, T.C. Mendes, J.R.P. Leak detection in petroleum pipelines using fuzzy system. *J Petrol Sci. Eng*. 49:223-38, 2005.

Morozov, S.A., Kovtun, S.N., Budarin, A.A., Dvornikov, P.A., Kurdryaev, A.A., Kondratovich, F.V., Shutov, P.S., Shvetsov, D.M., Konoplev, N.P.. Development of Acoustic Leak Monitoring System. *Atomic Energy*, vol. 103, no 6: 925 - 931, 2007.

Jing-Pin, J., Ren-Yuan, F., Cun-Fu, H. Bin, W. Modal analysis of acoustic leak signal in pipelines using time-frequency analysis. *Front. Mechanical engineering.* China 2, 146-150, 2006.

Shimanskii, S.B., Strelkov, B.P., Anan´Ev, A.N., Lyubishkin, A.M., Iijima, T., Mochizuki, H., Kasai, Y., Yokota, K., Kanazawa, J. Acoustic method of leak detection using high-temperature microphone. *Atomic Energy*. vol. 98, no 2: 89-96, 2005.

Verde, C. Visairo, N. Gentil, S. Two leaks isolation in a pipeline by transient response. *Advances in Water Resources*. 2007.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Timed Automata Models for Batch Scheduling with Sequence-Dependent Changeovers

Subanatarajan Subbiah[a], Sebastian Engell[a]

[a]*Process Dynamics and Operations Group,*
Department of Biochemical and Chemical Engineering,
Technische Universität Dortmund, Emil-Figge Straße 70, 44227 Dortmund, Germany.

Abstract

The standard approach to solve scheduling problems in the process industries is to use mathematical model formulations such as MI(N)LP. Recently, an alternative approach that has gained attention is to use reachability analysis for timed automata (TA) to solve such problems. In this contribution, we discuss an application of the TA based framework to model and solve batch scheduling problems with sequence-dependent changeovers. The resources and the jobs are modeled as sets of interacting TA in a modular fashion and are composed to form a global automaton which represents the complete model of the scheduling problem. The solutions for the scheduling problem are computed by performing a cost-optimal reachability analysis in the global automaton.

Keywords: Multi-product batch plants, sequence-dependent changeovers, scheduling.

1. Introduction

In multi-product and multi-purpose batch processing industries, where a variety of products have to be produced with scarce available resources, the problem of scheduling with sequence-dependent changeovers frequently arises. The processing units require changeover times to switch the production from one product to another. In the presence of significant sequence-dependent changeovers, the utilization times of the processing units are strongly influenced by the sequence in which the products are produced. In this case it is necessary to have efficient scheduling models that include setup times and sequence-dependent changeover times. Most of the solution approaches proposed in the last years solve such problems by modelling them as mathematical programming formulations (MILP or MINLP) and applying commercial solvers to solve them. The representation of time in the models plays a crucial role for the performance of the solvers. In [1], the authors propose a MILP formulation in discrete-time where the key feature is to use specific variables for each unit to track the unit-operation events. Continuous time formulations based on single-time grid, multiple-time grid, and slot based approaches can be seen in [2], [3] and [4]. In [5] the authors present a rule-based heuristic using a genetic algorithm to solve single-stage multiproduct batch processes with jobs up to 200 orders. For exact approaches such as MI(N)LP and CP the application is hindered by the effort needed to formulate mathematical models and requires experience in algebraic modelling.

An alternative approach to model scheduling problems is to use the framework of timed automata and to solve the optimization problem using reachability analysis. Timed Automata (TA) are finite state automata extended by the notion of clocks to model discrete event systems with timed behavior. The formalism of TA has been

originally proposed in [6] and has been extended with the notion of costs, referred to as priced or weighted timed automata in [7]. Previous work on the TA based approach to scheduling problems with makespan minimization on hard job shop benchmarks were reported in [8] and [9]. A particular appeal of this approach is the modular and partly graphical modeling which enables inexperienced users to build models. Another advantage is the availability of powerful search algorithms that can be modified and extended for special purposes.

1.1. Background of Timed Automata

A short and informal definition of timed automata is given here, for complete definition of the syntax and semantics please refer to [7]. A timed automaton is defined by a tuple. $\boldsymbol{A} = (\boldsymbol{L}, \boldsymbol{C}, \boldsymbol{\Theta}, \boldsymbol{inv}, l_0, \boldsymbol{F})$ in which:

- $\boldsymbol{L}$ represents the finite set of discrete locations, with l_0, $\boldsymbol{F} \in \boldsymbol{L}$, where l_0 represents the initial location and $\boldsymbol{F}$ represents the set of final locations.
- $\boldsymbol{C}$ represents the set of clocks assigned to the TA.
- $\boldsymbol{\Theta} \subset \boldsymbol{L} \times \boldsymbol{\gamma} \times \boldsymbol{Act} \times \boldsymbol{U}(\boldsymbol{C}) \times \boldsymbol{L}$ represents the set of transitions between the locations where, $\boldsymbol{\gamma}$ is a set of guards specified as conjunctions of constraints of the form $c_i \otimes n$ or $c_i - c_j \otimes n$, where $c_i, c_j \in \boldsymbol{C}$, and $\otimes \in \{\leq, ==, \geq, <, >, \neq\}$ and $n \in \mathbf{N}$.
- $\boldsymbol{Act}$ represents the set of actions (e.g. invoking a new event or changing the value of a variable) while a transition is fired.
- $\boldsymbol{U}(\boldsymbol{C})$ represents the set of clocks that are reset to zero after an enabled transition fires. A transition between a source location l and target location l' with a guard $g \in \boldsymbol{\gamma}(\boldsymbol{C})$, performing an action $a \in \boldsymbol{Act}$ and resetting the clocks $r \in \boldsymbol{U}(\boldsymbol{C})$ is denoted by (l, g, a, r, l'). A transition can occur only when the guard conditions are satisfied and the invariant conditions of the target location evaluate to true.
- $\boldsymbol{inv}$ represents a set of invariants that assign conditions for staying in locations. The invariant conditions must evaluate to true for the corresponding location to be active. The automaton is forced to leave the location when the invariant evaluates to false.

In this contribution, we discuss the application of the TA framework to model and to solve batch scheduling problems with sequence-dependent changeovers. Modelling a scheduling problem using the TA framework is performed in a modular fashion. The resources and the jobs are modelled individually as sets of interacting TA. The interactions between the sets of automata are established by synchronizing the transitions between the automata. Synchronized transitions have the same synchronization labels and fire only simultaneously in the corresponding automata. The sets of interacting automata are composed using the procedure named *parallel composition* thereby forming one composed global automaton. Usually, the parallel composition is only performed on-the-fly to reduce space complexity of the reachability analysis. The composed automaton represents the complete model of the problem and a cost-optimal reachability analysis is performed to derive the schedules. The reachability analysis starts from the initial location of the composed automaton that represents the initial state of the system and evaluates the successor states created by a successor relation. This enumerative process is continued until the specified final target location is reached with minimal cost. In the search, various reduction techniques can be used in order to explore the state space quickly [8, 9].

2. Case study

The TA based framework explained above is applied to the medium size ice-cream manufacturing plant described in [10]. The case study is a two-stage process where 8 different end-products (*A* - *H*) have to be manufactured in batches. A schematic representation of the process is shown in Fig. 1. In the first stage, the raw materials stored in the warehouses are transported to the mixing department where they are processed according to the product recipes and stored in intermediate storage units. The intermediate products are cooled down and after a minimum waiting period has elapsed, the frozen products are transferred to the packing lines. In the second stage the products are packed and delivered to the customers. The intermediate products are unstable and have a maximum shelf time period within which they should be transferred to the packing lines from the storage units.

The plant consists of a single processing line with a production rate of 4500kg/Hr and 6 intermediate storage vessels (S_1 - S_6) in the first stage. The vessels S_1 and S_2 have maximum capacity of 8000kg each and the vessels S_3 - S_6 have maximum capacities of 4000kg each. The second stage consists of 2 packing lines P_1 and P_2 with production rate depending on the product packed. A routing constraint exists such that the two 8000kg vessels are coupled only to packing line P_1 and the four 4000kg vessels are coupled only to packing line P_2. The products (*A* -*D*) are packed only in packing line P_1 and products (*E* -*H*) are packed only in packing line P_2. Thus products (*A* - *D*) can be stored in storage vessels S_1 and S_2 only and products (*E* - *H*) can be stored in storage vessels S_3 - S_6 due to the routing constraint stated above. For the rest of the paper we consider that one batch of product measures 4000kg and the task durations are considered in minutes. Different instances of the case study are considered where production orders with different amounts of various products have to be produced within the minimum makespan, where a production order is equal to one batch of the corresponding product. The main features of the case study are: (a) the processing units in the first stage and the packing lines in the second stage are subject to sequence-dependent setup times and changeover procedure, (b) the intermediate storage units are limited have to be shared and (c) the intermediate products are unstable and have a minimum waiting time and a maximum shelf-life period, introducing a timing constraint between the processing and the packing task.

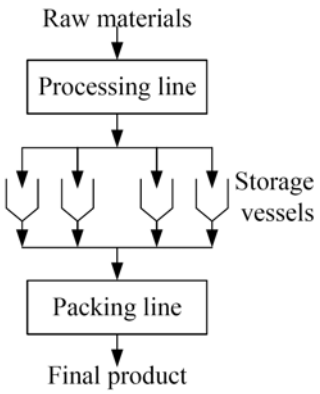


Figure 1. Process flow diagram

3. TA model for the case study

This section explains the framework to derive a TA model for the scheduling problem by constructing interacting TA for the recipes and for the resources.

3.1. Recipe automata

For each production order defined by its recipe number, one *recipe automaton* and a clock c_i are used. The automaton for recipe *A* is shown in Fig. 2. Each operation is defined by a *waiting* location and by an *executing* location connected by transitions. Starting the execution of the mixing task by occupying the processing line is represented by the transition α_{1A} between the locations *Wait Proc_A* and *Exec Proc_A*. Finishing the execution of the task in the processing line and transferring the intermediate material to one of the storage units S_1 or S_2 is modeled by the transition

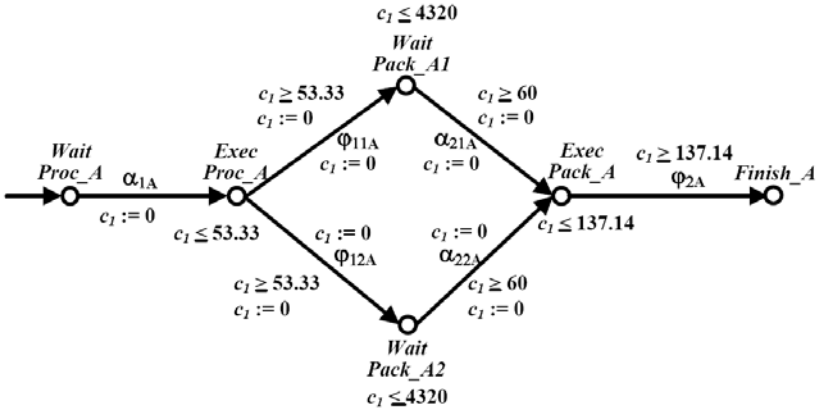
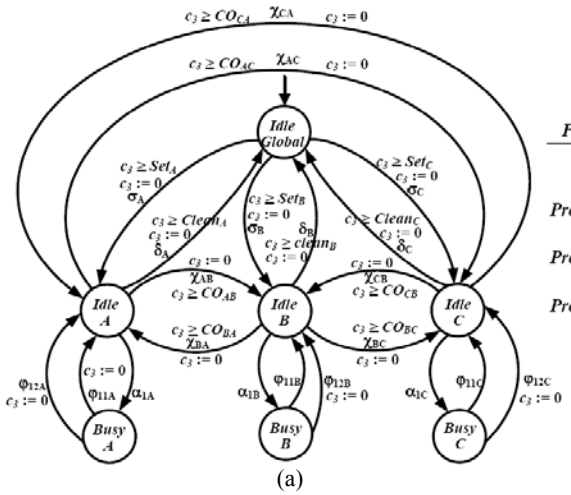


Figure 2. Recipe automaton for product A

φ_{11A} and φ_{12A}, respectively. Locations *Wait Pack_A1* and *Wait Pack_A2* represent that the intermediate material is stored in the storage unit S_1 and S_2, respectively and waiting to be packed in the packing line P_1. Starting the execution of the packing task in P_1 by consuming the intermediate material from the storage unit S_1 and S_2 is represented by α_{21A}, and α_{22A}, respectively. The additional location *Finish_A* represents the completion of the production order to produce one batch according to recipe *A*. The clock c_1 is introduced to model the timing behavior of the recipe. The invariants in the execute locations of the tasks force the automaton to leave the location once the task durations have expired. The guard conditions on the transitions that represents finishing of tasks which are labeled with φ ensures that the task is executed for the corresponding duration only. The clocks are reset to zero at every transition. The recipes *A*, *B*, *C* and *D* have the same automaton structure as the recipe structure is identical. Similarly the recipes *E*, *F*, *G* and *H* share the same automaton structure.

3.2. Resource automata

For each resource, a separate resource automaton with a clock is created by defining locations *idle* and *busy* for all possible configurations and operations which the resource executes. The automaton for the processing line with tasks representing processing of recipes *A*, *B* and *C* is shown in Fig. 3(a). The changeover table shown in Fig. 3(b) defines the changeover times (e. g. a minimum time period of CO_{AB} is required to changeover from configuration *A* to configuration *B*).



(a)

From \ To	idle	Product A	Product B	Product C
idle	-	Set_A	Set_B	Set_C
Product A	$Clean_A$	-	CO_{AB}	CO_{AC}
Product B	$Clean_B$	CO_{BA}	-	CO_{BC}
Product C	$Clean_C$	CO_{CA}	CO_{CB}	-

(b)

Figure 3. Resource automaton for the processing line and the changeover matrix.

Transitions that represent the setting up of the resource from the global configuration to a particular configuration are labeled by σ and transitions that represent resetting or cleaning of the resource from a configuration back to the global configuration are labeled by δ. Transitions that represent changeovers between any two non-identical configurations are labeled by χ. The resource automata and the recipe automata interact by synchronized transitions using the synchronization labels α and φ. In the resource automaton, the α labeled transitions represent allocation of the resource and the φ labeled transitions represent the release of the resource. Guard conditions based on *clock* valuations on the setup transitions ensures that the transition can take place only after the respective defined setup time. Similarly, guard conditions on the cleanup transitions and changeover transitions ensure that the corresponding transition is enabled only after a minimal reset time or a minimal changeover time.

3.3. Storage automata

For each common intermediate storage vessel, a *storage automaton* is created with a shared integer variable *V*. The automaton for the storage vessel S_1 which is used as a common intermediate storage unit in the recipes *A*, *B*, *C* and *D* is shown in Fig. 4. The integer variable V_1 represents the number of batches of the material present in S_1 and b_i represents the amount of product *i* processed. The maximum capacity of the storage unit is depicted as (V_1). The *storage automaton* consists of an *Empty* location and, for each possible material that can be stored in the storage vessel, a *Filled* location is defined.

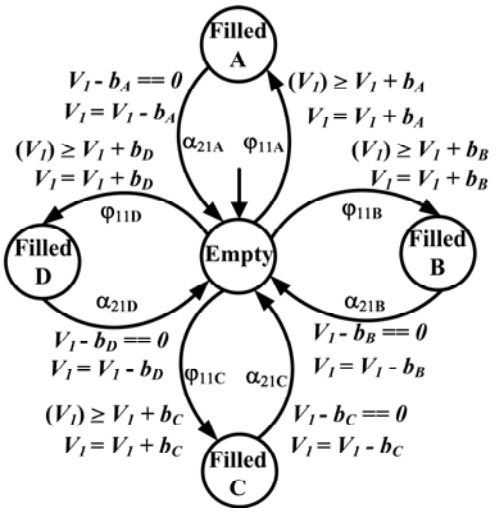


Figure 4. Storage automaton for the vessel S_1

The *storage automata* interact with the resource automata and recipe automata using the synchronization labels and φ, where φ represents allocation of the storage unit and α represents release of the storage unit.

4. Computational experiments and results

The plant description with information on the resources, separate recipe files for 8 different recipes in the form of RTN and a table of production orders with reference to recipes are given as input to TAOpt, a tool developed at the Process Dynamics and Operations group [9]. From these inputs, TAOpt creates sets of recipe automata and resource automata in a modular fashion and the reachability analysis is performed on-the-fly. The search algorithm used to explore the reachability tree is a combination of depth-first search and best-first search. Various reduction techniques such as weak non-laziness and sleep-set method were used to reduce the search-space. A detailed description of the reachability analysis and of the search-space reduction techniques can be found in [9]. In addition to the above mentioned reduction techniques, based on assumptions on the setup times, cleanup times and changeover times for a resource with possible configurations *i*, *j* and *k*, partial traces in the reachability tree that lead to sub-optimal schedules are pruned. The assumptions on the setup times, changeover times and cleanup times for configurations *i*, *j* and *k* are as follows:

Table 1. Results for the test on makespan minimization: (T_{cpu}) computation time in CPU sec., (C_{max}.) makespan in minutes, (Created nodes) total number of nodes created to reach the solution, (Visted nodes) total number of nodes explored to reach the solution.

Orders	First feasible solution				Best feasible solution			
	T_{CPU}	Nodes		C_{max}	T_{CPU}	Nodes		C_{max}
		Visited	Created			Visited	Created	
10	0.98	37255	68240	1667.61	1.01	37984	69676	1607.21
20	2.24	73706	138425	2452.84	2.26	74382	139765	2392.84
32	5.23	109095	225957	4333.75	8.63	109154	226023	4273.72
46	6.32	112264	208412	5506.13	44.73	545750	1403457	5446.13
48	6.40	118452	216790	5780.41	24.63	405391	979564	5720.41

- $Setup_i + Changeover_{ij} > Setup_j$
- $Changeover_{ik} + Cleanup_k + Setup_j > Changeover_{ij}$
- $Cleanup_i + Setup_j > Changeover_{ij}$
- $Changeover_{ij} + Changeover_{jk} > Changeover_{ik}$

The TA based modeling and solution approach was applied to various problem instances of the case-study ranging from 10 to 48 orders. The computational equipment used for the tests is a 3.06 GHz Xeon machine with 2 GB memory. For all tests the computation time was restricted to 3000 CPU seconds and the node limit was restricted to 9 million to avoid memory overflow. The results obtained for various instances are shown in Table 1. For all instances a feasible schedule could be obtained within 7 CPU seconds. This shows the TA based approach can be used in online and reactive scheduling where it is important to derive good feasible schedules within short computation times. The node limit of 9 million caused the termination in all experiments. We observed that the search scheme initially explored sub-optimal regions of the solution space. We therefore plan to use rule-based strategies in the search algorithm to guide the exploration of the solution space to promising regions.

5. Summary and outlook

This work demonstrates an application of the TA based approach to model and to solve the problem of scheduling batch processes with resources subject to sequence-dependent changeovers. Various other objective functions, such as minimization of changeover costs, can also be accommodated in the proposed approach by extending the automata with the notion of costs using Priced TA. Future work will investigate problems with various other objective functions and to embed rule-based techniques into the reachability analysis of TA.

The authors gratefully acknowledge the financial support from the NRW Graduate School of Production Engineering and Logistics at TU Dortmund.

References

1. Jeffrey D. Kelly and Danielle Zyngier, 2007, *Ind. Eng. Chem. Res.*, 4964 - 4973.
2. Chi-Wai Hui, et al. , 2000, *Comp. Chem. Engg.*, 2705 - 2717.
3. Pedro M. Castro, et al. , 2006, *Ind. Eng. Chem. Res.*, 6210 - 6226.
4. M. Erdirik-Dogan and I. E. Grossmann, 2008, *Ind. Eng. Chem. Res.*, 1159 - 1183.
5. Yaohua He, Chi-Wai Hui, 2008, *Comp. Chem. Engg.*, 3067 - 3083.
6. R. Alur and D.L. Dill, 1994, *Theoretical Computer Science*, vol. 126, 183-235.
7. K. Larsen et al. , 2001, *Proc. 13th Int. Conf. on Computer Aided Verification,* 493 - 505.
8. Y. Abdeddaim and O. Maler, 2006, *Theoretical Computer Science,* Vol. 354, 272 -300.
9. S. Panek, S. Engell, and O. Stursberg, 2006, *Control Engg. Practice*, 1183 - 1197.
10. Peter M. M. Bongers et al. , 2006, *Proc. 16th ESCAPE and 9th PSE,*1917-1922.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Use of Event Correlation Analysis to Reduce Number of Alarms

Fumitaka Higuchi,[ac] Ichizo Yamamoto,[a] Tsutomu Takai,[b] Masaru Noda,[c] Hirokazu Nishitani [c]

[a]*Idemitsu Kosan Co., Ltd, 2-1, Anesakikaigan, Ichihara, Chiba 299-0194, Japan*
[b]*Yamatake Corporation, 1-12-2 Kawana, Fujisawa 251-8522, Japan*
[c]*Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma, Nara 630-0192, Japan*

Abstract

Event correlation analysis was applied alarm data at an Idemitsu Kosan ethylene plant in Japan to reduce the number of plant alarms. This is a data-mining method that detects statistical similarities among discrete occurrences of alarms or operations. By grouping correlated events based on degree of similarity, a policy for reducing alarms can be designed more easily than by analyzing individual alarms and operations. By using event correlation analysis, we were able to divide event data of alarms or operations of the ethylene plant into a limited number of groups. The results of the analysis were helpful for identifying unnecessary alarms, such as sequential alarms, buried in a lot of noisy plant data. This method is useful for reducing the number of alarms.

Keywords: Alarm Reduction, Top Ten Approach, Event Correlation Analysis, Ethylene Plant, EEMUA

1. Introduction

With the advance of distributed control systems (DCS) in the chemical industries, it has become possible to install many alarms cheaply and easily. While most alarms help operators detect and identify faults, some are unnecessary. A poor alarm system may cause alarm floods and nuisance alarms, which reduce the ability of operators to cope with plant abnormalities because critical alarms are buried in a lot of unnecessary alarms.

Since 1998, Idemitsu Kosan has actively worked on reducing the number of unnecessary alarms at its plants using the top-ten worst alarm method. This method collects data from event logs of alarms generated during operation and makes a list of frequently generated alarms. Then, the alarms are reviewed one by one, starting from the most frequently generated alarm, and the root causes that triggered them are identified. This method is effective for early stages of alarm reduction. In fact, average alarm frequency standards proposed by the EEMUA (2007) are achieved in some plants. However, when the proportion of the worst ten decreases, it becomes difficult to achieve effective improvements. For this reason, a novel alarm reduction method is required for further plant alarm reduction.

In this study, event correlation analysis was applied to alarm reduction for the Idemitsu Kosan ethylene plant in the Chiba complex. The event correlation analysis, originally proposed by Nishiguchi and Takai (2005), is a data-mining method that detects statistical similarities among alarms and operations. By grouping correlated

alarms and operations based on the degree of similarities, a strategy for reducing alarms can be found more easily than by analyzing individual alarms. This helps us to identify sequential alarms, repeated alarms caused by routine operations, and so on within a huge amount of alarm and operation data.

2. History of alarm reduction activities by Idemitsu Kosan

2.1. Ethylene plant in Chiba complex

Idemitsu Kosan started the ethylene plant at the Chiba complex in 1985. The ethylene plant is operated by two board operators using a DCS. The total numbers of tags in the DCS are 775 for process control and 2461 for process monitoring. The alarm data collection system was implemented at the ethylene plant in 2003. Figure 1 shows event generation frequencies of alarms and operations per board operator in the ethylene plant over ten minutes in 2003. In the ethylene plant, furnace decoking operations cause many repeated alarms.

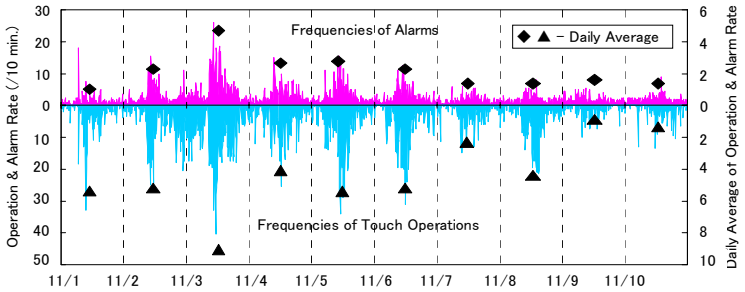


Fig. 1 Frequencies of alarms and operations in the ethylene plant

2.2. Alarm reduction by top-ten worst alarm method

Idemitsu Kosan applied the top-ten worst alarm method to the ethylene plant as a part of its total productive maintenance (TPM) activities in 2003. Figure 2 shows the ratio of the top ten alarms to all 16271 alarms generated in November of 2008 in the ethylene plant.

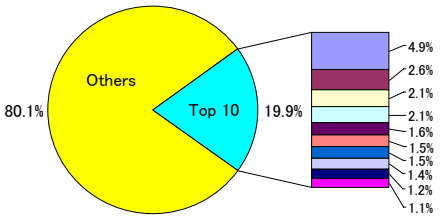


Fig. 2 Ratio of top ten alarms to all alarms generated in ethylene plant

We reviewed alarms one by one from the top of the list and reduced the number of alarms by changing alarm settings, retuning PID controller parameters, utilizing an alarm suppression system, and so on. Implementation of the programmable logic controller, where alarm settings are automatically changed according to plant states, drastically decreased a large number of alarms generated by unsteady state operations.

The top-ten worst alarm method was effective when the ratio of the top ten alarms was more than 20%. Figure 3 shows the transition of the frequency of alarm generation in the ethylene plant from 2003 to 2008. Though we continued the alarm reduction activities by the top-ten worst alarm method, the alarm generation frequency has not

been reduced since 2003. Because the ratio of each alarm in the top-ten worst alarm list became very small, it was difficult to achieve further effective improvement. Thus, a novel alarm reduction approach was desired to reduce the number of alarms.

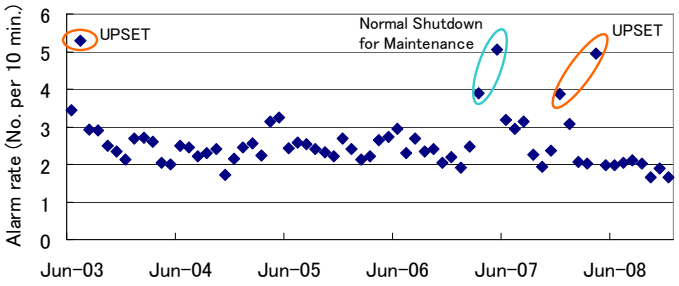


Fig.3 Average alarm rate in ethylene plant per board operator

3. Event correlation analysis

Event correlation analysis is a knowledge extraction (data mining) method intended event data composed of discrete events and the time they occur. This method uses event data from the plant to quantify the degree of similarity and time lag between two events by evaluating the cross correlation function. Then, the event tags are grouped using hierarchical clustering based on the degree of similarity.

3.1. Event log data

Plant data used in this method consists of occurrence time and tag name of an alarm or an operation, which we call an "event" hereafter. The plant data is converted into sequential event data $s_i(k)$. When event i occurs between $Ts + (k\text{-}1)\Delta t$ and $Ts + k\Delta t$, $s_i(\mathrm{k}) = 1$. Here, Δt is the unit time, k denotes the discrete time, and Ts is the time of the first event. Figure 4 shows an example of the event data and the sequential event data $s_i(k)$.

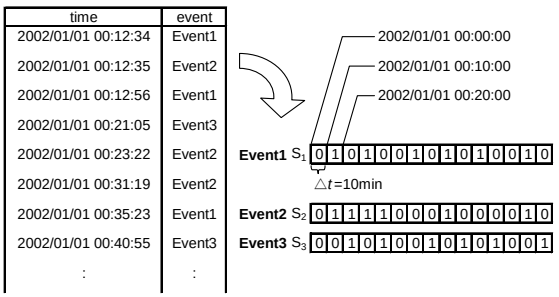


Fig. 4 Event log data

3.2. Cross correlation function

The cross correlation function between two sequential event data $s_1(k)$ and $s_2(k)$ for time lag m is defined by Eq. (1). Figure 5 shows the cross correlation function between two sequential event data. Here, K is the maximum time lag and T is the time period of the event data.

$$c_{12}(m) = \begin{cases} \sum_{n=1}^{T/\Delta t - m} s_1(n) s_2(n+m) & 0 \le m \le K \\ c_{21}(-m) & -K \le m < 0 \end{cases} \tag{1}$$

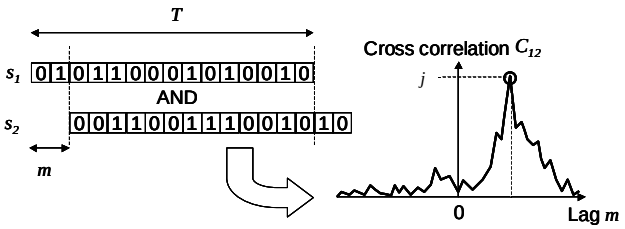


Fig. 5 Cross correlation function between two events

3.3. Definition of similarity

The probability of the occurrence of event i, which is denoted by p_i, is calculated by Eq. (2).

$$p_i \cong \frac{\sum_{k=1}^{T/\Delta t} s_i(k)}{T/\Delta t} \tag{2}$$

Assuming that any two events s_1 and s_2 are independent, the probability that two events occur simultaneously can be calculated by Eq. (3).

$$p_{12} = p_1 \times p_2 \tag{3}$$

Then, the probability distribution that two events occur simultaneously j times with time lag m can be expressed by the following binomial distribution. Here, j indicates the maximum value of $c_{12}(m)$.

$$P(c_{12}(m) = j) = {}_{T/\Delta t}\mathrm{C}_j \, p_{12}^{j} (1 - p_{12})^{T/\Delta t - j} \tag{4}$$

If p_{12} is very small, Eq.(4) can be approximated by Poisson distribution.

$$P(c_{12}(m) = j) \cong \frac{\nu^j e^{-\nu}}{j!}, \qquad \nu = T/\Delta t \cdot p_{12} \tag{5}$$

The probability that two events occur simultaneously at least j times during time period T with time lag m is given by Eq. (6).

$$P(c_{12}(m) \geq j) \cong \sum_{i=j}^{T/\Delta t} \frac{\nu^i e^{-\nu}}{i!} \cong 1 - \sum_{i=0}^{j-1} \frac{\nu^i e^{-\nu}}{i!} \tag{6}$$

Therefore, the total probability that two independent events occur simultaneously more than j times with time lag m ($-K \leq m \leq K$) is given by Eq. (7).

$$P(c_{12}(m) \geq j \,|\, -K \leq m \leq K) \cong 1 - \left(\sum_{i=0}^{j-1} \frac{\nu^i e^{-\nu}}{i!} \right)^{2K+1} \tag{7}$$

Finally, the similarity S_{12} between two events s_1 and s_2 is calculated by Eq. (8).

$$S_{12} = 1 - P(c_{12}(m) \geq j \,|\, -K \leq m \leq K) \tag{8}$$

Larger similarity means stronger dependency or a closer relationship between the two events.

3.4. Grouping of events

After calculating similarities between all combinations of any two events in the plant log data, all events are classified into groups. By grouping events, it becomes possible to stratify and visualize the distance between events. In this study, the hierarchical clustering method was used.

3.5. Alarm reduction method

The following method is applicable to reduce unnecessary alarms and operations found by the event sequence analysis.

(1) Find a root cause of unnecessary alarms and operations based on the cause-effect analysis. Eliminating the root cause reduces unnecessary events.
(2) Find the first alarm in the sequential alarms, and retune settings of alarms following the first alarm. A state-based alarm is also effective.
(3) A programmable logic control is useful for reducing repeating alarms generated by routine operations.

4. Case study

To verify the effectiveness of the event correlation analysis method for reducing unnecessary alarms and operations, the analysis was applied to the Idemitsu Kosan ethylene plant in the Chiba complex. The plant log data gathered in one month included 1771 tags of alarms and operations. The total number of generated alarms and operations was 51640.

4.1. Results of grouping of alarms and operations

By applying the event correlation analysis to the plant log data, 1771 tags of alarms and operations were classified into 657 groups, where Δt, K, and the minimum threshold for identifying similarity between two events were set as 10 min., 30 min., and 0.95, respectively. Table 1 shows the result of the event correlation analysis. The total number of alarms and operations in the top 20 groups accounted for 45% of all events.

Table 1 Top 20 worst groups

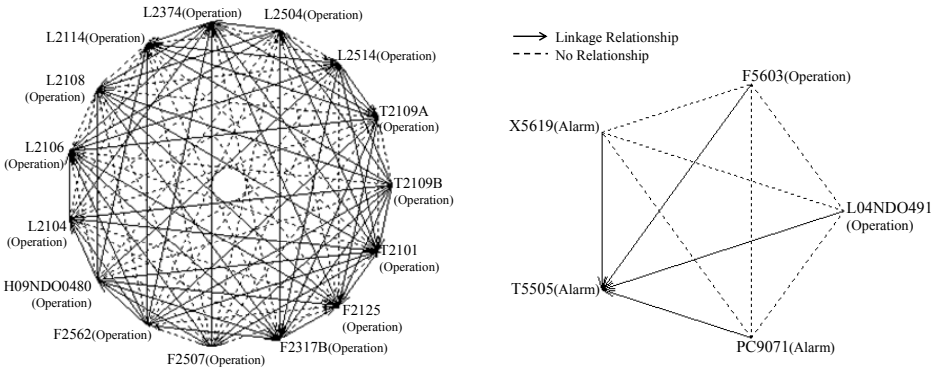
		Number of Events			Number of Tags		
No.	Group No.	Total	Alarm	Operation	Total	Alarm	Operation
1	163	4198	0	4198	15	0	15
2	358	3570	3570	0	3	3	0
3	390	1314	0	1314	8	0	8
4	225	1223	83	1140	10	4	6
5	388	1158	47	1111	5	1	4
6	441	1072	141	931	13	7	6
7	179	1034	213	821	9	4	5
8	318	1003	0	1003	4	0	4
9	234	938	938	0	1	1	0
10	160	910	62	848	5	1	4
11	54	848	25	823	8	1	7
12	40	800	800	0	4	4	0
13	293	799	472	327	9	8	1
14	600	731	28	703	8	2	6
15	180	702	77	625	5	1	4
16	391	675	220	455	9	7	2
17	309	659	169	490	5	3	2
18	25	626	171	455	8	6	2
19	294	611	93	518	5	2	3
20	181	611	36	575	8	2	6

Figure 6 (a) shows 15 events and their relationships in the first group in Table 1. The total number of events in this group accounted for 8% of all generated events. According to an interview with the operators in charge, the events in this group occurred during furnace decoking. This means that 8% of the events could be reduced by improving the decoking operations. This indicates that the event correlation analysis is effective for reducing many alarms and operations.

Figure 6 (b) shows relationships among multiple alarms included in the seventeenth group in Table 1. Those alarms occurred in sequence. To reduce these sequential alarms, it may be effective to change their set points.

4.2. Identification of routine operations

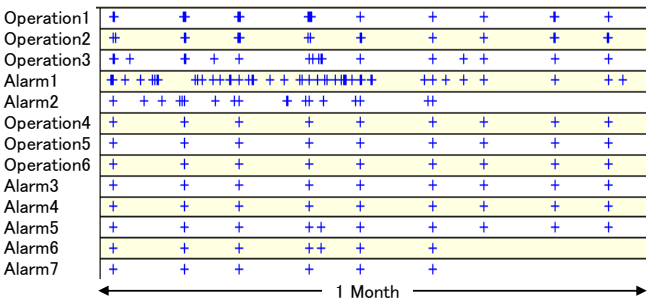
When many operations are included in a group, these may be routine operations. Table 2 shows the points when alarms and operations occurred in the sixth group in Table 1. As shown in Table 2, alarms and operations in this group appeared synchronously. According to interviews with operators, these events occurred once or twice a week and were caused by routine operations. We automated these routine operations by a programmable logic controller. As a result, 70% of the alarms caused by these operations were eliminated.



(a) The top group (b) A group containing multiple alarms

Fig. 6 Results of event correlation analysis

Table 2 Generation patterns of events in a group



5. Conclusion

Event correlation analysis was applied to the Idemitsu Kosan ethylene plant in Japan. Using this method, we effectively identified unnecessary alarms and operations within a lot of event data. This method is useful for reducing the number of alarms.

References

The Enginnering and Equipment Materials Users' Association (EEMUA), 2007, Alarm Systems A Guide to Design, Management and Procurement, EEMUA, London

J. Nishiguchi and H. Tsutsui, 2005, A New Approach to Process Alarm Reduction Using Statistical Point Processes, *SICE Annual conference 2005*, pp.443–448

T.Takai, 2007, Alarm Management in Chemical Plants, *Human Factors,* Vol.12, No.1, pp.10–23

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Control strategy with distributed action for minimization of transients in distillation column

Leandro O. Werle[a], Cintia Marangoni[a], Joel G. Teleken[a], Claudia Sayer[a], Ricardo F. Machado[a*]

[a]*Santa Catarina Federal University, Chem. Eng. Dep. University Campus, Mail Box: 476, Florianópolis-SC, Brazil, Zip Code: 88010-970 *Email: machado@enq.ufsc.br*

Abstract

Distillation columns require well adjusted control systems to quickly reject disturbances. To assist this purpose, a new distributed control strategy using electrical resistances on intermediate plates of a distillation column was proposed. Usually the control of distillation columns is accomplished centralized at the bottom and at the top. In the present work the effect of the distributed control is evaluated, through the addition of thermal energy using an electric resistance located at an intermediate plate, when a disturbance is applied in the composition of the feeding stream. In order to validate the proposed strategy, a pilot distillation unit composed of 13 trays was used, processing an ethanol-water mixture with a feed flow rate of 300 L/h. The objective of this paper is a reduction in the operation transients. In addition, a thermal balance was used to verify if distribution of thermal energy occurs between the bottom of column and the plate where the distributed control is applied. The results showed that the introduction of the distributed heating along the column allows faster dynamics, representing a valid option for the reduction of transients.

Keywords: distillation column, distributed action, minimization of transient.

1. Introduction

The distillation process is widely used in industries, especially in petroleum refineries. The dynamic behavior of the process is intrinsically non linear and variables are highly coupled, being an essentially multivariable problem and with operation restrictions (Miladi and Matjaba, 2004).

Distillation columns require well adjusted control systems to quickly reject disturbances. However, it is known that well designed and adjusted control systems are not enough to eliminate the operation transients. The transient formation in a distillation column occurs when the process is disturbed or when an external factor induces the modification of the operation point of the unit. In the first case there are factors such as variable coupling, nonlinearities, time-delay, high time constants and process constrains. Whereas the second case involves aspects like the mixture to be distillated, feed composition changes and operation transitions due to market alterations.

Procedures of campaign changes are problematic in distillations, requiring to drive the system to a new operation regime as fast as possible to minimize the formation of out of specification products. Although several researches involve quite complex control approaches, only few works approach the matter of transient minimization. Usually, the distillation column control is carried out in a centralized way in the variables on the bottom, top and lateral outlets, generating a high transition time when the process is

disturbed. In previous works (Marangoni et al., 2005) a distributed control strategy was proposed, by using electrical resistances as an energy source on the trays.
It is known that in the operation of distillation columns the largest cost is in the amount of applied energy. In this study we chose to work with electrical resistance and a plate heat exchanger (reboiler). It should be emphasized that, instead of conducting the heat through the electrical resistance, it could do so through the use of heat exchangers or condensers, to perform the heat exchange. According to Agrawal and Herron (1998) the appropriate use of reboiler and intermediate heat exchangers can promote significant reductions in the cost of operating of a system of distillation. The authors perform a detailed analysis of some configurations using such equipment.
In addition, several studies involve heating at intermediate trays from the point of view of diabatic distillation (Rivero, 2002), aiming energetic saving. Some columns in petroleum refineries use circulating refluxes in order to promote a better separation. Nevertheless, in both of these applications intermediate heating is not used as part of the control loop. This paper analyzes the behavior of a distillation column when a disturbance is applied in the composition of the feed stream using two different approaches: conventional and distributed control strategy. A thermal balance was used to verify if in the second approach the thermal energy added to the column is distributed between the bottom of column and the plate where the new control strategy is applied.

2. Methodology

The experiments were carried out in a pilot distillation unit, composed of 13 trays, instrumented with a fieldbus digital communication protocol, processing an ethanol-water mixture with a feed flow rate of 300 L/h.

2.1. Experimental Unit

The pilot unit operates continuously, with the feed stream inserted in the 4^{th} tray, with a total height of 2.70m, built in modules with 0.15m height and 0.20m diameter.

2.2. Control System

The control configuration of the distillation column was formulated based on Nooraii et al (1999) research. The following control loops were defined: (1) bottom level control through the bottom product flow rate adjustment; (2) reflux accumulator level control by manipulating the top product flow rate; (3) feed flow rate control as a function of the adjustment of the same stream flow rate; (4) feed temperature control through the fluid flow rate adjustment in the heat exchanger of this stage; (5) last tray (distillate) temperature control by means of the manipulation of the reflux flow rate; (6) reboiler temperature control through the vapor flow rate in the heat exchanger of this stage and (7) temperature control of pre-defined stages of the column through the adjustment of the dissipated power in the tray electrical resistance.
The first, second and third loops represent the column mass balance (inventory) control. The fifth and sixth loops comprise the quality control – in this case represented by the temperature. When these two loops are used combined we called it conventional control. When these two loops are combined with the seventh loop described above, we consider it as the distributed strategy.
PID controllers were adjusted for the temperatures of the bottom and of the last tray, characterizing the conventional control. The controllers used for the bottom level and the feeding flow rate consist of a cascade system composed by a feedback PI and another anticipative (feedforward). To adjust the controllers the classical methods of Cohen-Coon, ITAE and Zieger-Nichols (Seborg et al., 1989) were applied. The distributed control was carried out by using an intermediate tray, tray 2.

2.3. Experimental Procedure

For this study disturbances were introduced in the feed composition. The performance of the distributed approach (control on the bottom, top and tray 2) was compared with the conventional configuration. Experiments were conducted with initial volumetric compositions of ethylic alcohol in the feed of 10% (v/v), the feed temperature was controlled at 92ºC, the feed flow rate was 300 L/h and the pressures on the bottom and top varied around 1.25 and 0.25 bar, respectively. Composition measurements were carried out during the experiments. The distributed heating was carried out by means of electrical resistances with maximal power of 3.5 kW each. For the identification of the most sensitive stage for the application of the distributed control sensibility analysis methods were used: successive plates, sensitivity symmetry and maximum sensitivity, as detailed in Marangoni (2005). The sensibility analysis was accomplished with base on the studies of Luyben (2006). The results showed that in the stripping section plates 2 and 3 were the most sensitive plates, and in the rectifying section trays was 5 and 7. Plates 3 and 5 are adjacent to the feeding plate, receiving disturbances directly in their flow rate, for this reason these plates were not selected for the distributed heating. Between the remaining "most sensitive" plates (2 and 7), plate 2 was chosen for application of the distributed control, since the stripping section is the region that demands more energy. This decision is also based on the operation mode of diabatic distillation columns in which the rectifying section receives heat, and the rectification section releases heat through heat exchangers. The choice of using a single plate for the distributed heating also aimed to avoid more than one simultaneous effects.

2.4. Energy Balances

An energy balance, shown in Equation 1, was used to calculate the heat inserted by the reboiler. Assuming that the losses to the environment are negligible, on the steady state:

$$Q_{reb} = (Q_c + Q_d + Q_b) - (Q_{res} + Q_f) \quad (1)$$

Where: Q_{reb} is the heat inserted into the bottom by the reboiler, Q_c is the heat withdrawn by the condenser, Q_d and Q_b are, respectively, the heats that come out of the column by the destilate and bottom product streams, Q_{res} is the heat added by the electrical resistance and Q_f is the heat of the feed stream, all in [kW].

3. Results

The disturbance applied in the feed composition was a positive step, the volumetric fraction of ethanol varied from 10% to 20% (v/v), with the addition of ethanol to the feed reservoir. Due to the largest volume of the reservoir the contributions of the streams composition of the top and bottom that return to the feed reservoir are small (negligible) in the composition of the feeding. It is important to mention that in this paper the reference values for the control loops were maintained after the introduction of the disturbance.

Figure 1a shows the effect of the disturbance on the control loop of the temperature of the reboiler. It can be observed that after the control system assimilates the disturbance, a temperature decrease occurs in both configurations, however the distributed control resulted in a lower decrease (0.5ºC), when compared to the conventional one (1.2ºC), this is explained by the contribution of the resistance to the formation the vapor phase.

In the same way as in the stripping section, the rectifying section is also affected by the disturbance of the feed composition. This effect can be observed in Figure 1b, the increase of the top temperature, observed for both configurations, is due to the thermal load added in excess by the reboiler in order to keep the reference values, since the operation points were not adjusted after the disturbance. The disturbance affected the

liquid-vapor equilibrium, leading the vapor to drag great amounts of water to the top of the column. This resulted in temperature increases for both approaches, due to the saturation of the reflux valve in both configurations (Figure 1b).

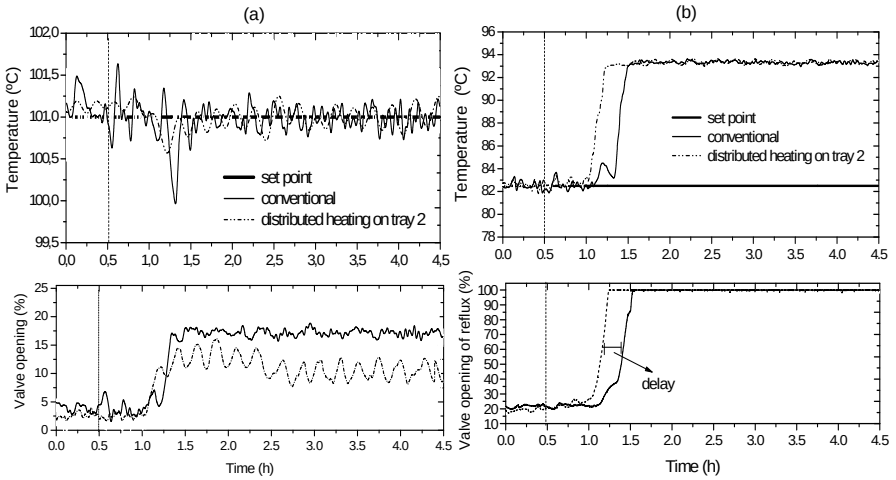


Figure 1 - Response of the temperature control loops: a) reboiler and b) last stage.

In the results shown in Figure 2, temperature of tray 2, it might be observed that 30 minutes after the application of the disturbance the temperature of tray 2 starts to decrease for the conventional configuration, due to the increase of the amount of the more volatile component.

In the distributed configuration, on the other hand, this decrease was not observed due to the instantaneous energy supply by the electrical resistance. The energy provided by resistance R2 to the system, increasing the vapor phase, may have favored the transport of the vapors rich in water to the top of the column, explaining the premature saturation of the reflux valve of the temperature control loop of the last stage (Figure 1b) and the dynamic differences between both configurations.

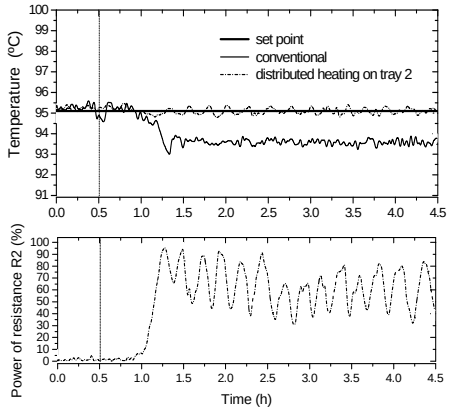


Figure 2 - Effect of the disturbance in the temperature control of tray 2.

Comparing these results with those of Marangoni (2005) a similar behavior might be observed for these three control loops analyzed when a disturbance was introduced in

the composition of the feed stream. Finally, the analysis of the behavior of the manipulated and controlled variables in Figures 1 and 2, shows that the introduction of a heating point distributed along the column improves considerably the performance of control loops.

The disturbance applied volumetric fractions of ethanol in feed stream, affects the fraction of this component in the stripping section before the rectifying section. After the alteration of the composition at the bottom the change is propagated through the vapor phase to the upper trays, establishing a new composition at the top, which in turn affects the composition at the bottom through the liquid that flows down the column. The equilibrium between the phases is reached after a few hours.

Figure 3 shows the derivative of the volumetric fraction of ethanol in the top. These results show that the distributed control is faster (around 1 hour) than the conventional control, mainly due to its different dynamics.

The high temperature reached at the top of the column led to very high reflux flow rates (Figure 1b), which in turn ceased the production of distillate. Therefore, the new steady state reached after the disturbance cannot be considered appropriate for the operation since it does not offer conditions for an adequate separation of the components. Showing, this way, that this kind of disturbance needs a new adjustment of the process operation point. According to Hurowitz et al. (2003), disturbances in the feed composition consist in the main challenge for distillation column dynamics.

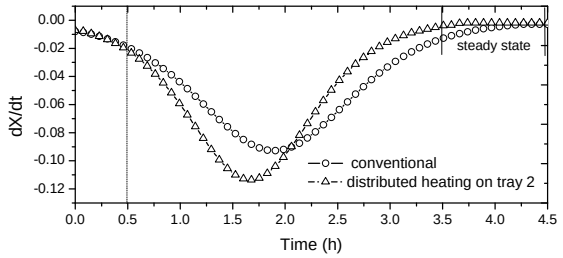


Figure 3 - Derivative of the volumetric fraction of ethanol in the top product.

The effect of this disturbance, increase of the fraction of the more volatile component in the feed stream, which leads to an increase of the fraction of this component in the bottom section, can be compared to the effect of a positive disturbance on the thermal load of the reboiler, leading to an increase of the vapor flux throughout the whole column, vaporization of the less volatile components and raise of the their fractions in the top, coupled with a temperature increase throughout the whole column.

The thermal energies of the streams, calculated by the global energy balance (Equation 1) are shown in Table 1. Comparing the absolute values of the added energies it might be observed that in the distributed control the energy added in the bottom in the conventional control is distributed between the bottom and the resistance of tray 2.

In the Figure 1a it might be observed that after the disturbance the bottom valve opening operated with smaller values in the distributed control compared to the conventional one, compensating this way, the energy added in plate 2. In the conventional approach, after the disturbance, the overall heat supply of reboiler increases 45.41 kW, while in the distributed approach the increase was only 43.34 kW, and 2.28 kW were added by the electrical resistances, showing this way the distribution of overall heat supply. This distribution of heat was also observed by Werle et al. (2007) when used the distributed control, applying a different type of disturbance (in the feeding flow rate).

Table 1 - Thermal energies (kW) calculated at steady state.

Conventional	Q_{reb}	Q_f	Q_{res}	Q_c	Q_d	Q_b
Before	9.80	21.71	----	9.75	0.34	21.41
After	55.21	20.68	----	55.17	1.11	19.61
Distributed healting						
Before	9.70	21.68	0.05	9.61	0.40	21.42
After	53.04	20.68	2.28	55.49	1.15	19.54

The great increase of the thermal energy of the reboiler, after the disturbance for both studied control configurations is a result of the increase of the ethanol concentration, keeping the feed stream temperature constant.

4. Conclusion

The distributed control approach led to a reduction of approximately 1 hour of the operation transient when compared to the conventional process. In addition, the distributed control also resulted in a less oscillatory behavior in most of the control loops. Both these advantages observed when the distributed control was used were obtained with nearly the same overall heat supply of the conventional approach, distributed along the column. Thus, the introduction of the distributed heating along the column has shown itself as a valid option for the transient reduction.

5. Acknowledgments

The authors are grateful for the financial support from the National Agency of the Petroleum (ANP), and the Studies and Projects Fund (FINEP), by means of the Human Resources Program of ANP for the Petroleum and Gas sector – PRH-34-ANP/MCT.

References

A. Noorai, J. Romagnoli, J. Figueroa, 1999, Process, identification, uncertainty characterisation and robustness analysis of a pilot scale distillation column. Journal of Process Control, v. 9, n 3, p. 247-264.

C. Marangoni, 2005, Implementação de uma Estratégia de Controle com Ação Distribuída em uma Coluna de Destilação. Tesis. Programa de Pós-Graduação do Dep. de Eng. Química e Alimentos, Univ. Federal de Santa Catarina, Florianópolis.

D. E. Seborg, T. F. Edgar, D. A. Mellicamp, 1989, Process Dynamics and Control, John Wiley & Sons, Singapore.

L. O. Werle, C. Marangoni, F. R. Steinmacher, R. C. Soares, P. H. H. Araújo, R. A. F Machado, C. Sayer, A. Bolzan, 2007, Uso de aquecimento distribuído em uma coluna de destilação: efeito da perturbação na vazão de alimentação. In: 4° Congresso Brasileiro de P&D em Petróleo e Gás - PDPetro.

M. M. Miladi, J. Mutjaba, 2004, Optimisation of design and operation policies of binary batch distillation with fixed product demand, Comp.&Chem. Eng. v.15, p.2377-2390

R. Rivero, 2002, Application of the exergy concept in the petroleum refining and petrochemical industry. Energy Conversion and Management, v. 43, p. 1199-1220.

R. Agraval, D.M. Herron, 1998, Intermediate reboiler and condenser arrangement for binary distillation column. AICHE Journal, v. 44, p. 1316-1324.

S. Hurowitz, J. Anderson, M. Duvall, J.B. Riggs, 2003, Distillation Control Configuration Selection, Journal of Process, v. 13, p. 357-365.

W. L. Luyben, 2006, Evaluation of criteria for selecting temperature control trays in distillation columns. J. Proc. Cont. v.16, p.115-134.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Experimental startup of a distillation column using new proposal of distributed heating for reducing transients

Leandro O. Werle[a], Cintia Marangoni[a], Joel G. Teleken[a], Claudia Sayer[a], Ricardo F. Machado[a*]

[a]*Santa Catarina Federal University, Chemical Eng. Dep. University Campus, Mail Box: 476, Florianópolis-SC, Brazil, Zip Code: 88010-970 *Email: machado@eng.ufsc.br*

Abstract

This paper proposes the use of distributed heating along the trays of a distillation column, which is important during transient periods, particularly during startup. The experiments were carried out in a pilot distillation column with 13 sieve trays processing a binary mixture composed of water and ethanol. It was verified that the use of a heating system that is able to act in a distributed way allows a reduction in the time required for the startup procedure. This process was chosen due to its characteristics of inertia, non-linear behavior, and long transient periods, which make the objective an even greater challenge since this process is hard to control with traditional methods. Most of the previously published studies on startup involve simulations, whereas the main contribution of this paper is that is provides experimental results.

Keywords: distillation column, startup, distributed heating.

1. Introduction

Distillation columns have been extensively studied and this type of research could be considered as complete (Eden et al., 2000). However, the control of a distillation column can be split into two different problems: steady state and dynamic control. Dynamic control is necessary when the column is operating out of its normal operation conditions, which generates strong nonlinearities in the process. Typical situations are the startup and shutdown operations, in which the plant is far from under its normal production conditions, and therefore these may display very disparate behavior.

The startup of a continuous distillation column presents a series of operational problems, as described by Kister (1979), which until now have been difficult to minimize. Most of the proposals address controlling this process phase. Fabro (2005) proposed the use of intelligent control techniques such as a neural networks, fuzzy systems and genetic algorithms for distillation column startups, allowing the desired steady state to be reached as quickly as possible. Along the same research lines a new on-line control algorithm has been proposed by Barolo et al. (1993), based on Generic Model Control, for improving the automatic startup of a binary distillation column. Scenna and Benz (2003) described multiple steady states for different initial conditions as starting points. As can be observed, several methods and configurations for distillation column startup have been proposed in the literature by different authors. In this study we try to make use of a new control strategy initially proposed by Marangoni (2005), which uses a distributed corrective action based on heating the distillation column trays, and evaluate this new configuration during the startup process. The startup of a distillation column can be considered in three distinct steps: the heating of

all stages, the introducing of a reflux stream in manual mode, and the changing of the control loops to automatic mode. The main goal of this study is to analyze these three steps of the column startup, which need special care and require enormous amounts of energy.

2. Methodology

The experiments were carried out in a pilot distillation unit, composed of 13 trays, instrumented with a digital fieldbus communication protocol, processing an ethanol-water mixture (sub-cooled liquid) with a feed flow rate of 300 L/h.

2.1. Experimental Unit

The pilot distillation unit operates continuously, with the feed stream inserted in the 4th tray, with a total height of 2.70 m, built in modules with 0.15 m height and 0.20 m diameter. A schematic diagram of the unit is shown in Figure 1.

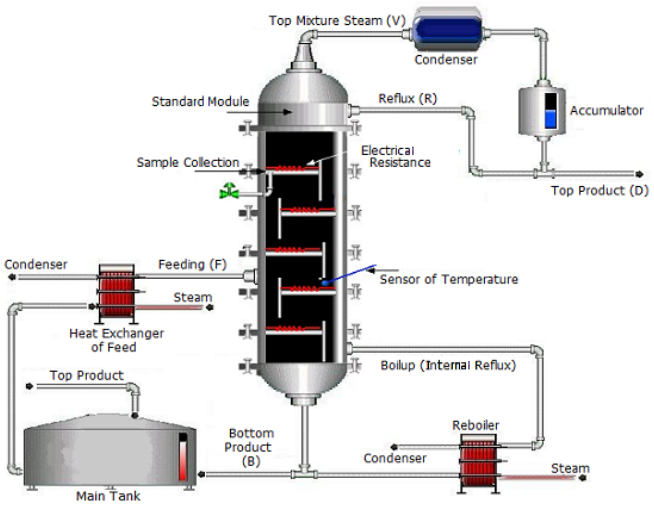


Figure 1 - General scheme of the equipment and pilot distillation unit.

The main feature of this distillation unit, which makes it different from the others, is the use of heating points supplied by electrical resistances distributed along the column, which decentralize the heat supply from the bottom.

2.2. Experimental Procedure

In the conventional approach, the column is operated with heat supply only at the reboiler, with the manual mode of the vapor valve opening fixed at 8% (20kW), as represented in Figure 3b. When the proposed configuration is applied, the heat is supplied at the bottom by the reboiler (with the same previous overall heat supply) and at an intermediate point, by way of the electrical resistance on the 2nd tray, started with 50% of its capacity (1.75kW) The procedure adopted for the column startup in the conventional operation is based on a study by Steinmacher et al. (2004). In this study the startup of the distillation column was considered to have three distinct steps: the heating of all stages, the introducing of the reflux stream in manual mode, and the changing of the control loops to automatic mode. In the first step, at the beginning of the procedure, volumetric compositions of ethanol in the feed and accumulator tanks

(partially filled with mixtures in the desired compositions) were 15 and 80% respectively. During the startup, the feed stream is continuously introduced and controlled at 300 $L.h^{-1}$. To do this experimentally the process is closed, that is, the bottom and top product streams form the feed stream in the main tank (Figure 1). Thus, the startup is initiated with the withdrawal of bottom product (bottom level control). The liquid mixture from the feed descends throughout the column until it reaches the bottom, where the heat exchanger heats and vaporizes this stream. The vapor mixture ascends through the column, heating it tray-by-tray until it reaches the condenser. Here step 1 is finalized. The condensed mixture is received by the accumulator, which is already filled with the material at the desired composition (80% v/v ethanol) for the reflux stream. The reflux is total until a stable situation is achieved (temperature deviation is null over time) which characterizes the end of step 2. At this moment, the top product withdrawal is started and the accumulator level bottom and top temperature control loops are changed to automatic mode, starting step 3.
The steady state is determined when the accumulator temperature and top composition do not vary over time. These variables were used as evaluation parameters. Volumetric composition measurements corroborate this analysis. The composition was measured with a densimeter for alcohol.
The startup procedure with the distributed heating approach was almost the same. The only difference was the use of an electrical resistance with its total power (3.5 kW) at one intermediate plate, enabled from the beginning of the experiment.
For the identification of the most sensitive stage for the application of the distributed control sensitivity analysis methods were used: successive plates, sensitivity symmetry and maximum sensitivity, as detailed in Marangoni (2005). The results showed that in the stripping section plates 2 and 3 were the most sensitive plates. In this study, plate 2 was chosen for application of the distributed control to avoid more than one simultaneous effect. Rectifying trays was not used because at the beginning of startup only the stripping section has enough liquid content in the trays to allow heating through electrical resistance.

3. Results

The temperature profiles of a stripping tray (tray 3) and a rectifying tray (tray 13) are represented in Figure 2. In these experiments a difference was noted in the tray behavior along the sections of the column during the heating. The *stripping* trays show a linear heating region, and which is not observed in the *rectifying* trays. Tray 3 was chosen to illustrate the stripping section as it is the tray which best represents the heat addition effect. This stage is successive to tray 4 which receives the feed stream, as also indicated by the sensitivity analysis.
The results shown in Figure 2 indicate that before the beginning of the reflux, the temperature reached at the top of the column is slightly higher for the distributed configuration than for the conventional one. However, the temperatures after the introduction of reflux were almost the same in both situations. This confirms that the steady state and the reflux flow rate (30L/h) were the same in both cases.
After the linear region (linear phase), shown in Figure 2, an interval can be observed. This is the time required to heat the tray contents. In this step, the heating is slow until the bubble point temperature of the mixture is reached, and the tray temperature increases rapidly. When using the distributed strategy, the time saving is the same for both the third and thirteenth trays (0.32h), when compared with the conventional strategy.

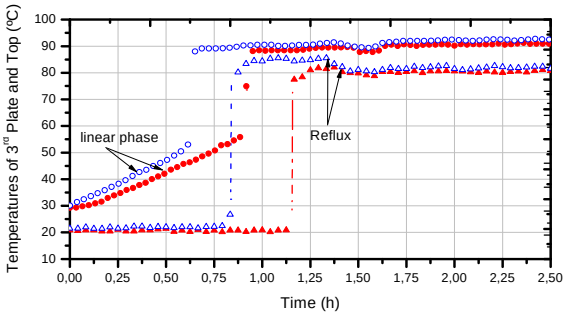


Figure 2 - Temperature profiles of the third tray: conventional (●) and distributed heating applied to tray 2 (○); and of the thirteenth tray, conventional (▲) and distributed heating applied to tray 2 (Δ), during the process startup.

An analysis of the angular coefficients of the linear regression in the feed and tray 3 temperatures for conventional and distributed modes was carried out. The results are summarized in Table 1. In all of the curves the multiple regression coefficient obtained was higher than 0.997. It is observed that, for both temperatures, angular coefficients are higher when the distributed strategy is used. These results, along with those shown in Figure 2, indicate that the distributed heating resulted in a better performance, a faster action and a reduction in the time required to reach the desired temperature at the top the column.

Table 1 - Coefficients of the linear phase of the heating at the temperatures of the feed stream and of tray 3.

Model	*Linear Regressiom*	
	Conventional	*Distributed Heating*
Feed temperature	Y=34.090.X + 31.235	Y=40.517.X + 31.527
Tray 3	Y=31.532.X + 29.761	Y=35.155.X + 30.153

Figure 3 shows the profile of the bottom temperature before and after the control loop is changed to automatic mode with the two studied configurations. The distributed heating again showed a faster action when compared with the conventional process. Using the intermediate heating, the desired temperature value (98°C) was reached 0.75h faster than using the conventional configuration. This is of particular interest as it shows that the action of the resistance increased the heating rate of the bottom temperature.

In order to better evaluate the effect of the distributed heating, the evolution of the derivatives of the bottom temperatures were compared for both experiments. These derivatives were obtained from sigmoidal functions, adjusted to describe the bottom temperature in each of the cases, as shown in Figure 3c. It was verified that the proposed distributed configuration allowed the steady state to be reached faster. A period of 1.4h was required for the bottom temperature to become stable with the conventional configuration and only 0.8h for the distributed configuration. This represents only 57% of the time required to reach the desired bottom temperature.

In addition, in Figure 3b it can be observed that after changing the control loop to automatic mode the bottom valve operated with smaller openings, thus compensating the energy added to plate 2.

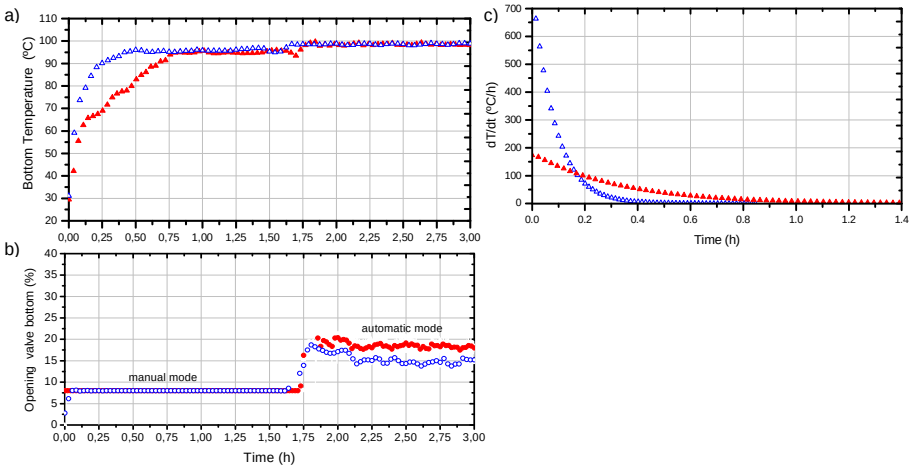


Figure 3 – a) Column bottom temperature, conventional (▲) and distributed heating applied to tray 2 (Δ), b) manipulated variable conventional (●) and distributed heating applied to tray 2 (○) and c) derivatives of the bottom temperatures as a function of time.

The reduction in the time required to reach the steady state when the distributed heating is applied is confirmed by the measurements of the volumetric ethanol fractions in the bottom column stream and their derivatives (dX/dt) to characterize the steady state, as shown in Figures 4a and 4b, respectively.

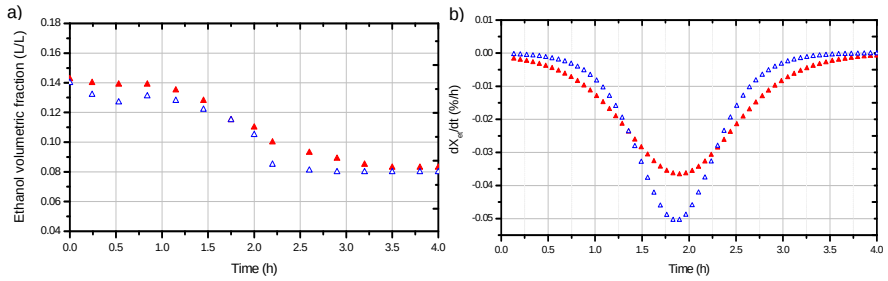


Figure 4 - a) Volumetric ethanol fractions in the bottom stream of the column, b) Derivatives of the volumetric ethanol fraction as a function of time, conventional (▲) and with distributed heating applied to tray 2 (Δ).

It can be observed that there is a change in the bottom ethanol fraction which starts at 0.14% and ends at 0.08%. This is due to the introduction of reflux. Through this observation it is possible to analyze the period out of steady state - as illustrated by the derivative of this variable in relation to time. Is important to note that reflux is introduced in the 13th stage – more distant from the bottom. This means that the interaction between the bottom and top is responsible for the longer time required to reach steady state (or to maintain the column out this condition). As illustrated by other results, Figure 4b shows that the distributed configuration allows that the transition period is shorter than the conventional one (around 0.75h).

In Figure 5 the results obtained in terms of the time required for each column startup step are summarized. The effect of the distributed heating in relation to the conventional heating can be clearly observed, particularly during steps 1 and 3.

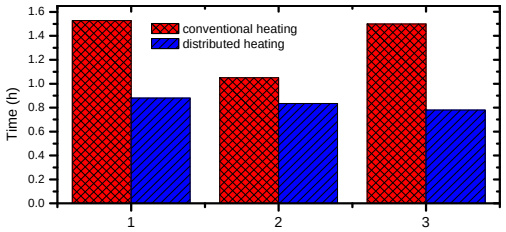


Figure 5 - Startup steps: 1 Heating all stages, 2 Manual Reflux, and 3 Reflux + Bottom product in the automatic mode.

In all three stages the reduction in the time, due to the effect of the addition of the intermediate heating point, was significant when compared to the conventional procedure, being in the order of 43%, 21% and 50%, respectively. Thus, a reduction of approximately 40% in the total startup time was achieved in the distillation column operating with heat distribution, which suggests considerable gains in the process.

4. Conclusions

Comparison between the two operation approaches, showed that the temperatures of the trays of the stripping section increase at a higher rate when the distributed heating with a resistance on plate 2 is used, reaching a permanent regime in a much shorter time than in the conventional process. Thus, the introduction of distributed heating along the column is demonstrated to be a valid option for reducing the startup time. In this study, a reduction of approximately 40% in the startup time was observed in relation to the conventional process, allowing faster dynamics and lower operation costs, with almost the same amount of energy as the conventional configuration. The reduction in the time required for startup, resulting in a considerable decrease in the production of out-of-specification products, compensates for the slight increase in the energy supply.

5. Acknowledgments

The authors are grateful for the financial support from the National Agency of the Petroleum (ANP), and the Studies and Projects Fund (FINEP), by means of the Human Resources Program of ANP for the Petroleum and Gas sector – PRH-34-ANP/MCT.

References

C. Marangoni, 2005, Implementação de uma Estratégia de Controle com Ação Distribuída em uma Coluna de Destilação. Tesis. Programa de Pós-Graduação do Dep. de Eng. Química e Alimentos, Univ. Federal de Santa Catarina, Florianópolis.

F. R. Steinmacher, C. Marangoni, R.F.A Machado, P. H. Araújo, 2004, Startup evaluation and definition of operational conditions of a distillation column. Annals of the Brazilian Congress of Petroleum and Gas.

H. Z. Kister, 1979, When tower startup has problems - Hydrocarbon Processing. Chem. Eng. Science, v. 2, p. 89–94.

J. A. Fabro, L.V.R. Arruda, F. Neves Jr., 2005, Startup of a distillation column using intelligent control techniques, Comp. & Chem. Eng., v.30, p.309–320.

M. Eden, A. Koggersbøl, S. B. Jørgensen, 2000, Dynamics and control during startup of heat integrated distillation column, Comp. & Chem. Eng., v. 24, p. 1091–1097.

M. Barolo, G. B. Guarise, S. Rienzi, A. Trotta, 1993, On-line startup of a distillation column using generic model control. Comp. & Chem. Eng., v. 17, p. 349-354.

N. J. Scenna, S. J. Benz, 2003, Start-up operation of reactive columns with mulitple steady states: ethylene glycol case, Ind. Eng. and Chem. Research, v.42, p.873-882.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Failure Diagnostics Using Data Mining Tools

Carlos A. Vaz Jr,[a] Ofélia de Q. F. Araújo,[a] José Luiz de Medeiros[a]

[a]*Departament of Chemical Engineering, Escola de Química da Universidade Federal do Rio de Janeiro. Sala E-201, Bloco E, Centro de Tecnologia, Ilha do Fundão, CEP: 21949-900 Rio de Janeiro, Brazil*

Abstract

The increase in the use of electronic instrumentation on the chemical industry produces a large operational database. Several methodologies have been applied to the task of review and extract relevant information from these data. A software based on principal component analysis (PCA), hierarchical classifiers and prototypes method was developed. This software detects operational failures using time series generated by the process. The process is an industrial pipeline network for naphtha transport. Two different failures were analyzed: sensors failures and leaks. The software presents a satisfactory failure detection performance. After the leak was detected, the software must proceed to locate and quantify it. The leak parameters – its diameter and location – were estimated correctly for severe failures. The success of detection proceeds is strongly influenced by the severity of the problem. This successful rate reaches more than 95% for severe failures.

Keywords: pipeline, leak detection, sensor failure detection, data mining

1. Introduction

The transport of liquids by pipelines is applied to the displacement of several products, such as water, oil, gas and fuel derivatives. Although highly reliable, pipelines may present severe accidents in some situations, like fatigue and collapse of the tube material. Pipes transporting toxic and/or flammable products are inherently associated with a high potential risk factor (Silva et al., 1996).

Pipelines used for a long time with both high pressure and flow rate may suffer material fatigue and collapse. This situation produces leaks of material and environmental damage, loss of inventory, fires, explosions and the possibility of loss of human lives. Preventing leaks is very important to the safe operation of a pipeline and so are procedures to minimize the extent of its consequences in case of a leak occur. Then, leaks become a constant concern of the chemical industry, regulatory agencies, and society.

In addition to leaks, another kind of pipeline's disruption is the failures in sensors. The development of micro-electronics and its industrial applications create low expensive equipments. Thus, the number of sensors in use in many different industry areas has grown strongly, increasing the possibility of failure in any of these electronic components (Singhal, A., Seborg, D., 2002).

Identifying sensors failures is important for the correct operation of complex pipelines networks. Many methods for detection of anomalies are based on comparison of real time process data and predictions generated by a mathematical model. The fail is detected if significant differences between the process values and model predictions are observed. The models applied to produce process predictions can be

"phenomenological" or "black box". "Black box" models do not require physical description of the process. Examples of "black-box" models are "neural networks", "AR" (autoregressive models) and "ARMA" (autoregressive moving average) (Vandaele, W., 1983).
In this article, tools based on principal component analysis (PCA), hierarchical classification and prototypes were applied to detect, localize, quantify and identify failures.

2. Methodology

This item describes the mathematical methods used in this work. The methods presented were applied to allow detect the presence of sensors failures or leaks. The methodologies also identify the type, origin and severity of the anomaly.

2.1. Hierarchical Classification

Hierarchical classifiers try to separate or classify observations contained in a wide range of data. Groups are formatted to promote the separation. The members of each group show high similarity between them. Similar observations are placed in the same group, and "non-similar" objects are placed in different groups (Hart P., Duda, R. O., 2000). Hierarchical classifier was applied to identify observations "with" and "without" the presence of failure. The similarity between two observations can be understood as inversely proportional to spatial distance between these points. Or, the spatial distance between two points are proportional to the "dissimilarity" between the observations.
In a two dimensional space (R^2), the distance between point "i" and point "j" in the plane is traditionally defined using Euclidean metric. This metric can be extended to a generic spatial dimension "n" (R^n) using Equation 1.

$$d_{ij} = \left[\sum_{k=1}^{n} \left(x_{i,k} - x_{jk} \right)^2 \right]^{1/2} \quad (1)$$

The Euclidean distance is the most applied distance metric, but it is not the only definition for "distance" between two points. The "City Block" metric (Equation 2) and "Minkowski" (Equation 3), described by Duda and Hart (2000) can be applied too. In some situations, City Block and Minkowski show greater effectiveness in the characterization of similarity between two observations. The City Block and Euclidean are special cases of Minkowski equation, using "r" equal to 1 and 2 in Equation 3.

$$d_{ij} = \sum_{k=1}^{n} \left| x_{ik} - x_{jk} \right| \quad (2)$$

$$d_{ij} = \left(\sum_{k=1}^{n} \left| x_{ik} - x_{jk} \right|^r \right)^{1/r} \quad (3)$$

Once defined the metric classification, the classification procedure is simple. A matrix of distances between all the data observations is created. Then, the two points with smaller distance between them are selected. These two observations are merged, creating a new point. A new matrix is prepared, and the procedure is repeated. The procedure continues until all data is allocated in two clusters. The observations on the same group must present high similarity between them, and strong dissimilarity when compared to data of another cluster. Finally, the identification of the two groups

("failure" and "without failure") is promoted classifying an observation previously known. Besides allowing the detection of failures, the hierarchical classifier also quantifies the severity of the fault.

2.2. Method of Prototypes

One approach to identification and location of the faults was based on method of prototypes. To implement this methodology, representative cases of major anomalies are selected, and these observations are called "prototypes". The prototype must be appropriated to identify similar situations.

After an operational anomaly has been detected, the experimental data of this failure is then compared with the prototypes, looking for an elevated similarity. A PCA model is adjusted for each prototype and for the experimental data. The models are compared using a similarity methodology called “Distance Similarity Factor” (Sdist). The prototype identification allows estimate the “type” (leak or sensors failures) and the “location” of the fault. The PCA and Sdist methodologies are described below.

2.3. PCA Methodologies

The principal component analysis (PCA) methodology can be described as the calculation of eigenvectors and eigenvalues of a covariance matrix obtained from the normalized data. The eigenvectors, especially those associated with larger eigenvalues; provide important information about the pattern of data distribution (Hart P., Duda, R. O., 2000).

The eigenvalues define the amount of information described in each eigenvectors. This article used eigenvectors enough to describe 99% of the pattern of data distribution. The eigenvectors are called "components", while the eigenvectors associated with the high eigenvalues are called "principal components".

The matrix of components (MC) is composed by the principal components, and the PCA model is represent by matrix "C" (equation 4).

$$C=MC*MCT \qquad (4)$$

The matrix "C" is multiplied by experimental data ($\hat{x}\exp_k$) resulting in the modeled data ($\hat{x}\,\mathrm{mod}_k$) (equation 5).

$$\hat{x}\mathrm{mod}_k = C\ \hat{x}\exp_k^T \qquad (5)$$

2.4. Distance Similarity Factor (Sdist)

The Sdist factor is the probability that the center of the data set A ($\overline{x}_A$) is at distance ϕ (or less) from the center of the data B (Singhal, A., Seborg, D., 2002). The Sdist is calculated using equation 6, where ϕ is the Mahalanobis distance.

$$Sdist = 2\left[1-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\phi} e^{-\frac{z^2}{2}}dz\right] \quad , \quad \phi=\sqrt{(\overline{x}_A-\overline{x}_B)\sum\nolimits_s^{*-1}(\overline{x}_A-\overline{x}_B)^T} \qquad (6)$$

2.5. Pipeline Network

To test the methodologies proposed here, a pipeline network simulator developed by Vaz Junior (2006) was employed. The computer simulation allows the realization of a large number of experiments and tests involving the pipeline, with different conditions of sensor faults and leaks.

The pipeline network employed here transport naphtha over a length of 30 km connecting the vertex "1" (origin) to the vertex "8" (destination). A network diagram is shown in figure 1.

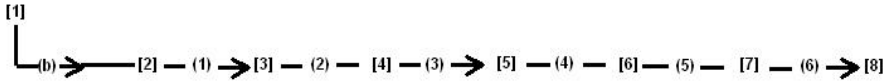


Figure 1: Pipeline Network, where: [x] vertices, (b) bombs, (x) tubes.

The pressurization and naphtha transportation in the pipeline are defined by specification of pressure (P) and external mass flow (W) directly in the vertices "1" and "8". The other parameters are calculated using fluid mechanics phenomenological equations.

To simulate the signal generated by real sensors, a time series containing random values was incorporated to the data generated by the simulator. The time series generated by the simulator present a 1-minute interval between each observation. Table 1 present all the sensors installed in the pipeline network.

Table 1: Sensors installed in the pipeline network

Sensor	Localization	Type	Sensor	Localization	Type
1	[1]	Flow	9	[7]	Press
2	[8]	Flow	10	[8]	Press
3	[1]	Press	11	(1)	Flow
4	[2]	Press	12	(2)	Flow
5	[3]	Press	13	(3)	Flow
6	[4]	Press	14	(4)	Flow
7	[5]	Press	15	(5)	Flow
8	[6]	Press	16	(6)	Flow

3. Results

This session describe the results obtained using the methodologies proposed here. Different approaches for detection, identification, localization and quantification of failures were tested. Several scenarios were studied, changing the location, severity and type of the tested failures. To simulate sensor failures, deviations in the form of step with magnitude 2%, 5%, 10% and 20%, were add to the signal from each sensor. Leaks caused by holes with a diameter of 1, 2, 4, 5, 8, 12, and 14 cm were simulated.

3.1. Sensor Failure Detection using hierarchical classification

A time series contain 450 observations of normal network operation and 50 observations of normal operation was employed (figure 2). Using Minkowski metric, more than 95% of the scenarios with sensors failures of severities 10% or 20% were correctly detected. The best performance was obtained using Minkowski metric (figure 3a). The degree of success of this methodology, however, is reduced in the presence of faults of low severity. The physical position of the sensor in the pipeline network also influences the efficiency of the detection tool. The fail detection in the sensors "3" and "10" presented lower efficiency.

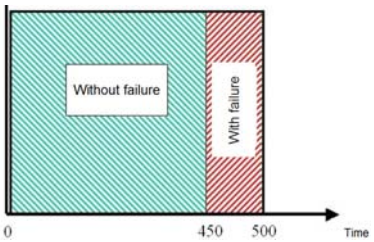
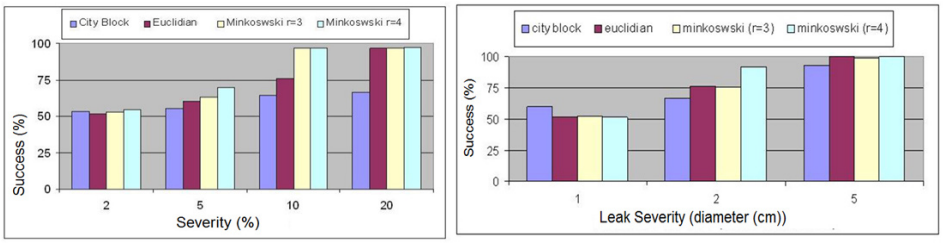


Figure 2: Phases with and without fault



(a) (b)

Figure 3: Degree of success to detect sensor failures (a) and leaks (b) using hierarchical classification

3.2. Leak Detection using hierarchical classification

Like sensor failure detection, hierarchical classification was employed to leak detection. Again, a time series consisting of 500 observations, of which 450 represent normal operating situation, and 50 represented a leak occurrence, was used. The metrics "Euclidian", "City Block", and Minkoswki" were tested.

Figure 3b presents the results for leaks with diameters of 1, 2 and 3 cm. The approach using hierarchical classification was not able to detect leaks of very small volume. For more severe leaks, the performance was satisfactory, especially using Minkoswki (r=4). The capability of the leak detector is also a function of the leak location.

3.3. Differentiating leaks from sensor failures occurrences

After an abnormal operation condition is detected, it is necessary to establish its type (sensor failure or leak). The method applied to detect failures is not able to differentiate between these two types of anomalies. A hybrid approach, using "hierarchical classification" and "method of prototypes" was applied. Twenty-two prototypes representing leaks and failures in sensors were created. The prototypes together with the data containing the detected anomaly are subjected to classification procedure. The similarity between the experimental data and the prototypes is able to identify the type of anomaly. This technique was successful in differentiating leaks from sensor failures, even in low severity situations (Table 2).

3.4. Localizing failures using prototypes

After determining the type of anomaly, the prototypes methodology is applied again in order to identify the location of the failure. The same 22 prototypes, 16 sensor failures and 6 cases of the "leak" are used. The most similar prototype is search using Sdist metric. This approach was effective to localize sensor faults and leaks in the pipeline network.

Table 2: Differentiating leaks from sensor failures occurrences

Leak (diameter - cm)	Success (%)	Sensor Failure (%)	Success (%)
1	33,33	1	100,00
2	83,33	2	100,00
4	100,00		
5	83,33		
8	83,33		
12	100,00		
14	100,00		

3.5. Quantification of sensor faults and leaks

Hierarchical classifications were applied to quantify sensor failures and leaks in pipeline network. The method can estimate the severity of operational deviation using the distance between the centers of the two clusters formed ("with" and "without" failure). Figure 4 shows the relation between the failure magnitude and distance between the two clusters. The Minkowski metric (r=4) has been used.

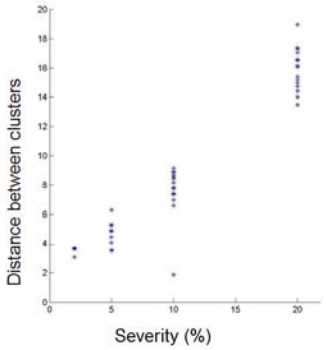


Figure 4: Distance between clusters vs. severity of the sensor failure

4. Conclusion

In this article, hierarchical classifiers were successfully used to detect abnormal operation in the pipeline network. The success of the methodology was strongly influenced by the severity of the failure and its localization. A hybrid approach, joining prototypes and hierarchical classifiers, was effective to identify the failure type (sensor failure or leak). Similar approach was applied to establish the location of the fault. Finally, the distance between the two clusters formed allows a quantification of the failure severity.

References

P. Hart, R. O. Duda, Pattern Classification, 2a ed. John Wiley Pro, 2000

A. Singhal, D. Seborg, 2002, Pattern Matching in Multivariate Time Series Databases Using a Moving-Window Approach, Ind. Eng. Chem. Res. v.41, p.3822-3838.

R. A. Silva, C. M. Buiatti, S. L. Cruz, J. A. F. R. Pereira, 1996, Pressure wave behaviour and leak detection in pipelines. Computers Chem. Engineering, v.20, p.S491-S496

W. Vandaele, 1983. Applied time series and Box-Jenkins Models. Academic Press

C. A. Vaz Junior, 2006, Detecção, Localização e Quantificação de Vazamentos: Uma Abordagem em Séries Temporais. Master in Science Thesis Escola de Química da UFRJ

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimization of wastewater filtration process in submerged membrane bioreactors: applicability of a dynamic model to scale up.

Alain Zarragoitia,[a, b] Silvie Schetrite,[a] Ulises J. Jáuregui-Haza,[c] Olivier Lorain,[d] Claire Albasi[a]

[a]*Laboratoire de Génie Chimique, UMR - CNRS 5503, 5 Rue Paulin Talabot, 31106 Toulouse cedex, France*
[b] *Departamento de Desarrollo Tecnológico, Centro de Química Farmacéutica (CQF), Calle 200 y 21, Apdo. 16042, Atabey, Playa. La Habana, Cuba*
[c]*Instituto Superior de Tecnologías y Ciencias Aplicadas (InSTEC), Ave. Salvador Allende y Luaces, Plaza, Ciudad de la Habana, Cuba*
[d]*POLYMEM, Toulouse cedex, France*

Abstract

The application of a hybrid mathematical model, which takes into account the effect of main variables in Submerged Membrane Bioreactor (SMBR) systems, was established in order to simulate and optimize the filtration process into SMBR, on the base of experimental data performed on a bench and pilot scale bioreactors. Numerical simulations of the wastewater treatment process were performed in order to find the optimal filtration conditions for both different scale bioreactors. Various operating conditions (idle-filtration time, aeration intensity, solids retention time, hydraulics retention time, and total suspended solids concentration) were tested. Different optimization criteria were considered to minimize the transmembrane pressure and energy consumption, and to maximize filtrate flow.

Keywords: Biological wastewater treatment; Submerged Membrane Bioreactor (SMBR); modeling; optimization.

1. Introduction

The wastewater treatment using a submerged membrane bioreactor (SMBR) allows obtaining a quite constant filtrate quality after biological degradation process and physical separation. The membrane bioreactor technology has been improved, and research on SMBR technology has increased significantly (Yang, 2006). MBR treatment of municipal wastewater yields high-quality water with reported removal percentages of 95%, 98%, and 99% (or greater) for chemical oxygen demand (COD), biochemical oxygen demand (BOD), and suspended solids (SS), respectively (Manem and Sanderson, 1996). Nevertheless, the filtration process is disturbed by the influence of complex factors that cause the membrane fouling.
On the other hand, mathematical modeling and simulation are powerful tools with which the specialists can predict the performances of potential systems under different operating conditions. For this reason, it is interesting to use models that can accurately describe the SMBR process for the design, and optimization of these systems.
The objective of this work is to show the practical application of a hybrid mathematical model, developed in a previous work (Zarragoitia, 2008), for optimizing the operation

of two SMBRs using different optimization criteria to minimize the transmembrane pressure (TMP) and energy consumption, and to maximize filtrate flow. The model was established considering Soluble Microbial Products (SMP) formation-degradation kinetic based on modified published models (Cho, 2003; Shane, 2007; Lu, 2001). A modification of Li and Wang's model (Li and Wang, 2006) allows to calculate the increase of the TMP evaluating the influence on fouling control of an intermittent aeration of coarse bubbles synchronized with the filtration cycles, and to analyze the effects of shear intensity on sludge cake removal. In order to describe the biological system behaviour a modified ASM1 model was developed to estimate the optimal operating conditions related to a specific bioreactor condition.

2. Experimental part

2.1. Experimental set-ups

Two different SMBRs were studied. In the first (bench scale), a U-shaped, hollow-fiber membrane (polysulfone, pore size: 0.1 μm, internal/external diameter: 0.4/0.7 mm) module with area of 0.3 m^2 (Polymem, Toulouse, France), was immersed in a bioreactor of 10.5 L of working volume. The municipal wastewater was continuously treated. The influent flow rate was controlled by the liquid level in the reactor. Filtration was operated in an intermittent sequence of filtration-relaxation. The TMP was continuously monitored (Sensor Keller). Filtrate flow was measured with an electromagnetic flow meter (Rosemount); a temperature sensor (PT 100; -50 to 250 °C) and a Mettler Toledo pH meter were used. The pH was maintained between 6.5 and 7.5. PC-based real time data acquisition hardware (IOTECK) and the software DASYLAB have been used for acquiring all data. The second bioreactor has four hollow-fibre membrane modules with an area of 2.5 m^2 and a bioreactor tank of 1 m^3 (Polymem, Labège, France). The membrane material is the same as the previous one. This bioreactor is operated in a largest plant, collecting wastewater from domestic and industrial areas. The bench scale bioreactor was operated with two types of aeration flows. An intermittent coarse bubbles flow injected closed to the fibres, to avoid the membrane fouling by reduction of cake formation, and a constant fine bubbles flow injected through a perforated membrane at the bottom of the reactor, providing mixing and biomass oxygenation. The membrane module was isolated to the contact with fine bubbles; thus, membrane fibers movement is only produced by the flow of coarse bubbles. The operation was stopped at 12 days of filtration or when the TMP reached 60 kPa under atmospheric pressure, and then a chemical cleaning was done.

2.2. Analytical methods

The Total Suspended Solids (TSS) concentration was determined using a PRECISA HA60 moisture analyzer. The activated sludge flocs size distribution was estimated with a laser granulometer (Mastersizer 2000, Malvern Instruments). The Extracellular Polymeric Substances (EPS) were quantified in influent and permeate samples. Proteins and polysaccharides were measured by spectrophotometric methods. The Lowry method was used for the quantification of proteins and humics. For quantitative analysis of carbohydrates, the anthrone method was used with glucose as standard. Respirometric measurements were used for the determination of readily and slowly biodegradable substrates considering Y_H = 0.67 gCOD / gCOD.

2.3. Simulation conditions and optimization procedure

A numerical simulation of the membrane fouling was performed using a program in Berkeley Madonna programming language. Membrane fouling development was

simulated by numerical iterations, which produces the overall TMP increase and the evolution in the distribution of the flux and the sludge cake layer across the membrane surface sections. The intermittent process of filtration and coarse bubbles injection were simulated by means of periodic functions. When the simulation results were compared with experimental data a good agreement was observed, thus, a D-Optimal Experimental Design (DOED) matrix was used for the optimization of the operating conditions for both SMBRs (Table 1). The values of the response variables in each experiment were estimated "in-silico" using our model. Finally, the predicted optimal operating conditions were validated by experimental way for bench scale SMBR. The validation of the optimum for the pilot scale SMBR is under study.

Table 1- Summary of the DOED.

	Bench scale SMBR, Experiments= 26, qa = 9 L/(m^2 s), J_T=0.24 m^3/(m^2 day)				
	tf (min)	tid (min)	Int_{CB} (min)	t_{CB} (min)	SRT (days)
Low actual	5	2	4	1	20
High actual	15	5	10	3	70
	Pilot scale SMBR, Experiments= 30, qa = 1 L/(m^2 s), J_T = 0.24 m^3/(m^2 day)				
	tf (min)	tid (min)	Int_{CB} (min)	t_{CB} (min)	tbw (min)
Low actual	5	0.5	0.2	0.5	
High actual	105	1.5	0.75	3	0.5

qa= aeration intensity, tf= filtration time, tid= idle time, Int_{CB}= time interval between two coarse bubbles injection, t_{CB}= coarse bubbles injection time, SRT=solids retention time, J_T= overall flux tbw=backwashing time

3. Results and discussion

3.1. Optimization of the operating conditions for the bench scale SMBR.

During the optimization, different criteria were considered, to minimize the transmembrane pressure and energy consumption, and to maximize filtrate flow. The equation 1 shows the objective function used to calculate and to minimize the energy consumption. On the other hand, table 2 shows a summary of the measured influent characteristics and initial operating conditions used during simulation. The table 3 shows the optimization results estimated by simulation as well as the predicted values of the optimized response variables.

$$EC = \left[\frac{0.666(Qfil * TMP * tf)}{FV * (tf + tid)} + \frac{0.666(Qair * \Delta Pair * t_{CB})}{FV * (t_{CB} + Int_{CB})} \right] time \quad \langle \mathrm{kW/m^3}h \rangle \quad \text{(eq. 1)}$$

where: EC= Energy consumption (kW/m3h), FV= Filtrate volume (l), Qair= aeration flow (l/h), Qfilt= Filtration Flow (l/h), ΔPair= aeration drop pressure (bar)

During the optimization, only five factors were considered in the DOE. Figures 1-2 show different examples of the optimization results, but considering only the values of tid and tf in the axis. The other factors values could be calculated in the same way, once a specific optimization criterion is chosen. Table 4 points out the existence of differences in the tf and SRT values for the case A and B, but it also suggests that different filtration-idle time cycles may lead in both cases to work in optimal conditions, this also change the synchronization between the filtration and the intermittent aeration. This result is reasonable and may find explanation in the fact that the cases A and B have different influent characteristics and biomass concentration

(Table 2). These differences modify the behaviour of the system due to the presence of different F/M ratios, different sludge viscosities, and biological variables.

Table 2- Summary of the measured influent and initial conditions.

	Measured influent and initial conditions, values used during simulation						
Case	X_{TSS0}	X_{TSSR0}	X_{S0}	X_{I0}	S_{S0}	S_{I0}	S_{SMP0}
A	25	6000	100	25	155	30	
B	30	8300	90	20	190	45	45

(Superscript 0 and R0 expresses "influent" and "initial value in the bioreactor" respectively), X_{TSS}= Total suspended solids in the mixed liquor (mg/l), X_S= slowly biodegradable products (mgCOD/l), X_I= inert organic materials (mgCOD/l), S_S= soluble biodegradable substrate (mgCOD/l), S_I= soluble inert matters (mgCOD/l), S_{SMP}=soluble microbial products (mgCOD/l)

Table 3- Optimization results calculated by simulation and the DOED for bench scale SBMR.

	Optimized values of the operating variables					Predicted values of the response variables		
Case	tf	tid	Int_{CB}	t_{CB}	SRT	FV (12d)	TMP (12 d)	EC (12 d)
A	10 min	2 min	8 min	2 min	40 d	627 L	30 kPa	0.294 kW/m^{3}h
B	15 min	2 min	8 min	2 min	30 d	762 L	37 kPa	0.262 kW/m^{3}h

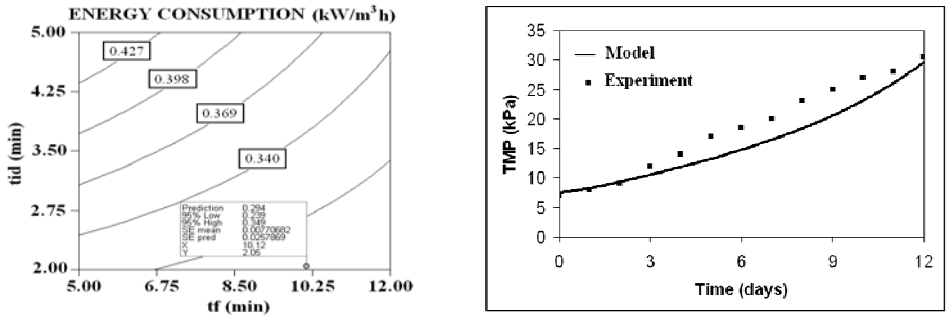


Figure 1- Optimization of operating conditions for the bench scale SMBR. Case A: Minimizing energy consumption [Left: Contour graph obtained during the optimization of the operating conditions (minimizing the energy consumption). Right: Experimental validation of the optimized values of the operating conditions].

The same observations can be made for the composition of mixed liquors, which are not the same for the cases A and B, and modifying the value of the specific filtration resistance of the sludge cake layer and, consequently, the fouling rate. The EC values were found between 0.15-0.96 kW/m^{3}h and the FV values between 432-800 l, therefore, the estimated optimal conditions facilitated, in both cases, to operate the bioreactor for an adequate volume of produced water with reasonable energy consumption (Table 3).

3.2. Optimization of the operating conditions for the pilot scale SMBR.

In the case of the pilot scale SMBR, the optimization study was carried out with the objective of modifying the usual values of the operating variables, in order to improve the operation efficiency of the SMBR. The results are shown in table 4.

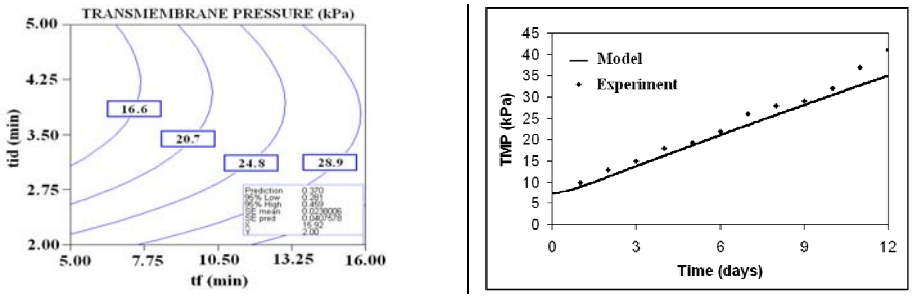


Figure 2- Optimization of operating conditions for the bench scale SMBR. Case B: Minimizing TMP [Left: Contour graph obtained during the optimization of the operating conditions (minimizing TMP). Right: Experimental validation of the optimized values of the operating conditions].

Table 4- Current and optimized values for the pilot scale SMBR (the values in brackets were used during the simulation with the optimized values of the operational variables).

Operating variables	Current value	Optimized value
tf (min)	5	7.61 (7.5)*
tid (min)	0.5	0.5
Int_{CB} (min)	0.25	0.5
t_{CB} (min)	0.25	1.76 (1.75)
tbw (min)	0.5	0.5

The analysis of the simulation-optimization results (Table 5) suggests that, for a better utilization of the SMBR, it would be convenient to increase the filtration time almost of 30% of its current value. In relation to the injection cycles of the coarse bubbles, the idle time between every injection could be increased in 80%, but at the same time, it is necessary to double the length of the aeration time. This operating way is more efficient because it decreases the energy consumption due to the aeration flow used to remove the sludge deposited on the membrane. Although, a higher idle time allows some temporary deposition of the sludge on the membrane, the simulation estimates that the increase of the injection time get possible to later efficiently the deposition of the sludge and the formation of the filtration cake on the membrane. On the other hand, a higher filtration time facilitates the production if a higher filtrate volume (increase of 11%) during the considered time, which decreases the estimated value of energy consumption by m^3 of treated water (decrease of 56%).

In relation to the value of the backwashing time, the model considers that the current value is correct (0.5 min). To evaluate the validity of the simulations and the quality of the model prediction, all results have still to be validated. Some simplifications have been made during the modeling of the system; therefore, it is necessary to characterize the influence of the assumptions and the limitations of the model on the optimization results. The optimization procedure presupposes the existence of a model calibrated to the real system, an accurate process of prediction by simulation is only possible after tuning the parameters of the model to a given plant. Therefore, the more sensible parameters must be recalibrated if the influent conditions and the biomass characteristics change considerably.

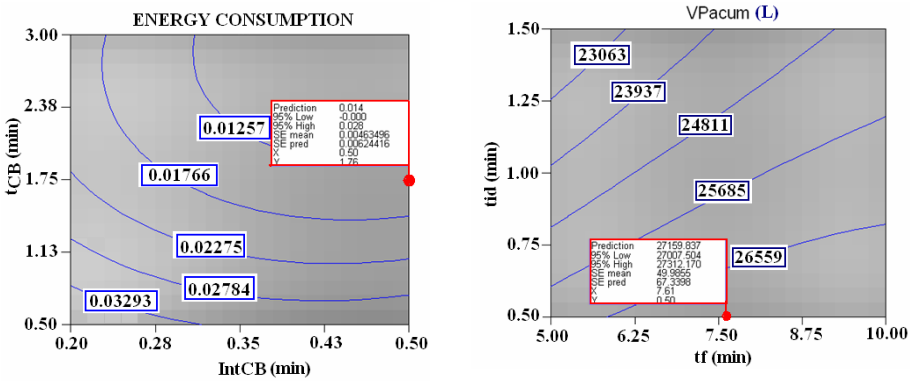


Figure 3- Examples of the contour graphs obtained during the optimization of the operating conditions for the pilot scale SMBR.

Table 5- Influence of the optimized operating variables on the response variables for the pilot scale SMBR.

	Reponse variables (at 15 days)		
	TMP (kPa)	VPacum (L)	EC (kW/hm^3)
Results with the current values of the operating variables	43.6	24000	0.055
Results with the optimized values of the operating variables	25	27160	0.024
TMP reduction	42.6 %	-	-
Volume increase	-	11.6 %	-
Energy consumption reduction	-	-	56.4 %

4. Conclusions

An optimization study to improve the filtration process on two different SMBRs at different scales has been successfully developed. A DOED was carried out "in silico" in order to optimize some operating variables for two different optimization criteria: minimizing energy consumption and minimizing TMP. Subsequently, the experimental verification process of the optimized operating conditions for the bench scale is in good agreement with the model predictions. Future tasks will focus on the system optimization, but, also including the aeration intensity and the filtration flux as factors to consider in the design of experiments. However, for better results a special attention is needed during parameters model calibration.

References

W. Yang, N. Cicek, J. Ilg, 2006, J. Membr. Sci., 270, 201–211.

J. Manem, R. Sanderson (Eds.), 1996, McGraw-Hill, New York.

A. Zarragoitia, S. Schetrite, M. Alliet, U. Jauregui-Haza, C. Albasi, 2008, J. Membr. Sci., 325, 2, 612–624.

J. Cho, K.-H. Ahn, Y. Seo, and Y. Lee. 2003,Wat. Sci. Tech., 47, 12, 177–181.

R. Shane, R. P. Merlo, W. Hermanowiczc, D. Jenkins, 2007, Water Res., 41, 947–958.

S.G. Lu, T. Imai, M. Ukita, M. Sekine, T. Higuchi, M. Fukagawa, 2001, Water Res., 35, 2038–2048.

X.-Y. Li, X.-M. Wang, 2006, J. Membr. Sci., 278, 151–161.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Control Structure Design for an Ethanol Production Plant

Gabriel V. N. de Andrade[a], Enrique L. Lima[b]

[a]*CHEMETCH - A Siemens Company, Rua da Quitanda, n. 50, 21º andar, Rio de Janeiro – RJ, CEP 20011-030, Brazil*
[b]*COPPE – Chemical Engineering Program – Federal University of Rio de Janeiro (UFRJ), Cidade Universitária, CP 68502, Rio de Janeiro – RJ, CEP 21945-970, Brazil*

Abstract

This paper studies two methodologies for control structure design. Both methodologies were applied to an ethanol production plant with an energy integration technology known as split-feed. A commercial process simulator was used. The proposed structures were tested to verify its performance. Finally, it was chosen the most efficient methodology for industrial applications.

Keywords: Plantwide control, ethanol, split-feed.

1. Introduction

Nowadays, there is great incentive for searching alternative energy sources to replace petroleum. Because of its renewable characteristics, ethanol has become an interesting option for automotive fuel. Also, industrial plants are now spending a great effort on reducing their energy costs. At ethanol production plants, a technology called split-feed has been used for this purpose. However, this technology still presents operating problems, demanding the development of control systems capable of dealing with such problems. A good alternative for this is the use of Plantwide control methodologies that, with detailed procedures, provide the necessary tools to build an efficient control structure.

The purpose of this paper is to apply plantwide control methodologies on an ethanol production plant and evaluate their practical use. Section 2 presents a short review of heat integration with split-feed and plantwide control. The ethanol production plant studied in this paper is presented in Section 3. In Sections 4 and 5, plantwide control methodologies are applied to the ethanol plant. A comparative analysis is made in Section 6 and the paper is concluded in Section 7.

2. The General Problem

2.1. Heat Integration with split-feed

The split-feed technology is used to thermally integrate distillation plants, where the feed is divided into two columns that operate at different pressures, as shown in Figure 1. The high pressure column condenser is responsible to generate the heat duty necessary to operate the low pressure column reboiler, reducing the consumption of heating and cooling fluids of the process.

Studies carried out by Lenhoff and Morari (1982), Chiang and Luyben (1988), Pohlmeier and Rix (1996) and Han and Park (1996) showed that several control

strategies have been tested in distillation systems with split-feed heat integration. Nevertheless, most of these control strategies were design for regular performance, dealing only with small process disturbances. The best results were achieved when a multivariable control scheme was used. Therefore, there is a need for development of more elaborated control structures that can handle the main problems of these complex distillation systems.

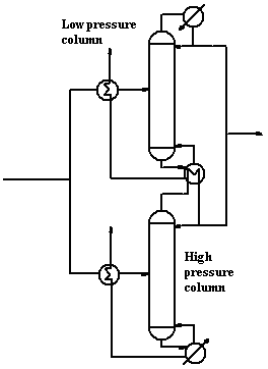


Figure 1: Split-feed arrangement

2.2. Plantwide Control

In the design of control systems the engineer must choose the best variables for measurement, manipulation and control. Also, it is very important to define how these variables should be interconnected, to guarantee the controllability of the plant as well as the good performance of the control system. Despite their great importance for design, these definitions and choices, which are part of the structural decisions (Skogestad, 2000), are usually made based on the engineer experience on the process.

Plantwide control is the design of the control system of an overall plant, where the engineer defines the control philosophy with emphasis on the structural decisions. Due to the complexity of the plantwide control problem, it is very difficult to use a mathematical approach, especially with respect to the dynamic and steady state modeling of the plant.

McAvoy and Ye (1994), Price et al. (1994), Stephanopoulos and Ng (2000) and Skogestad (2004) presented several methodologies for control structure design using the plantwide control concept. However, some of these methodologies use optimization tools that demand more complex computational approaches. Also, due to the lack of use of these methodologies in the design of industrial plants (maybe because the engineers still don't trust them), it is important to change this scenario by implementing them in different kinds of plants.

3. The ethanol production plant

In this paper, control structures will be proposed to an existing ethanol production plant (distillation and stripping sections), which is described below.

The unit has two sections, each one operating at a different pressure. The distillation section has a split-feed arrangement. With this heat integration, the plant has significant reduction on the utilities consumption (steam and cooling water), but the strong interaction between the two sections causes several problems to the control system.

Due to the great amount of streams and equipment in the plant, there are around 80 possible measurements and 50 manipulated variables, which give lots of possibilities of pairing. Therefore, in this paper, only the low pressure section will be considered, changing the scenario to 30 possible measurements and 15 manipulated variables. Figure 2 presents the low pressure section of the unit:

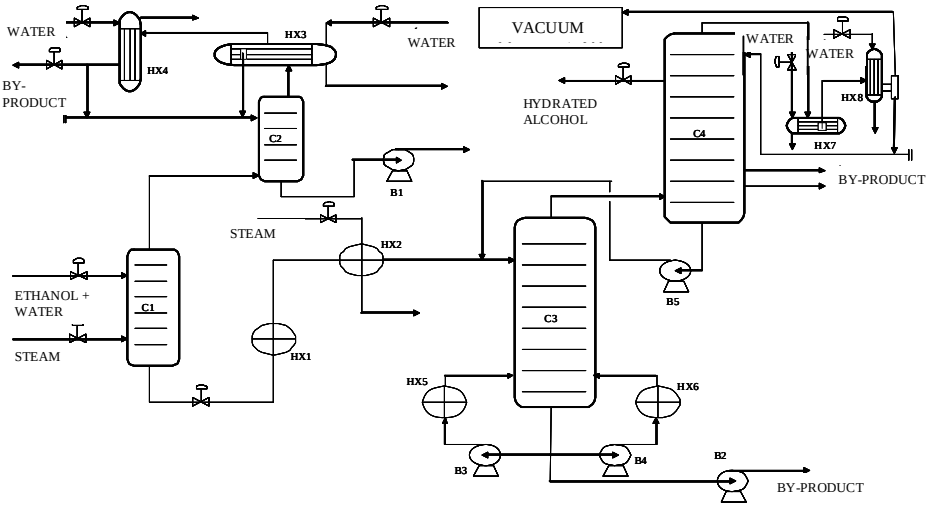


Figure 2: Low pressure section.

3.1. Simulation

The plant was simulated in a commercial process simulator software, ChemCad 5.6.4 (ChemStations). For the control loops, the PID module of ChemCad (traditional feedback controller) was used.

Since this section of the plant operates at a very low pressure, there is a specific vacuum system to control it. This system could not be simulated in ChemCad, so the pressure was kept constant in the simulation, even during the dynamic tests.

Looking for methodologies that can be easily understood and applied in the industry, the procedures proposed by McAvoy and Ye (1994) and Price et al. (1994) were chosen to be used in this work. This choice was based firstly on the fact that none of these methodologies necessarily uses optimization tools and/or multivariable control. The use of these tools would demand a lot of compilation and execution hours, as well as a computational structure capable of dealing with complex problems. In addition, both procedures are partially based on the engineers` knowledge of the process and also on analysis tools that can be easily applied.

4. Application of the methodology proposed by McAvoy and Ye (1994)

McAvoy and Ye (1994) proposed a systematic methodology for control structure design based on a Plantwide control approach, where single-input-single-output (SISO) control loops are used. The methodology is divided into four stages, as described below:

At the first stage, the fast response control loops are defined. Usually, these loops are simple flows and utility temperatures. The second stage has four steps: closing the level loops; interaction, stability and saturation analysis; steady state disturbance analysis and; tuning and testing via dynamic simulation. In the third stage, the quality control loops are defined, through a global material balance. The last stage consists of adding upper control layers, using real-time optimization tools, predictive control and others.

Besides the last stage, all the others were applied to the ethanol plant using the methodology of McAvoy and Ye (1994). The loops listed in Figure 4 were tuned and the control system was tested dynamically in order to verify its performance. The tests consisted of inserting disturbances (steps) on the unit feed flow and thermal conditions. The control system showed good results, rejecting satisfactorily the external disturbances, as can be seen in Figure 3:

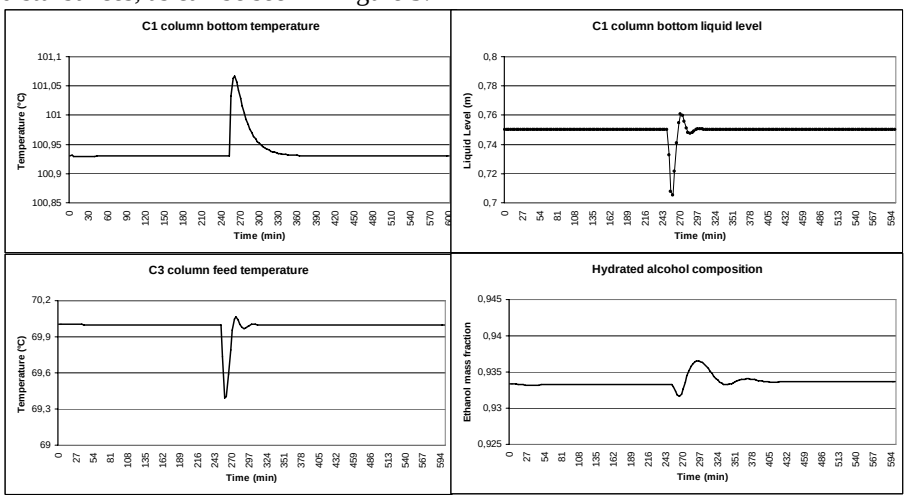


Figure 3: Tests with the control system proposed using the methodology of McAvoy e Ye (1994)

5. Application of the methodology proposed by Price et al. (1994)

Price et al. (1994) proposed a methodology for control structure design based mainly on the unit inventory management and the production rate.

The procedure starts by identifying the primary path of the process and choosing a stream to control the production rate. Then, a set of inventory control loops is added, making sure that all changes in the production rate will be propagated throughout the process.

According to the authors, the production rate manipulated variable must be chosen to obtain a self-consistent control chain, which is defined as a control chain capable of transmitting a production rate change throughout the plant, making adjustments in all inlet and outlet streams. When the control chain is not self-consistent, it is necessary to add extra control loops to help propagating production rate changes.

All the steps of the methodology of Price et al. (1994) were applied, resulting in three different control systems for the plant. As recommended by the authors, the control system with an internal flow controlling the production rate was chosen. The control loops, presented in Figure 4, were then tuned and tested via dynamic simulation, with the same disturbances used at the control structure proposed with the methodology of McAvoy and Ye (1994). Some of the obtained results can be seen in Figure 5.

	McAvoy and Ye (1994) Methodology	Price et al. (1994) Methodology
Manipulated variables	**Controlled variables**	
C1 inlet stream flow valve	C1 feed flow	C1 bottom liquid level
C4 side draw flow valve	C4 side draw flow	Production rate/C4 side draw flow ratio
C2 distillate flow valve	C2 distillate flow	C2 distillate flow
HX3 cooling water flow	HX3 cooling water temperature	-
HX7 cooling water flow	HX7 cooling water temperature	C4 top temperature
HX8 cooling water flow	HX8 cooling water temperature	-
C3 bottom stream flow	C3 bottom liquid level	C3 bottom liquid level
C4 bottom stream flow	C4 bottom liquid level	C4 bottom liquid level
C2 bottom stream flow	C2 bottom liquid level	C2 bottom liquid level
C1 bottom stream flow	C1 bottom liquid level	Production rate
Reboiler flow	C3 bottom temperature	C3 bottom temperature
Direct steam flow	C1 bottom temperature	C1 bottom temperature
HX2 steam flow	C3 inlet stream temperature	C3 inlet stream temperature
HX4 cooling water flow	C2 distillate temperature	C2 distillate temperature
HX7 cooling water temperature (setpoint)	C4 top temperature	-
C4 side draw flow (setpoint)	C4 bottom temperature	-

Figure 4: Control loops defined for each structure.

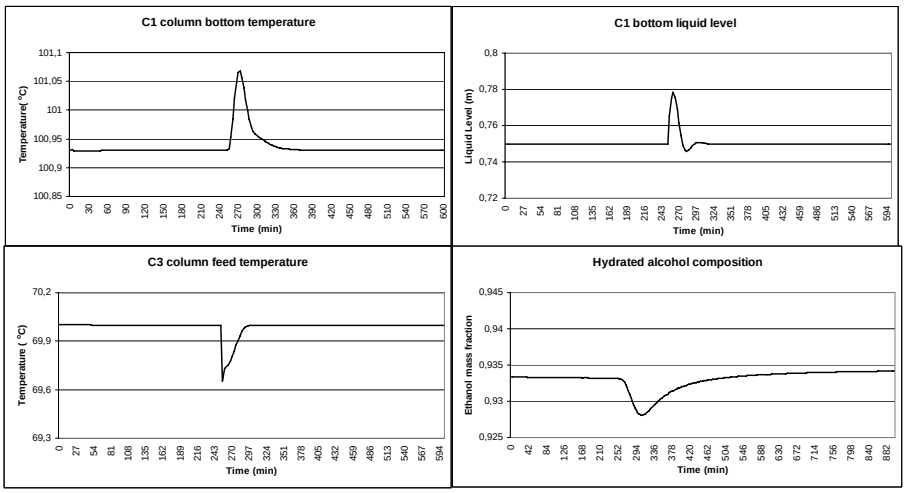


Figure 5: Tests with the control system proposed using the methodology of Price et al. (1994).

6. Comparative Analysis

The methodology proposed by McAvoy and Ye (1994) shows that it is possible to design a simple control structure in a fast and practical way, by using some analysis tools combined with the experience of the engineer on the process.

The methodology proposed by Price et al. (1994) presents a procedure that is easy to follow and understand. The methodology is based on the production rate and inventory controls, where the last ones are responsible for transmitting changes in the former throughout the process. McAvoy and Ye (1994) also give a significant importance to inventory control, dedicating one stage to determine level control loops. However, the production rate control is more detailed by Price et al. (1994).

Another difference between the methodologies is on the procedure to define the control loops that are not related to inventory and production rate control. McAvoy and Ye (1994) use interaction, stability and saturation analysis tools to determine the decentralized control loops. On the other hand, Price et al. (1994) complete their control structure by only defining product quality control loops.

With respect to production rate control, the methodology of Price et al. (1994), with the self-consistent concept, assures that changes in the production rate are totally propagated through the whole plant. Concerning this matter, McAvoy and Ye (1994) don't give any guideline to perform this kind of analysis.
Both methodologies strongly depend on the knowledge and experience of the engineer on the process. Therefore, it is very important to use some theoretical tools to check stability, saturation, interaction and any other problem that can occur with the control system. Even though both methodologies have a great applicability, the one proposed by McAvoy and Ye (1994) seems to be the most adequate for industrial applications, since it uses important analysis tools combined with process knowledge.
Also, the use of the self-consistent concept and the production rate manipulation analysis by Price et al. (1994) could be adapted to the methodology proposed by McAvoy and Ye (1994), generating a more complete procedure for control structure design.

7. Conclusions

The plantwide control literature presents several methodologies and procedures for control structure design. These methodologies, in most cases, have a process oriented approach, but some of them also use analysis and optimization tools.
The methodologies proposed by McAvoy and Ye (1994) and Price et al. (1994) were chosen to be studied on an ethanol production plant. For the first one, the stages were applied and the obtained control structure was tested to verify its performance in the presence of external disturbances. The control structure presented good results. For the methodology by Price et al. (1994), the given guidelines were followed to obtain a final control structure, which was also tested, showing good responses against external disturbances. Finally, a comparative analysis was made and the methodology by McAvoy and Ye (1994) was chosen the best for this industrial application.

References

Chemstations, CHEMCAD Simulation Suite.

T.P. Chiang, W.L. Luyben, 1988, Comparisons of the dynamic performances of three heat integrated distillation configurations, Ind. Eng. Chem. Res., v. 27, pp.99-104.

M. Han, S. Park, 1996, Multivariable Control of Double-effect Distillation Configurations, Journal of Process Control v. 6, n. 4, pp. 247-253.

A.M. Lenhoff, M. Morari, 1982, Design of resilient processing plants - I. Process design under consideration of dynamic aspects, Chemical Engineering Science, v. 37, pp.245-258.

T.J. McAvoy, N. Ye, 1994, Base Control for the Tennessee Eastman Problem, Computers Chem. Engng., v. 18, n. 5, pp.383-413.

J. Pohlmeier, A. Rix, 1996, Interactive Plant and Control Design of a Double-Effect Distillation Column, Computers and Chemical Engineering, v. 20, n. 4, pp. 395-400.

R.M. Price, P.R. Lyman, C. Georgakis, 1994, Throughput Manipulation in Plantwide Controle Structures, Ind. Eng. Chem. Res., v.33, pp. 1197-1207.

S. Skogestad, 2004, Control structure design for complete chemical plants, Computers and Chemical Engineering, v. 28, pp. 219-234.

G. Stephanopoulos, C. Ng, 2000, Perspectives on the synthesis of plant-wide control structures, Journal of Process Control, v.10, pp. 97-111.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Nonlinear Dynamic Process Monitoring using Canonical Variate Analysis and Kernel Density Estimations

P.P. Odiowei[a] and Y. Cao[a]

[a]*School of Engineering, Cranfield University, Bedford, MK43 0AL, UK*

Abstract

Amongst process monitoring techniques the Principal Component Analysis (PCA) and the Partial Least Squares Regression Analysis (PLS) assume that the observations at different times are independent. However, for most industrial processes, these assumptions are invalid because of their dynamic features. For dynamic processes, the Canonical Variate Analysis (CVA) based approach is more appropriate than the PCA and the PLS based approaches. The CVA model is linear and control limits associated with the CVA are traditionally derived based on the Gaussian assumption. However, most industrial processes are non-linear and the Gaussian assumption is invalid for such processes so that techniques based on this assumption may not be able to correctly identify underline faults. In this work, a new monitoring technique using the CVA with control limits derived from the estimated probability density function through kernel density estimation (KDE) is proposed and applied to the Tennessee Eastman Process Plant. The proposed CVA with KDE approach is able to significantly improve the monitoring performance compared to other methods mentioned above.

Keywords: Canonical Variate Analysis, Probability Density Function, Kernel Density Estimation, Process Monitoring

1. Introduction

Process monitoring is essential to maintain high quality products and process safety. Process monitoring techniques like the PCA and the PLS rely on static models and assume that the observations are independent in time and follow a Gaussian distribution. However, such assumptions are invalid for most chemical processes because measurements driven by noise and disturbances are strongly auto-correlated. Therefore the static PCA and PLS based approaches are inappropriate to monitor such dynamic processes. To address this issue, Ku et al.[1] proposed a dynamic extension of PCA (DPCA) by using lagged variables to develop dynamic models through parallel analysis. However, diagnosis of abnormal behaviour becomes more complicated with DPCA and the principal components extracted in this way are not necessarily the minimal dynamic representations[2].

More recently, monitoring techniques based on Canonical Variate Analysis (CVA) have been developed with control limits derived based on the Gaussian assumption[3,4,5]. CVA is a monitoring technique that is based on state variables which are linear combinations of the past measurements. Norvalis et al.[5] developed a process monitoring tool that combined CVA and knowledge based systems (KBS) with control limits of the T^2 metric. Juan and Fei[4] employed CVA for fault detection based on T^2 charts while

Chiang et al.[3] employed CVA to include both input and output variables to estimate state variables with control limits of the T^2 and Q metrics to measure the state and residual spaces.

Traditionally, control limits of the T^2 metric for the state space and noise space and the Q metrics are estimated based on an assumption that the latent or state variables follow a normal distribution. However, for most industrial processes, these assumptions are invalid because of their dynamic and non-linear properties. In such a case, the control limit estimated based on the Gaussian assumption is unable to correctly identify underline faults. The problem of monitoring non-Gaussian processes can be addressed by estimating the underline probability density function (PDF) of the T^2 and Q metrics through the kernel density estimation (KDE) to derive correct control limit[6,7,8]. Martin and Morris[6] presented an overview of the PCA and the PLS with control limits on T^2 and M^2 metrics. The M^2 metric was estimated based on the PDF and was more efficient than the T^2 metric. Chen et al.[7] adopted the PCA with KDE approach to monitor a gas smelter process and demonstrated that the KDE technique could be used as a tool to obtain nonparametric empirical density function for a more efficient process monitoring. Xiong et al.[8] presented a study of Independent Component Analysis with KDE to monitor a catalyzer reactor.

In this paper, control limits of the T^2 and Q metrics associated with CVA are estimated based on kernel density estimations resulting in a new extension of the CVA algorithm, the 'CVA with KDE' for nonlinear dynamic process monitoring. For comparison, PCA and PLS with and without KDE as well as CVA with and without KDE are considered in the present study. The monitoring performance is significantly improved by using the 'CVA with KDE' compared with other monitoring algorithms mentioned above. This paper is organised as follows: Section 2 explains the CVA model while section 3 describes monitoring metrics and their control limit derived through Kernel Density Estimations. The procedure of CVA with KDE is then summarised in section 4 and the application of the proposed CVA with KDE to the Tennessee Eastman Process plant described in section 5. Finally, the work is concluded in section 6. Up to authors knowledge, this is the first time the KDE approach has been applied to estimate control limits for CVA based monitoring techniques. This establishes the main contribution of this work.

2. Canonical Variate Analysis (CVA)

Canonical Variate Analysis (CVA) is a dimension reduction technique based on a linear dynamic state space model to estimate state variables for dynamic process monitoring. Assume $y(k) \in R^m$, for k=1, 2,…,N are process data collected under a normal operation condition. Variables that occur before time k, are described as the past variables ($y_p(k)$) while variables that occur after time k, are the future variables ($y_f(k)$). More specifically, the past and future output vectors are determined in (1) and (2), respectively. Based on these definitions, the past and future Hankel output matrices are obtained in (3) and (4), respectively.

$$y_p(k) = [y^T(k-1), y^T(k-2), y^T(k-3), \ldots. y^T(k-p)]^T \in R^{mp} \quad (1)$$

$$y_f(k) = [y^T(k), y^T(k+1), y^T(k+2), \ldots. y^T(k+f-1)]^T \in R^{mf} \quad (2)$$

$$Y_p = [y_p(k-1) \;\; y_p(k-2) \;\; y_p(k-3) \ldots y_p(k-M)] \in R^{mp \times M} \quad (3)$$

$$Y_f = [y_f(k) \;\; y_f(k+1) \;\; y_f(k+2) \ldots.. \; y_f(k+M-1)] \in R^{mf \times M} \quad (4)$$

where mp and mf are the lengths of the past and future output vectors respectively and M is the number of lags. CVA maximises the correlation between the past and future variables to explain the information in the process data. Precisely, the covariance matrices of the past and future Hankel matrices are $\sum_{pp}$ = E($Y_p Y_p^T$) and $\sum_{ff}$ = E($Y_f Y_f^T$) respectively, and the cross covariance between the future and the past Hankel matrices is $\sum_{fp}$ = E($Y_f Y_p^T$). However, the realisation of CVA requires the covariance of the past and future and the cross covariance between past and future Hankel matrices to be conditioned against any singularities by taking their square roots[2,4]. From the square root of the covariance matrices, a scaled Hankel matrix (H) is formed and the singular value decomposition (SVD) of H is shown in (5).

$$H = \sum_{ff}^{-1/2} \sum_{fp} \sum_{pp}^{-1/2} = U \sum V^T \tag{5}$$

The full singular value (SV) matrix $\sum \in R^{mf \text{ x } mp}$ contains the SVs in the diagonal of $\sum$ in a descending order. U $\in$ $R^{mf \text{ x } mf}$ contains the left singular-vectors and V $\in R^{mp \text{ x } mp}$ contains the right singular-vectors. The ratio of the specific SV to the sum of all the SVs is employed to determine n, which is the order of the system[4,9]. The state variables are linear combinations of the transformation matrix (J) which is given in as;

$$J = U_n^T \sum_{pp}^{-1/2} \tag{6}$$

In this way, the state variables estimated in (7) have identity covariance matrix, **I**.

$$x_k = U_n^T \sum_{pp}^{(-1/2)} y_p(k) = J y_p(k) \tag{7}$$

Hence the state variable will have n linear combinations of the past in the state space and (mp-n) linear combinations in the noise space.

3. Control Limit through Kernel Density Estimations

For monitoring, the T^2 and Q metrics are applied to the states derived in (7) to monitor the state and residual spaces respectively. The T2 metric is defined in (8)

$$T^2(k) = [x_k - \bar{x}]^T S^{-1} [x_k - \bar{x}] \tag{8}$$

for the state space as well and the noise space, where $\bar{x}$ is the mean of the sample from which the CVA model is developed, while S is the covariance and is estimated in (9).

$$S = \frac{1}{N-1} \sum_{k=1}^{N} [x_k - \bar{x}][x_k - \bar{x}]^T \tag{9}$$

N is the number of samples. The residual vector (e) is defined as;

$$e = (I - J^T J) y_p(k) \tag{10}$$

From (10) above, the Q-statistic is estimated as in (11) below.

$$Q = e^T e \tag{11}$$

One solution of monitoring non-Gaussian processes is to estimate the PDF directly for the T^2 and Q metrics by a non-parametric approach such as the KDE[6,7,8,10]. Assume x is a random variable and the density function is denoted by p(x). This means that

$$P(a < x < b) = \int_a^b P(x) dx \tag{12}$$

Therefore, an appropriate control limit can be derived for a specific confidence bound from (12) if p(x) is known. To estimate the PDF p(x) at point x, the kernel density estimator $\hat{p}_h(x)$ is defined in (13) based on the assumption that Xi is a set of random samples with size N, K is a kernel function and h is the bandwith.

$$\hat{p}_h(x) = \frac{1}{Nh}\sum_{i=1}^{N} K(\frac{x - X_i}{h}) \tag{13}$$

Then $\hat{p}_h(x)$ determined in (13) is the estimated PDF and based on the density estimation, appropriate control limits corresponding to the given confidence level is determined to detect the underline faults in the non-Gaussian process data.

4. CVA with KDE

The procedure of the proposed CVA with KDE algorithm consisting of the CVA modelling, the KDE estimation and the monitoring stage is shown below;

Building the CVA model from the normal operating process

1. Collect data from normal operation for CVA
2. Determine past and future output vectors as in (1) and (2)
3. Determine the Hankel output matrices as in (3) and (4)
4. Estimate the covariance of the past, future and cross covariance between past and future hankel matrices as demonstrated in section 2 above
5. Estimate the scaled hankel matrix and perform SVD as in (5)
6. Estimate the canonical variates as in (6)
7. Estimate the state variables as in (7)

KDE Estimation

8. Estimate the T^2 and Q metrics as in (8) and (11)
9. Estimate the PDF of the T^2 and Q metrics using KDE in (13)
10. Determine the control limits corresponding to a given confidence interval based on (12).

Monitoring Stage

11. Collect fresh data for monitoring
12. Calculate current states and residual using fresh measurements
13. Calculate the T^2 and Q metric of current states and residual
14. Employ the control limits estimated from step 10 for process monitoring

5. Application- Tennessee Eastman Process Plant

The Tennessee Eastman Process (TEP) plant[11] has 5 main units which are the reactor, condenser, separator, stripper and compressor[3]. The TEP data consists of two blocks; the training and test data sets each of which has 52 measured variables and 22 scenarios corresponding to Fault 0 to Fault 21, with Fault 0 being the data simulated at normal operating condition (no fault) and Faults 1 – 21 corresponding to data sets from the simulated fault processes, each with a specified fault. For the purpose of this study, six TEP faults have been employed which are Faults 2, 6, 10, 16, 19 and 21. All the results in this study have been based on a 99% confidence level. The superiority of the proposed CVA with KDE is compared with other monitoring techniques using the Q metric as shown in Figure 1 while Figure 2 shows the advantage of the new method being able to detect faults earlier than other monitoring techniques.

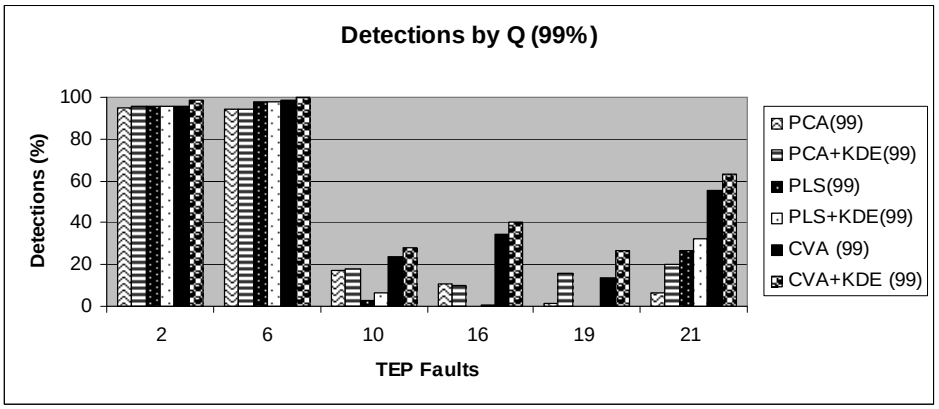


Figure 1 Comparison of All Monitoring Techniques (Q-99%)

For all six faults considered, Figure 1 clearly shows that the proposed CVA with KDE is able to significantly improve the monitoring performance compared with other monitoring techniques employed in this study, i.e. PCA, PCA with KDE, PLS, PLS with KDE and CVA only.

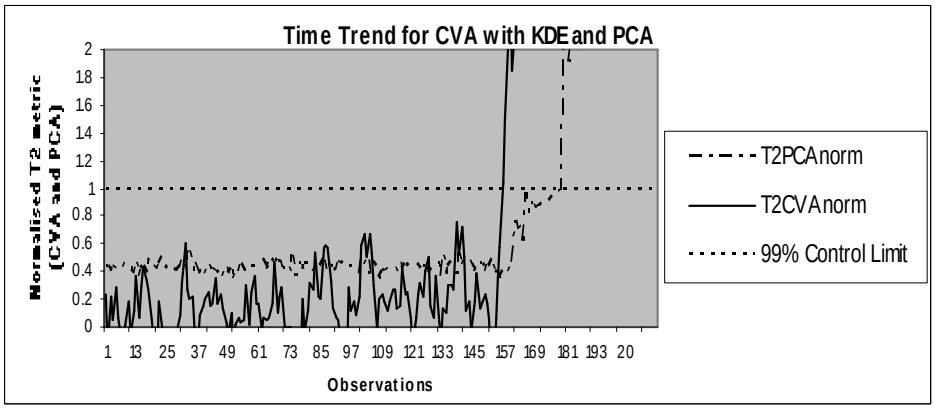


Figure 2 Time Trend for Fault 6 based on PCA and CVA with KDE

In addition to improving the performance significantly, the CVA with KDE approach is also able to detect faults earlier than other monitoring approaches so that operators have more time to deal with the underline fault before any disaster happens. For fault 6, Figure 2, shows two normalised T^2 metric curves generalised by the PCA and the CVA with KDE respectively. It shows that the PCA detects the fault at the 183rd observation while the CVA with KDE detects the fault at the 160th observation, which is about sixty nine minutes earlier than the PCA with a much more clear indication for a fault.

6. Conclusions

In this study a CVA with KDE approach is proposed for non-linear dynamic process monitoring, where due to the nonlinearity, the normal Gaussian assumption on measurements is invalid. This is the main contribution of this work. The Tennessee Eastman Process Simulation data was employed as the case study. The monitoring

performance of the proposed CVA with KDE was compared with PCA and PLS with and without KDE as well as CVA without KDE. The time trends were also investigated for the CVA with KDE compared with the time trend for PCA. The comparison demonstrates that the proposed CVA with KDE approach is able to improve the monitoring performance over other monitoring techniques mentioned above. Furthermore, the CVA with KDE is found to be able to detect faults earlier than the PCA. In summary, the CVA with KDE as a tool for nonlinear dynamic process monitoring is more appropriate and more efficient than the traditional CVA and other approaches mentioned above.

References

1. W. F. Ku, H. R. Storer and C. Georgakis (1995). Disturbance Detection and Isolation by Dynamic Principal Component Analysis. Chemometrics and Intelligent Laboratory Systems, Pages 179 – 196.

2. A. Negiz and A. Cinar. Monitoring of Multivariable Dynamic Processes and Sensor Auditing. Journal of Process Control, Vol. 8, No. 56, 1998, Pages 375 – 380.

3. L. H Chiang, E. L. Russell and R. D. Braatz (2001). Fault Detection and Diagnosis in Industrial Systems. Springer, London, Pages 85-98, 103-109.

4. L. Juan and L. Fei. Statistical Modelling of Dynamic Multivariate Process using Canonical Variate Analysis. Proceedings of IEEE International Conference of Information and Automation, Colombo, Sri Lanka, 15 -17, December 2006, Pages 218-221.

5. A. Norvalis, A. Negiz, J. DeCicco and A. Cinar (2000). Intelligent Process Monitoring by Interfacing Knowledge-Based Systems and Multivariate Statistical Monitoring. Journal of Process Control, Vol. 10, Pages 341 – 350.

6. E. B. Martin and A. J. Morris (1996). Non-Parametric Confidence Bounds for Process Performance Monitoring Charts. Journal of Process Control, Vol. 6, No. 6, Pages 349-358.

7. Q. Chen, P. Goulding, D. Sandoz and R. Wyne. Application of Kernel Density Estimates to Condition Monitoring for Process Industries. Proceedings of the American Control Conference, Philadelphia, PA, USA, 21-26 June 1998.

8. L. Xiong, J. Liang and J. Qian, J (2007). Multivariate Statistical Process Monitoring of an Industrial Polypropylene Catalyzer Reactor with Component Analysis and Kernel Density Estimation. Chinese Journal of Chemical Engineering, Vol. 15, No. 4, Pages 524 – 32.

A. Negiz and A. Cinar (1997). PLS, Balanced and Canonical Variate Realization Techniques for Identifying VARMA models in State Space. Chemometrics and Intelligent Laboratory Systems, Vol. 38, Pages 209 – 221.

T. Chen, J. Morris and E. Martin (2006). Probability Density Estimation via an Infinite Gaussian Mixture Model: Application to Statistical Process Monitoring. Appl. Statist, Vol. 55, Part 5, 2006, Pages 699 – 715.

N. L. Ricker (2001). Tennessee Eastman Challenge archive. Available at: http://depts.washington.edu/control/Larry/TE/download.html.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Multi-Period Continuous-Time Formulation for Integrated Scheduling, Blending, and Distribution of Refinery Products

Jie Li,[a] I. A. Karimi,[a] Rajagopalan Srinivasan[a,b]

[a]*Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117576*
[b]*Process Sciences and Modeling, Institute of Chemical and Engineering Sciences, 1 Pesek Road, Jurong Island, Singapore 627833*

Abstract

In this paper, we develop a slot-based multi-period mixed integer linear programming (MILP) formulation for an integrated treatment of recipe, specifications, blending, storage, and distribution, and incorporate many real-life features such as multi-purpose product tanks, parallel non-identical blenders, minimum run lengths, changeovers, etc. To enforce constant rates during blending runs, we develop a novel and efficient procedure that solves successive MILPs instead of a non-convex MINLP. We use fourteen examples of varying sizes and features to illustrate the superiority and effectiveness of our formulation and solution approach.
Keywords: Gasoline, recipe, blending and distribution, non-convex mixed-integer nonlinear programming (MINLP), property index

1. Introduction

The overall refinery operations (Pinto et al. 2000) involve crude oil storage and processing, intermediate processing, and product blending and distribution. Scheduling of crude oil operations (Reddy et al. 2004a,b; Li et al. 2007) has received the most attention so far. However, only limited work exists on the scheduling of product blending and distribution operations.

Gasoline is one of the most profitable products of a refinery and can account for as much as 60-70% of total profit. However, this process involves nonlinear blending and complex combinatorics, and can easily result in suboptimal schedules and costly quality give-aways. The large numbers of orders, delivery dates, blenders, blend components, tanks, quality specifications, etc. make this problem highly complex and nonlinear.

So far, several works in the literature (Pinto et al., 2000; Glismann & Gruhn, 2001; Mendez et al., 2006) have addressed the problem of product blending operations and incorporated some real-life features such as variable recipes, identical parallel blenders, etc. However, these works did not integrate the distribution operations with blending, and did not force the blending rate to be constant in a run and minimum run length. Jia and Ierapetritou (2003) proposed an MILP model for scheduling gasoline blending and distribution operations simultaneously. However, their model lacked many key operation features such as multiple parallel non-identical blenders, variable recipes, etc.

In this paper, we develop a multi-period MILP formulation that incorporates several real-life operation features such as multi-purpose product tanks, parallel non-identical blenders, minimum run lengths, changeovers, linear property indices (Li et al. 2007), piecewise constant profiles for blend component qualities and feed rates, etc. To ensure

constant rates in blending runs, we develop a novel schedule adjustment procedure that solves only MILPs and no non-convex MINLP.

2. Problem Statement

Consider a gasoline blending and distribution unit (GBDU) in a typical refinery (Figure 1). It employs I component tanks (i = 1, 2, …, I), B blenders (b = 1, 2, …, B), J product tanks (j = 1, 2, …, J), and some lifting ports. Each component has a distinct, and known quality or specification, and component i is stored in its own dedicated component tank i. At time zero, it has O orders (o = 1, 2, …, O) to fulfill during the coming scheduling horizon [0, H]. Each order o involving a single product has a time window [DD_o^L , DD_o^U] for delivery. Any delivery after DD_o^U incurs a demurrage cost (DM_o). The quality of blend components is specified in terms of various property indices such as RBN (Research Octane Number Index). With this, the GBDU problem addressed in this paper can be stated as:

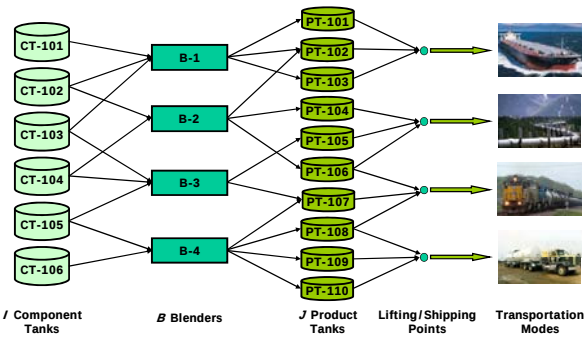


Figure 1 Schematic of gasoline blending and distribution

Given:
1. I components, their property indices and P products, their property indices limits.
2. I component tanks, initial inventories, capacity limits, feed flow profiles into tanks.
4. B blenders, allowable products, minimum blending length, and blending rates limits.
5. J Product tanks, allowable products stored in each tank, capacity limits, initial products and holdups, and delivery (lifting) rates for various products.
6. O orders, their constituent products, amounts, and delivery time windows.

Determine:
1. The blenders that each component tank should feed over time, and at what flow rates.
2. The products that each blender should produce over time, and at what rates.
3. The products that each product tank should receive over time, from which blender, and at what flow rates.
4. The orders that each product tank should deliver over time and their amounts.

Assuming:
1. Flow rate profile of each component from the upstream process is piecewise constant.
2. Component quality profile is also piecewise constant.

Other assumptions and some operating rules are not presented here.

3. Multi-Period MILP

Given the rate profiles of feeds into component tanks and profiles of component qualities, we divide the horizon H into T periods (t = 1, 2, …, T) of lengths H_t such that

the flow rates of components and component qualities are constant in each period and $H = H_1 + H_2 + \ldots + H_T$. We follow the approach used by Karimi and McDonald (1997) in their second model (M2). Thus, we divide each period into several process-slots (Liu and Karimi, 2007) of variable lengths. The process-slots are common to or synchronized across all units (tanks and blenders). Let K (k = 1, 2, …, K) be the total number of slots and $\boldsymbol{TK}$ = {(t, k) | slot k is in period t}. We define T_k as the time at which slot k ends and SL_k as the length of slot k. We fix some T_k to be the period ends with the upper bound of T_K being H. Denoting T_0 as the end of slot k = 0.

$$T_k = T_{(k-1)} + SL_k \qquad T_0 = 0,\ 0 < k \le K \qquad (1)$$

In the following, sets $\boldsymbol{BJ}$ = {(b, j) | blender b can feed product tank j}, $\boldsymbol{BP}$ = {(b, p) | blender b can process product p}, and $\boldsymbol{JO}$ = {(j, o) | product tank j may deliver order o}.

3.1. Blending and Storage

We use a dummy product tank (j = 0) to model the idle status of a blender. Thus, we have J real product tanks (j = 1, 2, …, J) and one (j = 0) dummy product tank. We define binary variables v_{bjk}, and u_{jpk}, and 0-1 continuous variables ue_{jk}, x_{bpk}, and xe_{bk}:

$v_{bjk} = \begin{cases} 1 & \text{If blender } b \text{ feeds product tank } j \text{ in slot } k \\ 0 & \text{Otherwise} \end{cases}$ $u_{jpk} = \begin{cases} 1 & \text{If product tank } j \text{ holds product } p \text{ in slot } k \\ 0 & \text{Otherwise} \end{cases}$

$ue_{jk} = \begin{cases} 1 & \text{If tank } j \text{ switches products at the end of } k \\ 0 & \text{Otherwise} \end{cases}$ $x_{bpk} = \begin{cases} 1 & \text{If blender } b \text{ processes product } p \text{ during } k \\ 0 & \text{Otherwise} \end{cases}$

$xe_{bk} = \begin{cases} 1 & \text{If blender } b \text{ ends the current run in slot } k \\ 0 & \text{Otherwise} \end{cases}$

Each blender must feed exactly one product tank (real or dummy) in each slot.

$$\sum_{j=0}^{J} v_{bjk} = 1 \qquad (b, j) \in \boldsymbol{BJ},\ 0 < k \le K \qquad (2)$$

3.2. Run Length

We define RL_{bk} as the length of the current run of blender b at the end of slot k, if the run does not end during slot k, and zero otherwise. Thus, RL_{b0} = 0, if a run has ended at time zero, otherwise it is the current run length at time zero. To compute RL_{bk}, we write,

$$RL_{bk} \le RL_{b(k-1)} + SL_k \qquad 0 < k \le K \qquad (3)$$

$$RL_{bk} \le H(1 - xe_{bk}) \qquad 0 \le k \le K \qquad (4)$$

Then, to ensure a minimum length (RL^L_{bp}) for each blend run, we demand,

$$RL_{b(k-1)} + SL_k + \left\{\max_p \left(RL^L_{bp}\right)\right\}(1 - xe_{bk}) \ge \sum_{p=1}^{P} RL^L_{bp} x_{bpk} \qquad (b, p) \in \boldsymbol{BP},\ 0 < k \le K \qquad (5)$$

3.3. Variable Blending Rate

If blender b is not idle during slot k, then Q_{bk} must be limited by the maximum (F^U_b) and minimum (F^L_b) of blender b.

$$Q_{bk} \le F^U_b \cdot SL_k \qquad 0 < k \le K \qquad (6a)$$

$$Q_{bk} + F^L_b \cdot H \cdot (v_{b0k} + xe_{bk}) \ge F^L_b \cdot SL_k \qquad 0 < k \le K \qquad (6b)$$

Eq. 6 allows the blending rate to vary from slot to slot during a run. Normally, this is not done in practice. However, enforcing this makes the formulation nonlinear and nonconvex. Therefore, we have decided to deal with this issue later.

3.4. Order Delivery

We define one binary variable (z_{jok}) as follows to denote order delivery operation.

$$z_{jok} = \begin{cases} 1 & \text{If product tank } j \text{ is delivering order } o \text{ during slot } k \\ 0 & \text{Otherwise} \end{cases}$$

If tank j is not delivering o during k, then the delivery amount (DQ_{jok}) must be zero:

$$DQ_{jok} \le TQ_o \cdot z_{jok} \qquad (j, o) \in \mathbf{JO},\ 0 < j \le J,\ 0 < k \le K \qquad (7)$$

Eq. 7 allowed order delivery to be intermittent from a tank. We propose a simple procedure later to correct this situation.

3.5. Inventory Balance on Component

Within any period, the feed rate (F_{it}) of i from upstream units is constant, so the inventory of component i at the end of slot k [V_{ik} ($V_i^L \le V_{ik} \le V_i^U$)] is:

$$V_{ik} = V_{i(k-1)} + \sum_{t \in \mathbf{TK}} F_{it} SL_k - \sum_{b=1}^{B} q_{ibk} \qquad 0 < k \le K \qquad (8)$$

3.6. Objective Function

The objective is to minimize the total operating cost including material (component), transition and demurrage costs.

$$\text{Minimize } TC = \sum_{i=1}^{I}\sum_{b=1}^{B}\sum_{k=1}^{K} c_i \cdot q_{ibk} + \sum_{b=1}^{B}\sum_{k=1}^{K-1} CB_b \cdot xe_{bk} + \sum_{j=1}^{J}\sum_{k=1}^{K-1} CT_j \cdot ue_{jk} + \sum_{o=1}^{O} DM_o \cdot d_o \qquad (9)$$

where, c_i is the price ($ per unit volume) of component i, CB_b is the cost ($ per occurrence) of transition on blender b, CT_j is the cost ($ per occurrence) of transition in product tank j, and DM_o is the demurrage cost ($ per unit time) of order o.

This completes our multi-period model (MPM), which comprises eqs. 1-9 and other equations such as product qualities, and inventory balance on product tank, which are not presented here. As mentioned before, it allows the blending rate to vary from slot to slot and order delivery to be discontinuous, which is undesirable in practice. Therefore, we need a procedure to adjust the solution from MPM to obtain a realistic schedule.

4. Schedule Adjustment

The optimal solution from MPM gives us the values of x_{bpk}, xe_{bk}, v_{b0k}, SL_k, and Q_{bk}. We use [x_{bpk}], [xe_{bk}], [v_{b0k}], [SL_k], and [Q_{bk}] respectively to denote their optimal values. The run lengths (CRL_{bk}) and volumes (CCQ_{bk}) are computed by,

$CRL_{bk} = CCQ_{bk} = 0$ if [xe_{bk}] = 1 *(10a,b)*

$CRL_{bk} = CRL_{b(k-1)} + [SL_k]$ if [xe_{bk}] = 0 *(11a)*

$CCQ_{bk} = CCQ_{b(k-1)} + [Q_{bk}]$ if [xe_{bk}] = 0 *(11b)*

Then, we compute the total volume (TCQ_{bk}) processed by a blender in a run as:

$TCQ_{bk} = 0$ if [xe_{bk}] = 0 *(12a)*

$TCQ_{bk} = CCQ_{b(k-1)} + [Q_{bk}]$ if [xe_{bk}] = 1 *(12b)*

The blending rate (R_{bk}) for each blending run at the slot where it ends is computed:

$$R_{bk} = \max\left(F_b^L,\ \frac{CCQ_{b(k-1)} + [Q_{bk}]}{CRL_{b(k-1)} + [SL_k]} \right) \text{ for } k \text{ with } [xe_{bk}] = 1 \,\&\, [v_{b0k}] = 0 \qquad (13)$$

Then, we set R_{bk} for all slots within each run to be the same as the one computed above. Now, to obtain a realistic schedule with the constant blend rates computed above, we fix x_{bpk}, xe_{bk}, and v_{b0k}. This allows us to fix, remove, or change some variables and constraints in MPM. For instance, eq. 6 becomes:

$$Q_{bk} = 0 \qquad \text{for } (b, k) \text{ with } [v_{b0k}] = 1 \qquad (14a)$$

$$Q_{bk} = R_{bk} SL_k \qquad \text{for } (b, k) \text{ with } [xe_{bk}] = [v_{b0k}] = 0 \qquad (14b)$$

$$Q_{bk} \le R_{bk} \cdot SL_k \qquad \text{for } (b, k) \text{ with } [xe_{bk}] = 1 \ \&\ [v_{b0k}] = 0 \qquad (14c)$$

The revised model (RMPM) comprises eqs. 1-2, 7-9, 14 and other equations, whose solution ensures that blending campaigns have constant blend rates that are within the limits on the blending rates and minimum run lengths at the same time. The schedule from RMPM may still show intermittent delivery of orders. When the delivery is over contiguous slots, then this can be easily revised by simply delivering at a constant rate until the entire order, which is distributed over contiguous slots, is fully delivered.

The complete procedure is described in brief as follows. We first solve MPM model which allows the blending rates to vary from slot to slot. Then, we compute CRL_{bk}, CCQ_{bk}, and blending rate R_{bk} for each run. Fixing x_{bpk}, xe_{bk} and v_{b0k}, we solve RMPM model to obtain a realistic schedule which ensures constant blending rate in a run. Finally, we correct intermittent delivery of orders to obtain final schedule.

Table 1 Solution statistics of various algorithms/codes for MPM for Examples 1-14

Ex	Order	Algorithm	Discrete Variables	Continuous Variables	Constraints	Total CPU Time (s)	Cost (k$)
1	5	DICOPT	130	334	1393	78.7	5149.73
		BARON	130	334	1393	14400*	5149.73
		Ours	130	328	1384	12.6	5149.73
2	10	DICOPT	329	735	3609	798	3658.11
		BARON	329	735	3609	14400*	3678.11
		Ours	329	729	3600	137	3658.11
4	15	DICOPT	486	883	4849	4074	4576.67
		BARON	486	883	4849	14400*	4717.13
		Ours	486	877	4840	800	4556.67
8	20	DICOPT	1159	1797	12878	14400*	12495.55
		BARON	1159	1797	12878	14400*	9492.56
		Ours	1159	1779	12848	10821	8329.13
10	25	DICOPT	1685	2458	19006	14400*	N/A
		BARON	1685	2458	19006	14400*	18223.18
		Ours	1685	2436	18968	13215	11649.17
14	45	DICOPT	4830	6493	57756	118800*	N/A
		BARON	4830	6493	57756	118800*	N/A
		Ours	4830	6439	57657	118800*	19207.48

Note: CPU time limit for MIP of DICOPT is set at 10800 s for Exs 1-10, 36000 s for Ex 11, and 108000 s for Exs 12-14
CPU time limit for MPM of ours is set at 10800 s for Exs 1-10, 36000 s for Ex 11, and 108000 s for Exs 12-14
CPU time limit for RMPM of ours is set at 3600 s for Exs 1-10, and 10800 s for Exs 11-14
Total CPU time limit of DICOPT, BARON, and ours is set at 14400 s for Exs 1-10, 46800 s for Ex 11, and 118800 s for Exs 12-14
* Reached total CPU time limit N/A: No feasible solution

5. Detailed Evaluation

Recall that forcing the blending rate to be constant during a run makes the formulation nonlinear and nonconvex, which can be solved using commercial MINLP solvers such as BARON and DICOPT. Our adjustment procedure obviated the need to solve

MINLPs. We solve 14 examples with widely varying sizes, structures, scale, and complexity to evaluate our formulation and solution approach. Table 1 shows the solution statistics only for Examples 1-2, 4, 8, 10, and 14. For Example 1, our procedure needs only 12.6 CPU s for the optimal solution of 5149.73 k$, but DICOPT needs 78.7 CPU s and 14400 CPU s for BARON. For Example 2, our procedure obtains the optimal solution of 3658.11 k$ within 137 CPU s, but DICOPT needs 798 CPU s. BARON gets a worse solution of 3678.11 k$ after 14400 CPU s. For Example 4, our approach obtains the optimal solution of 4556.67 k$, but DICOPT and BARON do not. For Examples 3 and 5 (not in Table 1), DICOPT does get the optimal solutions, but requires an order of magnitude longer solution times compared to our procedure. For instance, DICOPT takes 531 CPU s for Example 3 versus only 83 CPU s for our procedure. Interestingly, BARON also reaches the optimal solution for Example 3, but needs 14400 CPU s. For the remaining examples (Examples 6-14), our approach always obtains better solutions than both DICOPT and BARON within the allocated CPU time. For instance, our approach finds a solution of 8329.13 k$ for Example 8, while DICOPT obtains 12495.55 k$, and BARON gets 9492.56 k$. Moreover, our approach obtains a solution of 11649.17 k$ for Example 10, while BARON gets 18223.18 k$, and DICOPT cannot obtain a feasible solution.

6. Conclusion

We developed a slot-based multi-period continuous-time model for integrated scheduling of gasoline blending and distribution operations in a refinery and incorporated many real-life operating features and policies. We proposed an ingenious schedule adjustment procedure that requires only MILP solutions to ensure constant blending rate in a run. On 14 test problems of varying sizes and features, our proposed procedure was superior to commercial solvers (DICOPT and BARON). Much further work is needed, as our model still cannot optimally solve truly large problems involving more than 30 orders within reasonable time.

Acknowledgement: The authors would like to acknowledge financial support from The Agency for Science, Technology, and Research (A*Star) under grant 052 116 0074.

References

K. Glismann, G. Gruhn, 2001, Short-term scheduling and recipe optimization of blending processes, *Comput Chem Eng*., 25, 627-634.

Z. Y. Jia, M. Ierapetritou, 2003, Mixed-integer linear programming model for gasoline blending and distribution scheduling, *Ind. Eng. Chem. Res.*, 42, 825-835.

I. A. Karimi, C. M. McDonald, 1997, Planning and scheduling of parallel semicontinuous processes. 2. short-term scheduling, *Ind Eng Chem Res*., 36, 2701-2714.

J. Li, W.K. Li, I. A. Karimi, R. Srinivasan, 2007, Improving the robustness and efficiency of crude scheduling algorithms, *AIChE J.*, 52, 10, 2659-2680.

Y. Liu, I. A. Karimi, 2007, Scheduling multistage, multiproduct batch plants with nonidentical parallel units and unlimited intermediate storage, *Chem Eng Sci.*, 62, 1549-1566.

C. A. Mendez, I. E. Grossmann, I. Harjunkoski, P. Kabore, 2006, A simultaneous optimization approach for off-line blending and scheduling of oil-refinery operations, *Comput Chem Eng.*, 30, 614-634.

J. M. Pinto, M. Joly, L. F. L. Moro, 2000, Planning and scheduling models for refinery operations, *Comput Chem Eng.*, 24, 2259-2276.

P. C. P. Reddy, I. A. Karim, R. Srinivasan, 2004a, A new continuous-time formulation for scheduling crude oil operations, *Chem. Eng. Sci.*, 59, 1325-1341.

P. C. P. Reddy, I. A. Karimi, R. Srinivasan, 2004b, A novel solution approach for optimizing scheduling crude oil operations, *AIChE J.*, 50, 6, 1177-1197.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

NARX-Model-Based Control (NARX-MBC) for Citronellyl Laurate Esterification Reactor

S.A. Zulkeflee, N.Aziz*

School of Chemical Engineering, Engineering Campus, Universiti Sains Malaysia, Seri Ampangan, 14300 Nibong Tebal, Pulau Pinang Malaysia
**Phone: +604 -5996457, Fax: +604 -5941013, Email: chnaziz@eng.usm.my*

Abstract

Model-based control nowadays appears to be a very promising control strategy for various processes. However, this type of control strategy requires an accurate, low complexity, easy identifiability, structural flexibility and if possible, invertibility type of models. In this paper, a non-linear autoregressive exogenous-model-based control (NARX-MBC) has been designed and implemented to a batch citronellyl laurate esterification reactor. Multi-input-single-output (MISO) model has been developed to representing the process using NARX modelling approach. The performance of the developed NARX-MBC is evaluated and compared with a conventional PID controller. Overall, it is observed that the former has outperformed the latter.

Keywords: NARX, model based control, model identification, batch esterification.

1. Introduction

Batch reactors are frequently used in chemical, petrochemical or biochemical industries, for various quality products. Bioreactors are complex, as they generally involve subsystems that have numerous internal states, and require a number of parameters with highly nonlinear inter-relationships. As a result, the modeling and control of bioreactors is difficult and due to this reason, the application of model-based control strategies often gives better control performance (Arpornwichanop et. al., 2005).

In recent years, several researchers have implemented model based control strategies in a batch reactor (Konakom et. al., 2008; Zhang, 2008; Oliveira, 2004; Preub et. al., 2003; Aziz et. al., 2000). However, to the best of author's knowledge, there is no work carried out to implement model based control strategies for batch esterification process. In this work, NARX-Model-Based-Control (NARX-MBC) is designed and applied to control the reactor temperature of batch citronellyl laurate esterification process. Ability to track the desired temperature will lead the process reaching to the required ester conversion as demanded by the customers.

Nonlinear autoregressive with exogenous inputs (NARX) model is able to provide a powerful representation for time series analysis, modeling and prediction due to its strength to accommodate the dynamic, complex and nonlinear nature of real time series applications (Harris and Yu, 2007; Mu et. al.,2005). Due to such advantages, in this work NARX model has been chosen to represent the process under consideration. NARX model has been used as identification model that can capture the non-linear dynamics relating the inputs to the output of the system.

For control purposes, the developed NARX model is embedded in the internal model control strategy and known here as NARX-Model-Based-Control (NARX-

MBC). Finally, NARX-MBC performance is evaluated and compared with a conventional PID controller.

2. Citronelly laurate esterification process

In this work, the citronellyl laurate esterification process reaction (Aziah et. al., 2006) taking place in the jacketed batch reactor has been considered (Figure 1). The specification of reactor chosen in this study is similar to the batch reactor rig available at the Process Control Laboratory, School of Chemical Engineering, Universiti Sains Malaysia. This process is a well-mixed, heterogeneous with exothermic reaction scheme. Figure 2 shows the schematic that represent esterification of citronellyl laurate where citronellol and lauric acid are the reactants. Citronellyl laurate (ester) is the desired product while water is an unwanted byproduct. The immobilized *candida rugosa* lipase on to amberlite MB-1 is used as catalyst for the synthesis of the citronellyl laurate.

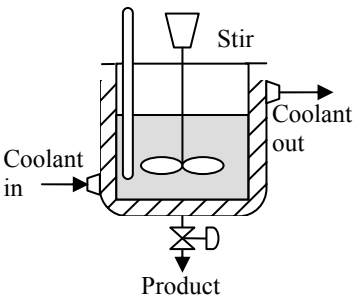


Figure 1: Schematic diagram of jacketed batch reactor

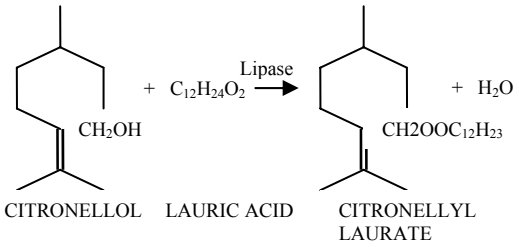


Figure 2: Schematic represents esterification of citronellyl laurate

The kinetic data for the model is obtained from analysis of the initial rate data and progress curve data and the reaction conducted with immobilized lipase follows an ordered bi-bi mechanism with dead-end of lauric acid (Garcia et. al., 2000).

3. NARX model

NARX model is a nonlinear generalization of a well known ARX model, which constitute a standard tool in linear black-box identification. A NARX model μ(θ) is characterized by the one-step ahead predictor:

$y(t|\theta)=F(x(t))$

where the vector $x(t)=[y(t-1),...Y(t-n),u(t-1),...u(t-m)]$ represents the memory of the model. Since the function F is, to a large extent, unknown it is approximated by an expression in a given function basis. The function ***F(.)*** can include both linear and nonlinear functions of ***x(t)*** as shown in Figure 3. Data for the input $u(t)$ and output $y(t)$ in this process is generated from the validated first principle models (Zulkeflee and Aziz, 2008) with operating conditions as shown in Table 1.

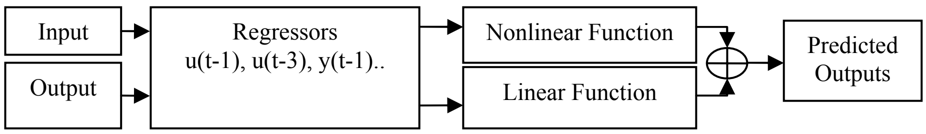


Figure 3: NARX blocks diagram.

Table 1: Operating Conditions and Calculated Parameters

Parameter	Values	Parameter	Values	Parameter	Values
A_{Ac}	18.20 L mol/s	Cp_{Ac}	420.53J/mol K	α	1
A_{Al}	24.04 L mol/s	Cp_{Al}	235.27J/mol K	β	1
A_i	0.32 L mol/s	$Cp_{Es\backslash}$	617.79J/mol K	U	2.86J/sm^{2}K
E_{Ac}	-105.40J mol/K	Cp_w	75.40J/mol K	A	0.077 m^2
E_{Al}	-66.09 J mol/K	V	1.5 L	ΔH_{rxn}	19.84 kJ
E_i	-249.94J mol/K	V_j	0.8 L	T_{ji}	294 K

NARX model use the wavelet network as a nonlinearity estimator. The wavelet network function is based on the following function expansion:

$$F(x)=(x- r)\ PL + a_{s1}f((b_{s1}(x- r))Q - c_{s1}) + a_{sn}f((b_{sn}(x- r))Q - c_{sn})+ ...$$
$$a_{w1}g\ ((b_{w1}(x-r))\ Q - c_{w1}) + a_{wn}wg\ ((b_{wn}w(x- r))Q - c_{wn}) + d$$

From the above equation, ***f*** is a scaling function and ***g*** is the wavelet function. ***P*** and ***Q*** are *m*-by-*p* and *m*-by-*q* projection matrices, respectively. The projection matrices ***P*** and ***Q*** are determined by principal component analysis of estimation data. Usually, $p=m$. If the components of ***x*** in the estimation data are linearly dependent, then $p<m$. The number of columns of ***Q***, *q*, corresponds to the number of components of ***x*** used in the scaling and wavelet function. Its structure and parameters are tabulated in Table 2.

Table 2: NARX model obtained with model analysis

Scaling Translation, *cs_k*	Scaling Coefficient, *as_k*	Scaling Dilation, *bs_k*
1 0 1 0 1 1 1 0 -1 0 -1 -2 0 1 1 0 0 0 1 -1 -1 1 1 0 1 1 -1 1 0 1 -1 0 0 1 1 1 1 1 1 -1 0 -1 0 -2 1 0 0 0 -1 0 0 1 1 -1 0 1 -1 -1 1 -1 -1 -1 2 0 1 -1 -1 0 1 1 -1 0 1 0 0 1 1 0 0 -2 1 0 0 1 0 0 0 1 1 -2 1 1 0 1 0 1 1 1 -1 0 1 0 1 1 1 1 1 -1 0 1 0 1 0 -1 0 1 0 1 0 1 0 1 0 1 2 2 1 -1 1 0 0 -2 2 -1 0 0 1 -1 -1 0 1 0 -1 -1 0 0 -1 0 0 1 0 0 -1 1 0 1 1 0 1 -1 1 0 1 -2 1 1 0 -2 1 0 0 1 1 -1 -1 0 1 -2 1 0	-0.0260 0.0382 -0.0579 0.0671 0.0336 0.0875 0.0468 0.0306 0.0670 -0.0220 -0.0298 0.0259 0.0034 0.0268 0.0009 -0.0290 0.0303 -0.0170 0.0011 -0.0014	[1;1;1;1;1;1;1;1;1;1;1;1;1;1; 1;1;1;1;1] Output Offset, *d* 320.9573 Regressor Mean, *r* [320.9097 320.85797 320.8021 0.0085 0.0084 0.0084 293.1141 293.1074 293.10076] Wavelet Dilation, *bw_k* [1;1;1;1;1;1;1;1;1;1;1;1;1;1; 1;1;1;1;1;1;1;1;1;1;1;1;1;1;1 ;1;1;1;1;1;1;1;1;1;1;1;1;1;1; 1;1;1;1;1;1;1;1;1;1; 1;1;1;1;1;1;1;1;1]

R-value for the state-space, ARX and NARX model validation achieved are 0.4849, 0.8382 and 0.9367 respectively. It is show that NARX performed better than the other

two models and thus will be used in the control section. Figure 4 shows the output profile for validation between identified NARX model and validation data.

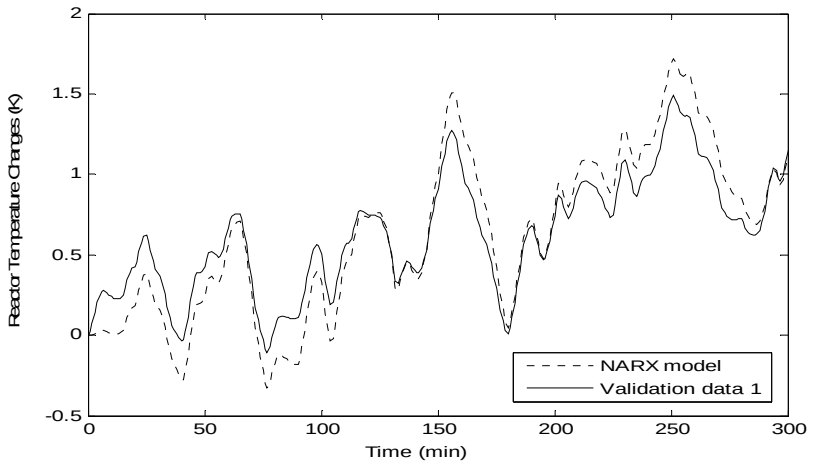


Figure 4: Simulation of model output compared against a validation data set

4. NARX Model-Based Control (NARX-MBC)

The IMC structure (Garcia & Morari, 1982) given in Figure 5 is a central to our discussions on designing the controller. Its conceptual usefulness lies in fact that it allows us to concentrate on the controller design without having to concern with control system stability provided that the process model $\tilde{p}(s)$ is a NARX model representation of a stable process. In this work, the process $p(s)$ is represented by the first principle model of the esterification batch reactor. The controlled variable of the system considered is temperature of the reactor (Tr) whereas the manipulated variable is cooling water flow (Fj). Meanwhile, the cooling water temperature (Tj) is considered as disturbance. The controller $q(s,\varepsilon)$ of this system is the inverse model of the process $p^{-1}(s)$.

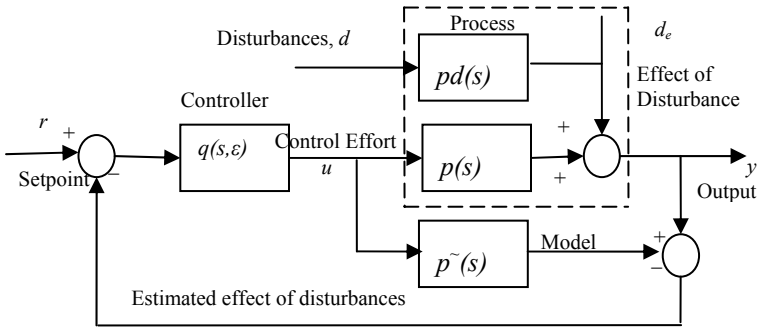


Figure 5: The IMC system

The inverse model is derived from first order plus time delay that developed using NARX model. Model parameters at different steady state points in esterification reaction are identified by Cohen Coon method (Seborg et . al. 2004) and the results are tabulated in Table 3. Worst case of model parameters of the process is chosen to develop first order model (Bhapa et. al. 2005). In this method the step test with largest process gain (i.e. $K_p = 82405$) and smallest time constant (i.e. τ=5.2047) were chosen as

the worst case for the process. The primary controller takes the form $q(s,\varepsilon)=p\text{-}1(s)f$, where f is a user specified low-pass filter;

$F(s,\tau_c)=1/(\tau_c+1)^r$

where r is sufficiently positive integer in order to guarantee that the IMC controller is proper. The r =1 and τ_c=0.01.

$$q(s,\varepsilon)=\frac{5.2047\,s+1}{82405\,(\frac{1}{0.01s+1})}$$

Table 3: Identified Model Parameters at different steady state points

Cases (L/min)	0-0.002	0.002-0.004	0.012-0.014	0.014-0.016	0.016-0.018
K_p	82405	81700	22490	19610	17633
Tp	26.863	20.193	5.7052	5.2047	7.8423

5. Results and Discussion

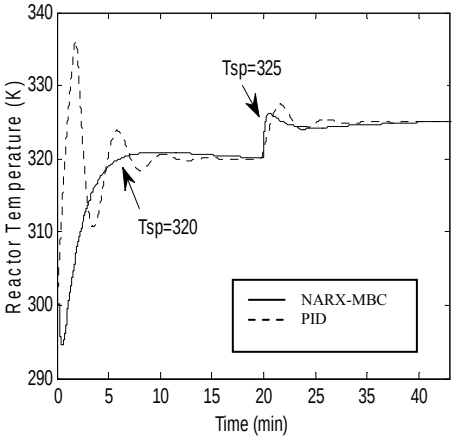


Figure 6: Set-point change for NARX-MBC and PID controller

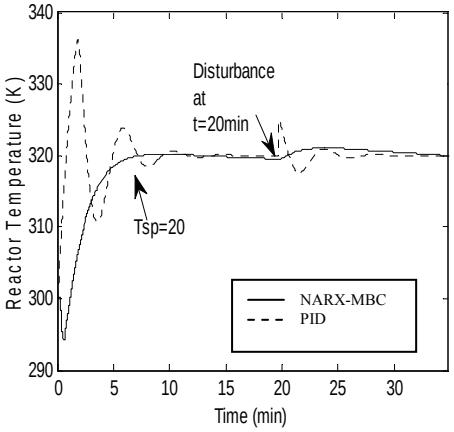


Figure 7: Disturbance rejection for NARX-MBC and PID Controllers

The responses for set point tracking and disturbance rejection for both NARX-MBC and PID controllers are shown in Figures 6 and 7 respectively. For both cases, NARX-MBC found to give faster and smoother closed loop compared to PID. The result also show that PID controller has difficulty in following the desired temperature profile and lead to a poor control performance. As shown in Figures 6 and 7, PID controller found to produce sluggish damping oscillations response for both set point change and disturbance rejection. On the other hand, NARX-MBC produce response with low amount of overshoot and less oscillate. NARX-MBC also able to settle faster compared to PID. Overall, NARX-MBC provides superior performance in set-point tracking and disturbance rejection than PID.

6. Conclusions

NARX-MBC has been designed and applied to the batch esterification process to produce citronellyl laurate. MISO NARX model to represent the process has been developed and the validation result showed that it was more than 93% accurate. The

validated NARX was then embedded in the control strategy proposed. The results showed that the NARX-MBC can cope with both set point change and disturbance rejection. It was then compared with PID controller and the former found to outperform the latter.

Acknowledgements

The authors wish to acknowledge the financial support by Ministry of Science, Technology and Innovation (MOSTI), Malaysia through the NSF scholarship for the first author and sciencefund project no 03-01-05-SF0090.

References

A. Arpornwichanop, P. Kittisupakorn and I.M Mujtaba, 2005, On-line dynamic optimization and control strategy for improving the performance of batch reactors, Chemical Engineering and Processing 44, 101-114.

S.N. Aziah, A.H. Kamaruddin and W.S Long, 2006, Studies of reaction parameters on synthesis of citronellyl laurate ester via immobilized *Candida rugosa* lipase in organic media, Bioprocess and Biosystems Engineering 29, 253-260.

N. Aziz, M.A. Hussain and I.M. Mujtaba, 2000, Performance of different types of controllers in tracking optimal temperature profiles in batch reactors, Computers and Chemical Engineering 24, 1069-1075

P.K. Bhaba, S. Sathishbabu, A. Asokan and T. Karunanithi, 2005, Real Time Implementation of Wiener Model PI (WMPI) Controller in a Conical Tank Liquid Level Process,ICCBPE/SOMChe, 297-302.

T.J. Harris and W. Yu, 2007, Controller assessment for a class of nonlinear systems, Journal of Process Control 17 607-619

C.E. Garcia and M. Morari, 1982, Internal model control: A unifying review and some new results. Ind. Eng. Chem. Process Des. Dev. 21 (2), 308-323

T. Garcia, A. Coteron, M. Martinez and J. Aracil, 2000, Kinetic model for the esterification of oleic acid and cetyl alcohol using immobilized lipase as catalyst, *Chemical Engineering Science, 55*, 1411-1423

K. Konakom, P. Kittisupakorn and I.M. Mujtaba, 2008, Batch control improvement by model predictive control based on multiple reduced-models, Chemical Engineering Journal.

J. Mu, D. Rees and G.P. Liu, 2005, Advanced controller design for aircraft gas turbine engines. Control Engineering Practice 13 1001–1015

R. Oliveira, 2004, Combining first principles modelling and artificial neural networks: a general framework. Computers and Chemical Engineering 28, 755-766.

K. Preub, M.V. Lann, M. Cabassud and G. Anne-Archard, 2003, Implementation procedure of an advanced supervisory and control strategy in the pharmaceutical industry, Control Engineering Practice, 11, 1449-1458.

D.E. Seborg, T.F. Edgar and D.A. Mellichamp, 2004, Process Dynamics and Control, 2nd edition.

J. Zhang, 2008, Batch-to-batch optimal control of a batch polymerisation process based on stacked neural network models. Chemical Engineering Science 63 1273– 1281

S.A. Zulkeflee and N. Aziz, 2008, Modeling and simulation of citronellyl laurate esterification in batch reactor. In Proceedings of Regional Symposium on Chemical Engineering (RSCE) in conjunction with Symposium of Malaysian Chemical Engineers, Kuala Lumpur, Malaysia (SOMChE), Vol II, 955-960.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Nonlinear model predictive control of a distillation column using NARX model

K. Ramesh, S. R. Abd Shukor and N. Aziz*

School of Chemical Engineering, Engineering Campus, Universiti Sains Malaysia, Seri Ampangan, 14300 Nibong Tebal, Pulau Pinang, Malaysia
** Phone: +604 -5996457, Fax: +604 -5941013, Email: chnaziz@eng.usm.my*

Abstract

Distillation column is an important process unit in petroleum refining and chemical industries, and needs to be controlled close to optimum operating conditions because of economic incentives. Nonlinear model based control (NMPC) scheme is one of the best options to be explored for proper control of distillation columns. In this work, NMPC scheme using sigmoidnet based nonlinear autoregressive with exogenous inputs (NARX) model has been developed to control distillation column The Unscented Kalman Filter (UKF) was used to estimate the state variables in NMPC and the nonlinear programming (NLP) problem was solved using sequential quadratic programming (SQP) method. The closed loop control studies have indicated that the NARX NMPC performed well in disturbance rejection and set point tracking.

Keywords: Distillation column, NARX model, Nonlinear model predictive control.

1. Introduction

Model Predictive Control (MPC) is an important advanced control technique which can be used for difficult multivariable control problems (Goodwin *et al.*,2001). The current generation of commercially available MPC technology is based on linear dynamic models, and is refereed by the generic term linear model predictive control (LMPC). Many processes such as high purity distillation column, multi-grade polymer reactors are sufficiently nonlinear to preclude the successful application of LMPC technology. This has led to the development of nonlinear model based controllers such as nonlinear model predictive control (NMPC) in which more accurate nonlinear model is used for process prediction and optimization.

Many authors have studied the performance of NMPC to control distillation using different nonlinear models namely semi-rigorous reduced order model (Maiti and Saraf, 1995), NARX model (Sriniwas *et al.*, 1995), Hammerstein model (Fruzzetti *et al.*, 1997), Recurrent Dynamic Neuron Network (RDNN) model (Shaw and Doyle III, 1997) and grouped neural networks (GNN) model (Ou and Rhinehart, 2003). Foss *et al*. (1998), in their case study on process modeling in Germany and Norway concluded that despite the commercially available modeling tools, the effort spent for all kinds of modeling activities is the most time consuming step in an industrial project where model based process engineering techniques are applied.

The NMPC problem formulation involves online computation of a sequence of manipulated inputs which optimize an objective function and satisfy process constraints. NMPC requires online solution of a nonlinear program (NLP) at each iteration. The solution of such NLP problems can be very time consuming, especially for large scale systems. An additional complication is that the optimization problem generally is

nonconvex because the nonlinear model equations are posed as constraints (Cannon,2004). Consequently, NLP solvers designed for convex problems may converge to local minima or even diverge. So it is necessary to find out an improved solution algorithm for nonconvex NLP problems.

In this paper, two multiple-input-single-output (MISO) sigmoidnet based NARX models were used to model the dynamics of the distillation column. An equilibrium model for distillation column was used as plant model in nonlinear system identification and in NMPC.

2. Sigmoidnet based NARX model

Two multiple-input-single-output (MISO) NARX models are developed in this study to model the dynamics of the distillation column. The reason for using two MISO models instead multiple-input-multiple-output (MIMO) model is that the MISO models provide better prediction compared to MIMO model (Eskinat *et al.*,1991) The first MISO model using reflux flow rate (L) and reboiler heat load (Q_R) as inputs, and top product composition (x_D) as output, while, the second MISO model using reflux flow rate (L) and reboiler heat load (Q_R) as inputs, and bottom product composition (x_B) as output. The models consisted of parallel combination of nonlinear and linear blocks.

In both the MISO NARX models used this work, sigmoidnet function with 2 units is used as nonlinear regression function containing two past output regressors and two past input regressors (one from each input). The parameters of sigmoidnet based MISO NARX models were estimated using the *ident* function system identification toolbox version 7.0 in MATLAB. The iterative prediction-error minimization method discussed in Ljung (1999) was used to calculate the model parameters. The data generated from equilibrium model for distillation column discussed in Ramesh *et al*. (2005) was used for parameter estimation and model validation. The model structure, parameter estimation, validation and model analysis are discussed in detail by Ramesh *et* a.l (2008).

3. NMPC

The developed nonlinear wavenet based Hammerstein model is of the following form.

$$X(k+1) = F[X(k), U(k)] \tag{1}$$

$$Y(k) = h[X(k)] \tag{2}$$

where $X(k) = [x_1(k)\ x_2(k)]^T$; $U(k) = [u_1(k)\ u_2(k)]^T$; $Y(k) = [y_1(k)\ y_2(k)]^T$

x_1 and x_2 are n-dimensional vector of state variables, u_1 (reflux flow rate) and u_2 (reboiler heat load) are m-dimensional vectors of manipulated input variables, and y_1 (top product composition) and y_2 (bottom product composition) are p-dimensional vector of controlled output variables. In this work, two separate MISO models were developed (one for each output) instead of using a MIMO model. The optimization problem is given by

$$\min_{\boldsymbol{U}(k\backslash k), U(k+1\backslash k), \ldots, U(k+M-1\backslash k)} \quad J = \phi[\boldsymbol{Y}(k+P\backslash k)] + \sum_{j=0}^{P-1} L\left[Y(k+j\backslash k, U(k+j\backslash k), \Delta U(k+j\backslash k)\right] \tag{3}$$

where $U(k + j\backslash k)$ is the input $U(k + j)$ calculated from information available at time k, $Y(k + j\backslash k)$ is the output $Y(k + j)$ calculated from information available at time k, $\Delta U(k + j\backslash k) = U(k + j\text{-}1\backslash k) - U(k + j\backslash k)$, M is the control horizon, P is the prediction horizon and φ and L are nonlinear functions of their arguments. The functions φ and L can be chosen to satisfy wide variety of objectives and in this study, the quadratic functions of the following form in considered:

$$L = [Y(k+j \setminus k - Y_s(k)]^T Q[Y(k+j \setminus k - Y_s(k)] + [U(k+j \setminus k - U_s(k)]^T R[U(k+j \setminus k - U_s(k)] + \Delta U^T(k+j \setminus k) S \Delta U(k+j \setminus k) \quad (4)$$

$$\phi = [Y(k+P \setminus k) - Y_s(k)]^T Q[Y(k+P \setminus k) - Y_s(k)] \quad (5)$$

where $U_s(k)$ and $Y_s(k)$ are steady-state targets for U and Y respectively, and Q, R and S are positive-definite weighing matrices. The principal controller tuning parameters are M, P, Q, R, S and the sampling period Δt.

The reflux flow rate bounds are set to be [0.1, 0.75] l/min. The lower bound for reflux flow rate 0.1 l/min was meant to keep the input physically meaningful, namely, the reflux flow rate should be positive and have some minimum value. The upper bound of reflux flow rate is approximately 150% of the nominal capacity, which would seldom occur in operation. The reboiler heat load bounds are set to be [0, 15] kW. The lower bound meant that the reboiler heat load should not be negative, whereas upper bound 15 kW was the maximum heater capacity of the reboiler.

The top product composition bounds are set to be [0.5, 1]. The lower bound for output 0.5 was meant that the top product purity should not be less than 50%. The upper bound 1 was meant that the maximum value of top product purity is 100% and beyond that is practically not meaningful. The bottom product composition bounds are set to be [0, 0.5]. The lower bound for bottom product composition 0 was meant that the maximum value of bottom product purity is 100%. The upper bound of bottom product composition 0.5 was meant that the bottom product purity should not be less than 50%. The desired product purity is achieved by solving the nonlinear optimization problem subject to the following inequality constraints.

$$U_{\min} \le U(k+j \setminus k) \le U_{\max}, \quad 0 \le j \le M-1 \quad (6a)$$

$$\Delta U_{\min} \le \Delta U(k+j \setminus k) \le \Delta U_{\max}, \quad 0 \le j \le M-1 \quad (6b)$$

$$Y_{\min} \le Y(k+j \setminus k) \le Y_{\max}, \quad 1 \le j \le P \quad (6c)$$

In addition, the nonlinear model equations are posed as a set of following equality constraints:

$$X(k+j+1 \setminus k) = F[X(k+j \setminus k), U(k+j \setminus k)], \quad 0 \le j \le P-1 \quad (7a)$$

$$Y(k+j \setminus k) = h[X(k+j \setminus k)], \quad 1 \le j \le P \quad (7b)$$

where $X(k\setminus k) = X(k)$ if the state variables are measured. It is important to note that input constraints are hard constraints in the sense that they must be satisfied. Conversely, output constraints can be viewed as soft constraints because their violation is necessary to obtain a feasible optimization problem.

NMPC calculation requires measurements or estimates of the state variables and in the present work, UKF was used to estimate the state variables in the NMPC problem. In UKF, the state distribution is represented by a Gaussian Random Variables (GRV), which is specified using a minimal set of carefully chosen sample points. These sample points completely capture the true mean and covariance of the GRV, and when propagated through the true non-linear system, captures the posterior mean and covariance accurately to the third order of Taylor series expansion for any nonlinearity (Wan and Merwe,2000). Finally, the NMPC problem was solved using *fmincon* function in MATLAB optimization toolbox version 3.1.1which uses a sequential quadratic programming (SQP) method.

4. Closed loop control studies

A sampling interval of 1 min was chosen for closed loop control studies. The NMPC parameters M, P, Q, R, S and the sampling period Δt are chosen by repeated tuning and the final values are; sampling period Δt = 1 min, prediction horizon P = 30 time steps (30 min), control horizon M = 6 time steps (6 min), Q = (1,0.5), R = (1,1) and S = (1,1). The equilibrium model for distillation column discussed in Ramesh *et al.* (2005) was used as the plant model

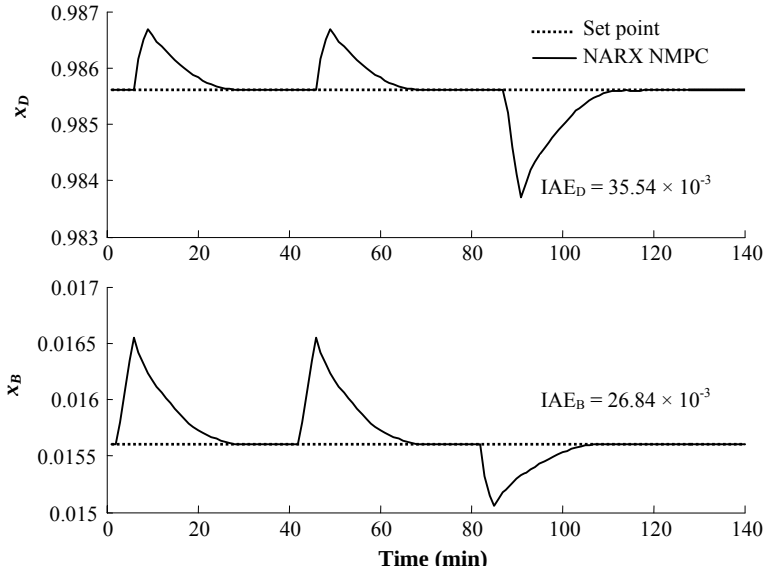


Figure 2: NARX NMPC responses of product compositions to a feed flow rate disturbance

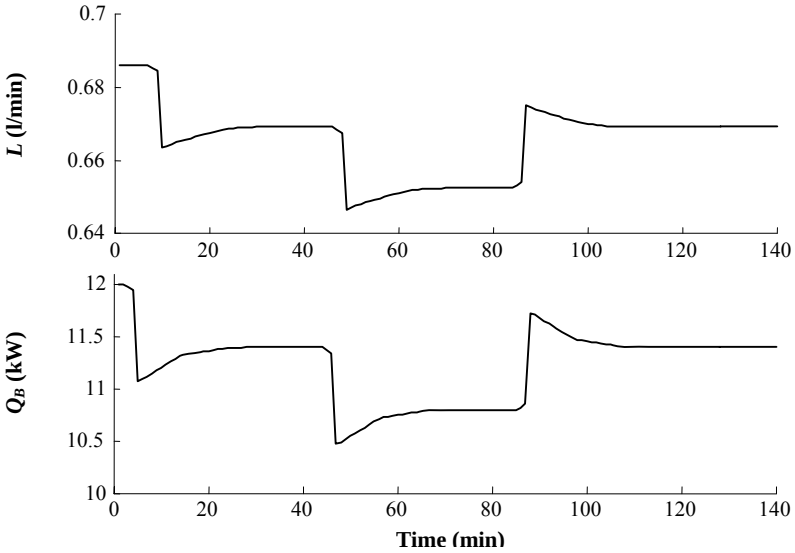


Figure 3: NARX NMPC responses of manipulated variables to a feed flow rate disturbance

The performance of the NMPC was studied by making three feed flow rate disturbances: a +20% increase at t = 0 min, again a +20% increase at t = 40 min and a -20% decrease at t = 80 min. The responses of product compositions for these feed flow rate disturbances are shown in Figure 2 along with the corresponding numerical values of IAE. It can be observed from the figure that NARX NMPC was reasonably rejected the feed flow rate disturbances. The corresponding responses of manipulated variables to feed flow rate disturbances are shown in Figure 3.

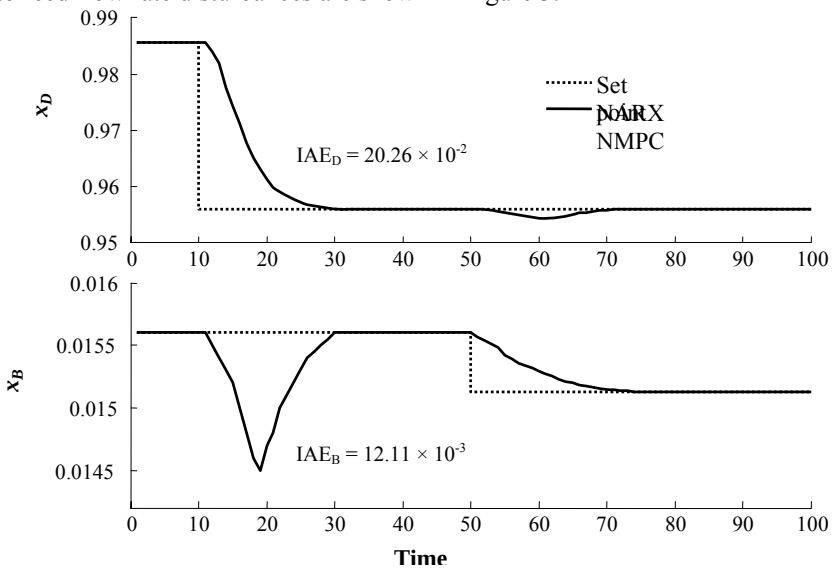


Figure 4: NARX NMPC responses of product compositions to -3% change in x_D at t =10 min and -3% change in x_B at t = 50 min

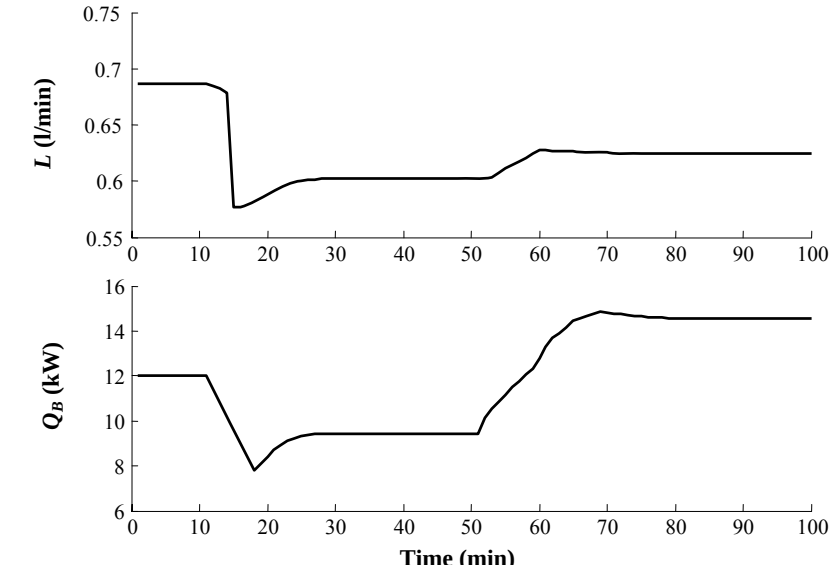


Figure 5: NARX NMPC responses of manipulated variables to -3% change in x_D at t =10 min and -3% change in x_B at t = 50 min

The performance of the NARX NMPC in set point tracking was studied by making -3% change in x_D at t = 10 min followed by -3% change in x_B at t = 50 min. The responses of product compositions for these set point changes are shown in Figure 4 and it can be noted from the figure that the NARX NMPC tracking the set point well. The corresponding responses of manipulated variables for set point changes are shown in Figure 5. It was noted that the reboiler should be operated very close to the upper limit of the Q_B in order to track the set point changes.

5. Conclusions

The sigmoidnet based NARX model NMPC was developed to control distillation column. The Unscented Kalman Filter (UKF) was used to estimate the state variables in NMPC and the NLP problem was solved using sequential quadratic programming (SQP) method. The closed loop control studies indicated that the developed NMPC technique performed well in controlling the distillation column by rejecting the disturbances in regulatory control and tracking the set point quickly in servo control.

Acknowledgement

Financial support by Ministry of Science, Technology and Environment (MOSTI), Malaysia through the IRPA – 8 project No: 03-02-4279EA019 is acknowledged.

References

M. Cannon 2004. Efficient nonlinear model predictive control algorithms. Annual Reviews in Control 28: 229-237.

E. Eskinat, S. H. Johnson and W. L. Luyben 1991. Use of Hammerstein models in identification of nonlinear systems. A.I.Ch.E. Journal 37(2): 255–268.

B. Foss, B. Lohmann and W. Marquadt 1998. A field study of the industrial modeling process. Journal of Process Control 8: 325–337.

K. P. Fruzzetti, A. Palazoglu and K. A. McDonald 1997. Nonlinear model predictive control using Hammerstein models. Journal of Process Control 7(1): 31-41.

G. C. Goodwin, S. F. Graebe and M. E. Salgado 2001. Control System Design. Upper Saddle River, NJ, Prentice Hall.

L. Ljung 1999. System Identification: Theory for the User. Upper Saddle River, NJ, Prentice Hall PTR.

S. N. Maiti and D. N. Saraf 1995. Adaptive dynamic matrix control of a distillation column with closed-loop online identification. Journal of Process Control 5(5): 315-327.

J. Ou and R. R. Rhinehart 2003. Grouped neural network model-predictive control. Control Engineering Practice 11(7): 723-732.

K. Ramesh, N. Aziz and S. R. Abd Shukor 2008. Development of sigmoidnet NARX model for a distillation column. Chemical Products and Process Modeling 3(2): Article 4.

K. Ramesh, M. Ramasamy, N. Aziz, S. R. Abd Shukor and M. Zailani Abu Bakar 2005. Comparison of rate- based and equilibrium models for multicomponent distillation. In Proceedings of International Conference on Chemical and Bioprocess Engineering 2005, Sabah, Malaysia.: 549-554.

A. M. Shaw and F. J. Doyle III 1997. Multivariable nonlinear control applications for a high purity distillation column using a recurrent dynamic neuron model. Journal of Process Control 7(4): 255-268.

G. R. Sriniwas, Y. Arkun, I. L. Chien and B. A. Ogunnaike 1995. Nonlinear identification and control of a high-purity distillation column: a case study. Journal of Process Control 5: 149-162.

E. A. Wan and R. V. D. Merwe 2000. The Unscented Kalman Filter for Nonlinear Estimation. In Proceedings of Symposium 2000 on Adaptive Systems for Signal Processing, Communication and Control (AS-SPCC), Lake Louise, Alberta, Canada., IEEE.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A neighbour in control technique for batch process monitoring

Carlos Alvarez, Adriana Brandolin, Mabel C. Sánchez

Planta Piloto de Ingeniería Química (UNS-CONICET), Camino La Carrindanga Km 7, Bahía Blanca 8000, Argentina

Abstract

In this work a Nearest in Control Neighbour (NICN) based strategy is presented for monitoring batch processes. The technique detects a fault if the T^2 statistic value exceeds the critical one for three successive observations. In this case, the identification of suspicious measurements is determined in terms of the distance between the current observation and its NICN. Furthermore a comparison between the NICN approach and the classical Multiway Principal Component Analysis (MPCA) is provided. Both techniques are applied to supervise the operation of a semibatch reactor for methyl-methacrylate polymerization. The analysis of results reveals that NICN strategy is suitable for fault detection of batch processes and has the advantage of avoiding ambiguous identifications when compared with classical techniques.

Keywords: Statistical Process Control, Multiway Principal Component Analysis, Original-variable Space.

1. Introduction

In the last few decades batch processes experienced a renaissance as products-on-demand and first-to-market strategies impel the need for flexible and specialized production methods. They are mainly devoted to the production of polymers, pharmaceuticals, foods, bio-chemicals and semiconductors.

Online process performance monitoring and product quality prediction in real time ensure safe and profitable operation of batch processes since they provide the opportunity to take corrective actions before the effects of excursions from normal operation ruin the batch.

Several multivariate statistical techniques have been proposed for online monitoring of batch processes. Multiway Principal Component Analysis (MPCA) and Multiway Partial Least Squares (MPLS) are some of the most commonly used in statistical process control. Using a model of the normal process behaviour in a reduced space of latent variables, these techniques monitor the process performance by comparing the time progression of the data with that of the average normal batch, which is obtained from the data base of normal operating conditions. A comprehensive discussion of PCA-based statistical techniques for monitoring batch processes can be found in the review work by Westerhuis et al. (2000).

Recently Alvarez et al. (2008) presented a novel approach for monitoring continuous processes in the original-variable space. The detection stage of the technique makes use of the T^2 statistic value to indicate the out of control state. In the identification stage, the distance between the Nearest In Control Neighbour (NICN) and the current observation is used to determine the contribution of each variable to the out of control state. Those variables whose distance measures exceed a certain threshold value are considered as

suspicious. The strategy has shown good performance for monitoring simple continuous processes.

In this work an extension of the NICN strategy to address batch monitoring is presented. Different threshold values are employed for identification purposes. The performance of the proposed technique is evaluated and compared with that provided by MPCA (Nomikos and MacGregor, 1994) for monitoring a polymerization batch reactor. The measures of performance are the number of false and missing alarms, the action to signal time, and the number of correct, ambiguous, wrong and null identifications.

The rest of the paper is structured as follows. In Section 2 the proposed strategy is described. Next the results of the performance comparison are presented and discussed. A "Conclusion" section closes the work.

2. A Nearest in Control Based Method for Batch Monitoring

A Multivariate Statistical Process Control strategy is devised based on the NICN concept. It has two stages that are performed off-line and on-line respectively.

2.1. Off- line Stage

Normal operation data of batch processes are commonly grouped in a three-way data matrix $\underline{\mathbf{Z}}$ ($I \times J \times K$), where I stands for the number of batches, J represents the number of variables and K indicates the number of observations of each variable during the batch run. After defining the Normal Operating Condition (NOC) data set, this stage involves the following steps:

a. Each column of $\underline{\mathbf{Z}}$ is centred and scaled to obtain matrix $\underline{\mathbf{X}}$. The mean $\bar{z}_{j,k}$ and the standard deviation $ds_{j,k}$ used to standardized the (j,k)-th column are stored.

b. The inverse of the correlation matrix of each time slide of $\underline{\mathbf{X}}$, $\mathbf{R}_k^{-1}$, is calculated and stored.

c. The T^2 statistic is calculated at each interval k for the members of the reference population. Then, the probability density-function of T^2 corresponding to interval k is estimated using a kernel smoothing technique. Given the significance level of the test, α, the (1- α) percentile of this distribution is selected as the critical value $T^2_{crit,k}$.

Dynamic Time Warping techniques (Ramaker et al. 2003) are successfully applied if the batches included in the NOC data set have different duration or some measurements are missed or delayed. Furthermore matrix inversion-related problems may appear if the number of batches is not sufficient, if the number of variables is quite large or if some of them are linear combinations of others. This problem might be addressed by applying regularization techniques to **R** in order to make it non-singular.

2.2. On- line Stage

During batch operation a measurement vector $\mathbf{z}_k$ of dimension J is tested for monitoring purposes at each time interval k. Testing $\mathbf{z}_k$ involves the following steps:

a. The standardized observation vector $\mathbf{x}_k$ is calculated using the mean $\bar{z}_{j,k}$ and the standard deviation $ds_{j,k}$ values corresponding to the j-th variable which are obtained from the NOC.

b. The value of the Hotelling's statistic for $\mathbf{x}_k$, $T^2_{x,k}$ is estimated as follows:

$$T^2_{x,k} = \mathbf{x}_k^T \mathbf{R}_k^{-1} \mathbf{x}_k \qquad (1)$$

where $\mathbf{R}_k^{-1}$ is an estimate of the correlation matrix (**Σ**) and is also extracted from the NOC.

c. The values of $T_{x,k}^2$ and $T_{crit,k}^2$ are compared to detect faults. Even if there are no systematic errors in measurements, the Hotelling's statistic value for an observation point may exceed the critical value $T_{crit,k}^2$ for a given significance level. This arises when one or more variables in the current observation vector do not behave as measurements in the normal population do. In this case an anomalous event is detected and next the faulty variables should be identified.

d. Fault identification. The knowledge about how far the faulty observation is from an in control allocation gives us an idea of the minimum distance and direction that need to be explained for an anomalous situation. This information can be obtained finding the nearest neighbour of the observation point that is in statistical control ($\mathbf{x}_{NICN,k}$). For this purpose an optimization problem is formulated whose objective is to minimize a distance measure between $\mathbf{x}_{NICN,k}$ and the measured point $\mathbf{x}_k$, subject to the constraint that the T^2-value for $\mathbf{x}_{NICN,k}$ is equal to or less than the critical one ($T_{crit,k}^2$). If the Mahalanobis distance is selected as the distance measure, the following optimization problem arises

$$\begin{aligned} &Min \quad (\mathbf{x}_k - \mathbf{x}_{NICN,k})^T \mathbf{R}_k^{-1} (\mathbf{x}_k - \mathbf{x}_{NICN,k}) \\ &s.t. \\ &\quad \mathbf{x}_{NICN,k}{}^T \mathbf{R}_k^{-1} \mathbf{x}_{NICN,k} \le T_{crit,k}^2 \end{aligned} \qquad (2)$$

The solutions of the first order optimality conditions for the constrained optimization problem are

$$\mathbf{x}_{NICN,k} = \pm \left(\frac{T_{crit,k}^2}{T_{x,k}^2} \right)^{1/2} \mathbf{x}_k \qquad (3)$$

As there exist only two possible solutions, the comparison of the objective function values at both solutions results easier than the evaluation of the second order optimality conditions to decide which one is the corresponding nearest neighbour.

The contribution of each variable to the fact that, T^2 for the current measurement is greater than $T_{crit,k}^2$, can be estimated in several ways using this information. In this work a simple method is implemented and evaluated. Since measurements are standardized, the movement in each direction that should be performed to reach the nearest in statistical control point from the measurement point is used as an estimate of the deviation degree of the variable. Therefore the directions whose changes are greater than a threshold value τ are associated with suspicious variables.

The procedure consists in comparing the absolute difference between x_k^j and $x_{NICN,k}^j$ for each variable *j* with a threshold value τ. If for a given variable this difference is greater than τ, it is identified as suspicious. Two values of τ are tested. One of them, τ_1, is the mean of the vector of absolute differences between $\mathbf{x}_k$ and $\mathbf{x}_{NICN,k}$. The other one, τ_2, is calculated using the mean plus the standard error of the same vector

$$\tau_1 = \frac{1}{J}\sum_{1}^{J}\left|x_k^j - x_{NICN,k}^j\right| \tag{4}$$

$$\tau_2 = \frac{1}{J}\sum_{1}^{J}\left|x_k^j - x_{NICN,j}^j\right| + \frac{ds(\left|\mathbf{x}_k - \mathbf{x}_{NICN,k}\right|)}{\sqrt{J}} \tag{5}$$

3. Application Results

The classic MPCA approach that uses contribution plots in the identification stage and the NICN method are applied for monitoring a non-isothermal semi-batch reactor for methyl-methacrylate emulsion polymerization.

For this purpose a NOC data set composed of 121 batches is obtained by simulation. It is assumed that eight measurements (monomer inlet flowrate Q_0, refrigerant inlet temperature T_{REF0}, jacket temperature T_{REF}, reactor temperature T, surfactant concentration in aqueous phase S_W, liquid holdup h_R, density ρ_R and molecular weight MW) are sampled every 2 minutes for an 80 minute run. The reactor model was developed and implemented in gPROMS code environment (Process System Enterprise, Ltd.) by Alvarez (2009).

Eighty two faults are simulated considering potential process-fault sources. Faults are arranged in eight groups. The first set of faults (1 to 23) comprises perturbations in Q_0. In some cases (1 to 5) Q_0 is fixed at a given value during the whole batch run. The remaining faults in this set (6 to 23) involve step increases or decreases of different sizes that occur at different times during the reactor operation.

The second set of faults (24 to 46) groups several perturbations in the value of T_{REF0}. For faults 24 to 28, T_{REF0} is fixed at a given value during the complete batch run. In contrast, faults in the range 29 to 46 correspond to ramp perturbations in T_{REF0}.

Finally, sets 3 to 8 present both positive and negative deviations in the initial conditions for the liquid holdup, initial surfactant amount, water volume fraction, initiator amount, monomer concentration and reactor temperature respectively.

The performance measures selected for comparison purposes are: the number of false alarms (*FA*), the number of missing alarms (*MA*), the action to signal time (*AST*), and the number of correct (*CI*), ambiguous (*AI*), wrong (*WI*) and null (*NI*) identifications.

The action to signal time accounts for the delay between the fault occurrence and its detection. In this work the *AST* is expressed in terms of the observations' ordinals. In order to obtain a common basis to average *AST* values for both techniques, alarm signals at k=42 (one period after finishing the operation) are considered for all missing alarms.

Furthermore the identification is considered correct when only the variable in which the fault was simulated is pointed out by the identification methodology. If this variable and others are indicated as suspicious, a misidentification occurs. A wrong identification arises if the real faulty variable is not identified as suspicious. Finally null identifications correspond to an empty set of faulty variables.

Results are summarized in Fig. 1 which presents the time intervals in which the faults are simulated (k_S) and detected (k_D) for α=0.05, the *AST* and the set of suspicious variables. They are represented using the numbers 2 or 3 depending on the thresholds τ_1 or τ_2 are surpassed. The level of significance of each statistic is selected in such a way that they have the same Average Number of Type I Errors under the null hypothesis.

Figure 1 shows that the last six simulations in group 1, which correspond to small positive and negative magnitude steps in Q_0, are not detected independently of the time interval in which the fault appears.

		MPCA																						NICN								
		k_S	k_D	AST_D	Q_0	T_{REF0}	T_{REF}	T	S_W	h_R	ρ_R	MW	k_D	AST_{SPE}	Q_0	T_{REF0}	T_{REF}	T	S_W	h_R	ρ_R	MW	k_D	AST_{T2}	Q_0	T_{REF0}	T_{REF}	T	S_W	h_R	ρ_R	MW
Faults Group 1	1	1	14	13	3						3		22	21	3				3				18	17	3							
	2	1	3	2	3								3	2	3	3		2	2				3	2	3							
	3	1	3	2	3								3	2	3	3		3	3		3	3	3	2	3							
	4	1	3	2	3								3	2	3	3							3	2	3							
	5	1	3	2	3			2			3		3	2	3	3	3	3	3		2		3	2	3							
	6	7	9	2	3				2				9	2	3		2			3			9	2	3							
	7	17	19	2	3							3	19	2	3				2				19	2	3							
	8	27	29	2	3								30	3	3								29	2	3							
	9	7	9	2	3								9	2	3	2	2		3			3	9	2	3							
	10	17	19	2	3								19	2	3								19	2	3							
	11	27	29	2	3								32	5	3				3				29	2	3							
	12	7	9	2	3								9	2	3								9	2	3							
	13	17	21	4	3								16	-1					3				21	4	3							
	14	27																														
	15	7	9	2									9	2	3				2				9	2	3							
	16	17	22	5	3			2			3												22	5	3						2	
	17	27																														
	18	7																														
	19	17																														
	20	27																														
	21	7																														
	22	17																														
	23	27																														
Faults Group 2	24	1	3	2		3					3		26	25			3					2	3	2		3						
	25	1																														
	26	1																														
	27	1	13	12		3			3		2																					
	28	1	3	2		3							15	14		3		3					3	2		3						
	29	1	15	14	3	3																	15	14		3		3				
	30	1	19	18		3																	19	18		3						
	31	1	27	26		3							25	24		3						2	22	21		3		3				
	32	1	14	13		3																	14	13		3	3					
	33	1	18	17		3	2																17	16		3						
	34	1	22	21		3	2	3															21	20		3						
	35	1	11	10		3					2	3	19	18		3							9	8	2	3		3				
	36	1	11	10		3						3	19	18		3						2	11	10		3						
	37	1	10	9		3							17	16		3				3			10	9		3					2	
	38	1	11	10		3							15	14		3	3			2	2		11	10		3						
	39	1	12	11		3	2				2		15	14		3			3	2			11	10		3						
	40	1	11	10		3							13	12	3	3							11	10		3						
	41	1	12	11		3							15	14		3		3					11	10		3						
	42	1	11	10		3							26	25	3	3	2						11	10		3						
	43	1	12	11		3							22	21	2	3	3		2	3		2	15	14		3						
	44	1	11	10	3	3							15	14		3	3						11	10	2	3		3				
	45	1	11	10		3							23	22		3	3						11	10		3						
	46	1	14	13		3					3		25	24		3		2	2				14	13		3	2				2	
F. Group 3	47	1	3	2						3	3		5	4		3				3			3	2					3			
	48	1	3	2						3	3		5	4		3			2	3			3	2					3			
	49	1	3	2						3			34	33			3			3		3	3	2					3			
	50	1	3	2						3	3		37	36			3				2		3	2					3			
	51	1	3	2						3			7	6	3	2	2	2		3			3	2					3			
	52	1	3	2						3			3	2	2	3				3	3		3	2					3			
F. Group 4	53	1											3	2					3				3	2								
	54	1											3	2	2				3				3	2					3			
	55	1																														
	56	1																														
	57	1																					3	2					3			
	58	1											36	35			3					3	3	2					3			
F. Group 5	59	1	17	16			3	3	3		3	3	7	6							3		15	14				3			2	2
	60	1	20	19			3	3		2		3	13	12				2			3	3	20	19			3	3				
	61	1																														
	62	1											28	27			3				2		28	27			3				2	
	63	1											34	33							3	3	25	24			3	2				
	64	1	24	23			3	3	2				15	14								3	24	23			3	2				
F. Group 6	65	1																														
	66	1																														
	67	1																														
	68	1																														
	69	1											24	23	2				3				26	25					2			3
	70	1											25	24	2				2		2	3										
F. Group 8	77	1	3	2				3	3		2	3	3	2		3	3	3				2	3	2				3				
	78	1	6	5				3					3	2				3	3				3	2				3				
	79	1	9	8	3								3	2		3		3					3	2				3				
	80	1	4	3									3	2			3	3	3				3	2					3			
	81	1	4	3	3		3	3					3	2		2	3	3	3		3		3	2				3				
	82	1	3	2			3	3	3		3	3	3	2		3	3	3					3	2				3	2			

Figure 1. Performance Results

The same happens with faults 14 and 17 although the step sizes (the first positive and the last negative) are considerable higher. This problem arises because faults occur during the last period of the reaction which is less sensitive and has high variability.

For faults 29 to 46, the *ASTs* are, in general, greater than those corresponding to previous faults for all the statistics. This behaviour is motivated by the fact that the simulated perturbations (temperature ramps) are not drastic enough to cause an alarm signal immediately. Faults in groups 5, 6 and 7 are changes in the initial conditions simulated in unmeasured variables. Therefore, they can be detected depending on their effects on measurements. As the identification stage provides no clear indication, operator's knowledge is always required in this sort of situations. It can be seen that many of the missing alarms occur in these groups. In fact, no one of the faults in set 7 was detected. In contrast, sets 3, 4 and 8 involve faults in the initial values of some measured variables, and the faulty variable is properly indicated in most of the cases. Results provided in Figure 1 are summarized in Table 1 for comparative purposes.

Table 1. Summary of performance results

		FA	*MA*	*AST*	*CI*	*AI*	*WI*	*NI*
MPCA	*D*	0	31	21	24	24	1	2
	SPE	1	30	23	4	44	3	0
NICN	T^2	0	25	18	41	15	1	0

The only *FA* obtained for the faulty batches' data set is given by *SPE* statistic for batch 13 at k=16. Table 1 shows that the smallest average value for *AST* and the lowest value of *MA* correspond to the T^2 statistic (NICN's metric). Furthermore it should be noticed that the highest values for *CI* as well as the lowest values for *AI* are achieved using the NICN approach. Regarding *WI*, there are no significant differences.

4. Conclusions

In this work a Statistical Control strategy for batch processes, which is applied in the space of the original variables, is presented. The distance between the observation and its nearest in-control neighbour is used to identify suspicious measurements.

A performance comparison between the NICN technique and MPCA for monitoring a batch polymerization reactor is conducted. Regarding the detection capability, higher values of *MAs* and *ASTs* are obtained when statistics are calculated in terms of latent variables for the tested set of faults. Furthermore, the proposed method shows a rewarding performance to correctly identify the fault source, avoiding ambiguous identifications. Regarding the threshold values, the higher threshold behaves better because it reduces the *AI* without increasing the *NI*.

References

C. Alvarez, 2009, Monitoreo de Procesos Batch. Aplicaciones a Reactores de Polimerización, PhD Thesis, Universidad Nacional del Sur.

C. Alvarez, A. Brandolin, M. Sánchez and L. Puigjaner, 2008, Proceedings of 2008 AIChE Annual Meeting.

P. Nomikos and J. Macgregor, 1994, Monitoring of Batch Process Using Multiway Principal Component Analysis, AIChE Journal, 40, 1361-1375.

J. Westerhuis, S. Gurden, A. Smilde, 2000, Generalized Contribution Plots in Multivariate Statistical Process Monitoring, Chemom. Intell. Lab. Syst., 51, 95-114.

H. Ramaker, E. N. M. Van Sprang, J. A. Westerhuis y A. K. Smilde, 2003, Dynamic Time Warping of Spectroscopic Batch Data, Analytica Chimica Acta, 498, (1-2), 133-153.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Fouling Management in Crude Oil Preheat Trains through Stream Split Optimization

Joana L. Borges,[a] Eduardo M. Queiroz,[a] Fernando L. P. Pessoa,[a] Fábio S. Liporace,[b] Sérgio G. Oliveira,[b] and André L. H. Costa,[c]

[a]*Federal University of Rio de Janeiro – UFRJ, School of Chemistry - Department of Chemical Engineering, Av. Horácio Macedo, 2030, CT, Bl. E, Cidade Universitária, CEP 21941-909, Rio de Janeiro - RJ, Brazil*
[b] *PETROBRAS/CENPES/PDEDS/Gás Natural/Célula de Otimização, Av. Horácio Macedo, 950 - Cidade Universitária, Rio de Janeiro, 21949-915, Brazil.*
[c]*Rio de Janeiro State University – UERJ, Institute of Chemistry - Department of Industrial Operations and Designs, Rua São Francisco Xavier, 524, Maracanã, CEP 20550-900, Rio de Janeiro – RJ, Brazil*

Abstract

In general, the first main step of petroleum refining consists in the distillation of the crude oil stream. In order to provide adequate fractionation, the crude stream must be fed in the atmospheric distillation column at about 380ºC. Aiming to reduce energy consumption, heat from hot streams of side products and pumparounds is transferred to the crude stream in a heat integration scheme, called crude preheat train. The final heating of the crude stream is executed in a furnace. However, during the operation of the preheat train; the thermal effectiveness of the heat exchangers diminishes due to fouling and, as a consequence, fuel costs increases. The large volumes of crude oil processed and the scenario of crescent energy prices justify the importance of this problem for the oil companies. Seeking to provide a solution to reduce the impact of this problem, this paper presents the exploration of stream splitting in crude preheat trains composed by several parallel branches. In this case, each branch may present different fouling levels, which allows the exploration of different distributions of the stream flow rates along the system, through a proper optimization algorithm. This optimization algorithm searches the set of stream splitters related to the maximization of the final temperature of the crude preheat train. A mathematical model of the preheat train works coupled to the optimization method. An important focus of this paper is to explore the introduction of constraints in order to guarantee feasible operating solutions, i.e., the optimum solution must attend different operational aspects related to bounds on fluid flow velocities and heat exchanger capacities. The performance of the proposed approach is illustrated through a typical example of a petroleum refinery.

Keywords: fouling, heat exchangers, oil refining.

1. Introduction

Conversion of crude petroleum into more valuable products involves many operations, such as, distillation, catalytic cracking, hydrotreating, delayed coking, etc. The first step frequently involves an atmospheric distillation column, sometimes preceded by a pre-flash column.

Crude petroleum must be pre-warmed to a satisfactory temperature before it is fed into those columns. Usually there is a huge consumption of energy at this stage. Therefore, aiming the reduction of fuel costs (with consequently decrease of gas emissions), energy integration of side-product streams and pumparounds with the crude stream is promoted. The final heating step is conducted in a furnace. Such structure is denominated preheat train.

An important operating problem of this kind of structure corresponds to the fouling of thermal surfaces. This phenomenon diminishes the thermal effectiveness and increase the furnace load. Recently, the literature presented several papers concerning the optimization of cleaning actions (Smaïli *et al.*, 2001; Lavaja and Bagajewicz, 2004).

In many preheat trains, the heat exchangers are organized in multiple parallel branches which can suffer fouling distinctly, because of different previous cleaning records or operating policies. Oliveira Filho *et al.* (2007) explored the optimization of stream splitters in multi-branch preheat trains aiming to maximize energy recovery and minimize operational costs.

In this context, this paper studies the application of the proposed technique in a real refinery, analyzing important questions in order to guarantee an actual operational solution: complexity of the control loops, maximum capacity of the coolers, impact of the variation in the pumparounds on the distillation performance and modification of fluid velocities in the fouling behavior. Economic results are also evaluated, seeking to measure the potentiality of the methodology.

2. Methodology

In a multi-branch preheat train, the final temperature depends on the flow distribution. Thus, the employed methodology is based on an association of a simulation algorithm to an optimization procedure. For each set of split fractions along the branches, a final crude temperature can be evaluated. The optimization seeks to find the best set of split fractions in order to reach the highest temperature (associated to the lowest fuel cost). The proposed scheme is illustrated in Figure 1.

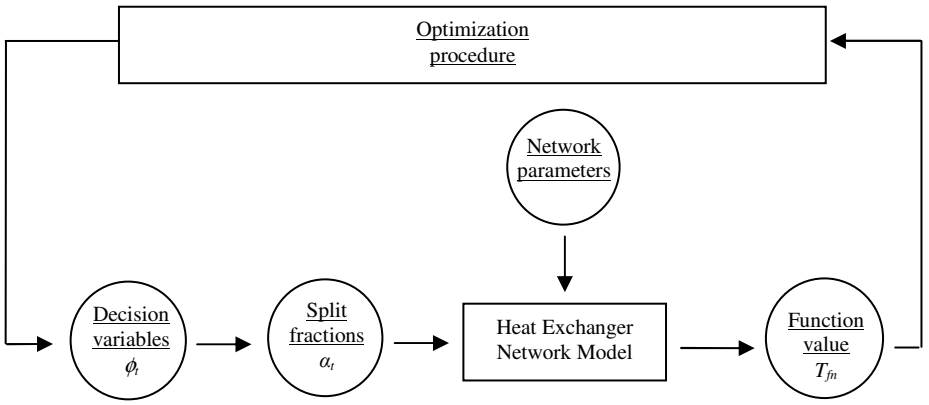


Figure 1 - Objective function evaluation scheme.

2.1. Feasibility issues

The feasibility of the application of the proposed approach in a real refinery involves the solution of some limitations of the existent hardware.

2.1.1. Complexity of the control loops

The maximum extension of the degrees of freedom would involves the insertion of control valves for all splitters. However, it can imply an unnecessary capital cost with process instrumentation. This analysis demands the investigation of different sets of control arrangements, comparing the energy reduction gain with the complexity of each control loop set.

2.1.2. Maximum capacity of the coolers

The maximization of the crude stream temperature implies, in certain situations, an increase of the inlet product stream temperature in a final cooler. This effect can influence the product storage (with a possible disproportional increase of cooling water demand). This aspect is handled by the insertion of constraints to avoid overcharging of the external heat exchangers.

2.1.3. Impact of the variation in the pumparounds on the distillation performance

Another possible disturbance that can be caused by the optimization procedure consists in the pumparound operation. In order to avoid a misbalance of the heat load along the distillation column, such problem is also solved by the insertion of bounds on these heat exchangers.

2.1.4. Modification of fluid velocities in the fouling behavior

Recently, Ishiyama *et al.* (2008) argued that the difference in flow rates of the branches, caused by the splitters misbalance, would accelerate the fouling process instead of contributing for the energy recovery. In fact, the fouling rate decreases with increasing flow velocity, as a consequence, the streams splitters optimization could generate results where a branch would receive a too low crude velocity that the fouling would consistently raise and the energy recovery of the system would be diminished. This aspect can be contemplated in the optimization through the imposition of velocities bounds in heat exchangers, thus avoiding very low velocities.

3. Example

The example describes a real network of a Brazilian refinery. The complete refinery crude preheat train presents 35 heat exchangers organized in five branches. The process flowsheet can be found in Figure 2. The first studied case (reference) corresponds to a simulation of the preheat train with all heat exchangers without fouling. The second case (base) is similar to the reference case, but considering the operational fouling of the refinery (at a specific day).

3.1. Analysis of the control loops

The investigation of the control loops is conducted by the examination of nine splitter schemes, described in Table 1, without constraints. For each scheme, the optimization was run and the final temperatures and the corresponding economic results are presented in Tables 2 and 3, respectively.

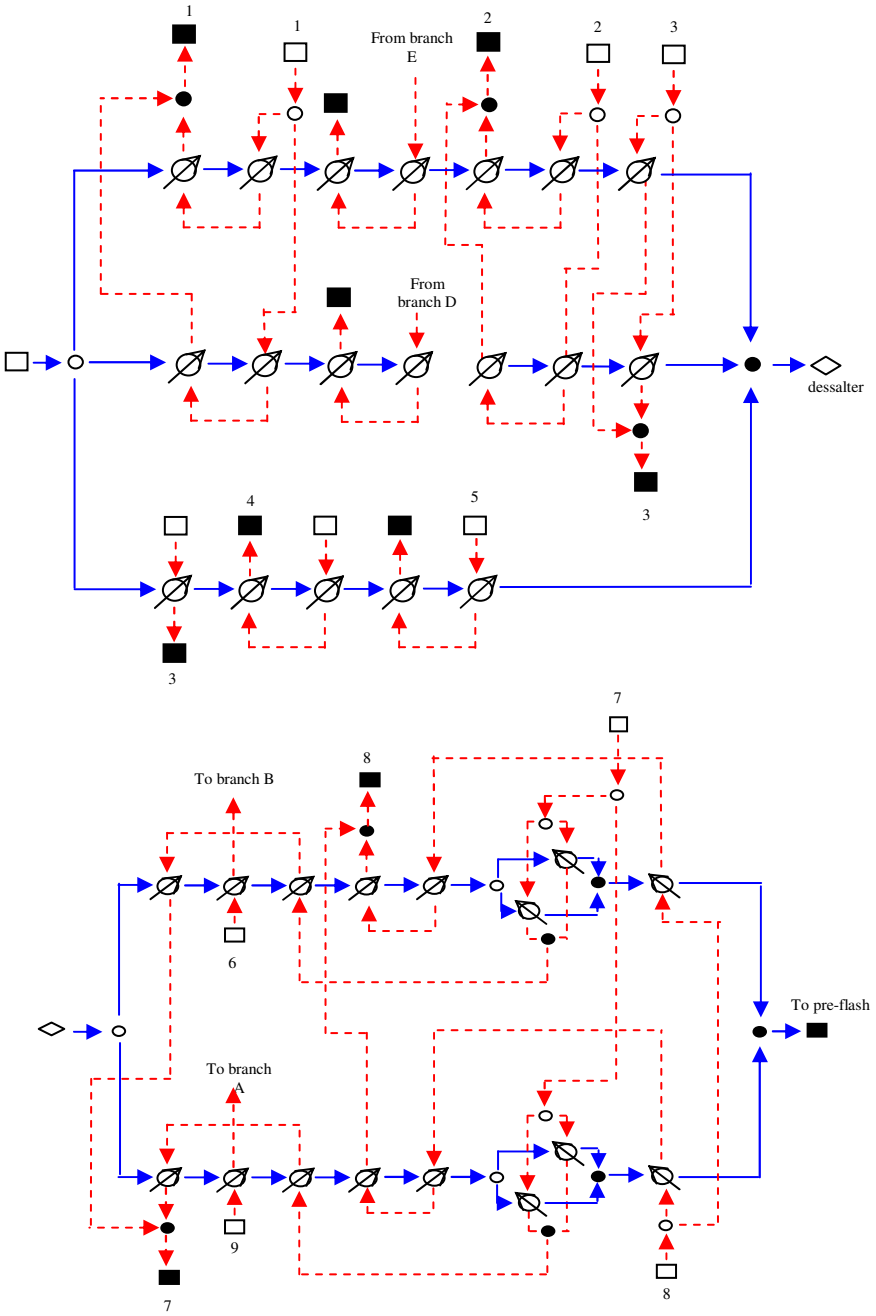


Figure 2 - Heat exchanger network flowsheet upstream and downstream the desalter - Cold streams: continuous lines - Hot streams: dashed lines - Supply nodes: white squares - Demand nodes: black squares - Splitters: white circles - Mixers: black circles

Table 1 – Optimized splitters on analyzed cases

Cases		**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Upstream the desalter	**Hot**	Yes	-	Yes	-	Yes	-	Yes	Yes	Yes
	Cold	Yes	Yes	Yes	Yes	Yes	-	-	Yes	-
Downstream the desalter	**Hot**	Yes	-	-	Yes	-	Yes	Yes	Yes	Yes
	Cold	Yes	Yes	Yes	Yes	-	Yes	-	-	Yes

Table 2 - Network final temperatures (ºC)

Reference	**Base**	**1**	**2**	**3**	**4**
273.94	265.15	266.22	265.84	265.85	266.22
5	**6**	**7**	**8**	**9**	
265.46	265.92	265.80	266.09	265.93	

Table 3 – Reduction of fouling costs (MMUS$/year)

Reference	**Base**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
-	-	0.17	0.11	0.11	0.17	0.05	0.12	0.10	0.15	0.12

The temperature gain ranges from 3.43% to 12.12% in relation to the total fuel costs due to fouling, evaluated by $(T_{reference} - T_{case}) / (T_{reference} - T_{base})$. The best results involve the optimization of all splitters (case 1), however similar temperature increases are obtained by simpler configurations (case 4 and 8). In terms of savings, it can be seen that even a 1 °C reduction, such as obtained, represents a great economy on operational costs of refinery due to fouling.

3.2. Feasibility constraints

The constraints discussed in the previous section were introduced in the optimization procedure by the insertion of penalty factors. The control loop scheme employed was case 1 (all splitters optimization).

Temperature constraints are the same in all cases: (i) Product streams: the limits were calculated based on a 20% upper variation on the heat transfer rate of the cooler needed to storage of products; (ii) Pumparounds streams: the limits appeared based on a variation on the heat load of the heat exchangers (upper and lower bounds). The constraints on the heat exchanger velocities were analyzed by several criteria, presented in Table 4.

Table 4 – Constraints in the heat exchanger velocities

Case	**Lower limit**	**Upper limit**
10	None	None
11	0.9 m/s in all heat exchangers	2.4 m/s in all heat exchangers
12	- 20% variation of base case	+ 20% variation of base case
13	Smallest lower limit of 11 and 12	Highest upper limit of 11 and 12
14	0.9 m/s in all heat exchangers	None
15	- 20% variation of base case	None
16	Smallest lower limit of 11 and 12	None

The final crude oil temperature in all cases is shown in Table 5. Attending the constraints implies just a small decrease of the possible temperature gain.

Table 5 – Optimized splitters on analyzed cases

Cases	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
Temperature (^oC)	266.22	266.21	266.19	266.09	266.17	266.20	266.17	266.17
Solution feasibility	-	Yes	No	Yes	Yes	No	Yes	Yes

It is important to mention that none of the streams had its temperatures violated with the optimization. Case 10 results were very similar to the optimizations without constraints once the optimized case originally was already inside the limits. Velocity constraints based on minimum velocity of 0.9 m/s and maximum velocity of 2.4 m/s (case 11) were not attended since a feasible case could not be found. However, it is important to mention that all heat exchangers of the refinery analyzed in operational conditions already violated those velocity bounds. That is the reason why it was proposed the establishment of a 20% velocity variation in relation to the base case.

4. Conclusion

Results show that optimization of hot and cold streams splitters can have a great impact on operational costs of refineries and should be further investigated. The careful analysis of operating aspects allowed the identification of realistic solutions for the fouling management in multi-branches preheat trains.

Acknowledgments

All authors thank Petróleo Brasileiro SA – PETROBRAS for the access to the real network example. E.M. Queiroz and F.L.P. Pessoa are grateful to CNPq for the research scholarship and A. L. H. Costa thanks Prociência/UERJ.

References

E. M. Ishiyama, W. R. Paterson, D. I. Wilson, 2008, Thermo-hydraulic channelling in parallel heat exchangers subject to fouling, Chemical Engineering Science, 63, 3400-3410.

J.H. Lavaja, M.J. Bagajewicz, 2004, On a new MILP model for the planning of heat-exchanger network cleaning, *Ind. Eng. Chem. Res.*,43, 3924-3938.

F. S. Liporace, S. G. Oliveira, 2005, Real time fouling diagnosis and heat exchanger performance, in: Proceedings of 6th International Conference on Heat Exchanger Fouling and Cleaning – Challenges and Opportunities, Kloster Irsee, Germany, 267-277.

L.O. Oliveira Filho, F. S. Liporace, E. M. Queiroz, A. L. H. Costa, 2007, Investigation of an alternative operationg procedure for fouling management in refinery crude preheat trains, *Applied Thermal Engineering*, submitted for publication, 2007.

F. Smaïli, V.S. Vassiliadis, D.I. Wilson., 2001, Mitigation of fouling in refinery heat exchanger networks by optimal management of cleaning, *Energy & Fuels*. 15, 1038-1056.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
©2009 Elsevier B.V. All rights reserved.

Optimal Sensor Location for Chemical Process Accounting the Best Control Configuration

David Zumoffen[a,b] and Marta Basualdo[a,b]

[a]*Grupo de Informática Aplicada a Ingeniería de Procesos (GIAIP), CIFASIS-CONICET-UNR. 27 de Febrero 210 bis, S2000EZP Rosario, Argentina. E-mail address: zumoffen, basualdo@cifasis-conicet.gov.ar*
[b]*Universidad Tecnológica Nacional (UTN), FRRO. Rosario, Argentina.*

Abstract

In this work a new methodology for solving simultaneously the problems of optimal sensor location and control structure selection for large scale chemical processes is presented. Here, it is considered the need of guaranteing the best plant-wide control structure before answering about which is the best sensor net able for achieving that objective. In this work it is demonstrated the importance of answer both questions as an integrated problem because of the strong impact in the initial investment and the future controlled process performance. Most of the previous works in this area analyze these problems as separated subjects. Here, genetic algorithms (GA) are used because they represent a valuable tool for support the decisions about the sensor placement, possible pairing of input output variables among a great number of combinations, since the interaction effect point of view. It allows to avoid the expert knowledge as decision criteria for pairing selection. The preliminary study is done on a simplified plant model obtained by subspace identification techniques (4sid). The final testing is performed on the rigorous dynamic model with the obtained plant-wide structure where the controllers tuning is performed through the internal model control (IMC) theory. The well-known case of the Tennessee Eastman (TE) benchmark is adopted for testing the methodology described here and compared with other strategies.

Keywords: sensor location, genetic algorithm, plant-wide control, RGA.

1. Introduction

In this work a new methodology for optimal sensor location accounting the control structure explicitly for large scale chemical processes is presented. It was found that from the optimal sensor location point of view the problem to be solved considers the sensor network design taking into account investment costs and observability topics (Musulin et al., 2005; Kadu et al., 2008; Singh and Hahn, 2005). Typical tools such as Kalman filtering, GA and pareto graphics were used on those papers. All of these works considered an open loop process or an existing controlled process. A similar situation occurs in the control structure selection area, focused in process interaction measures without considering any integration with the optimal sensor location problem. As can be seen, most of the previous works did not analyze this problem simultaneously, which is one of the new results introduced here. The TE benchmark control problem with the requirements imposed by Downs and Vogel (1993) is used for testing the proposed methodology. The steps for its implementation consist on: firstly performs the process stabilization which is carried out by three level control loops. Secondly, the subspace identification techniques (4sid) are applied for obtaining a linear

state-space model accounting those more critical requirements such as production rate, G/H ratio and B composition in purge. The problem to solve is the proper measurements selection with the minimum number of sensors located on optimal positions from the control point of view. In this context, it is proposed to find a control structure with smaller interaction chosen among 495 variables combination. Usually, the literature suggests heuristic criteria in order to reduce the number of combinations. The objective here is to develop a systematic approach that rigourously drives to the plant-wide control structure without heuristic considerations. Since the system is non square, the main ideas given by Chang and Yu (1990) were accounted, even though they could applied their technique only on plants of low dimensionality. In this paper the extension of Chang and Yu (1990) methodology is done with the help of GA. Finally, the IMC theory in steady-state for full controller design and the previously identified linear model support the decision of the outputs selection (to make the process square). As a result an interesting improvement can be achieved for both, control performance and hardware requirements, compared with the results given by McAvoy and Ye (1994). The most prominent simulation results are included according to some of the tests imposed for the TE. In addition, as future work is considered to analyze the advantages obtained when the expert system for diagnosis and fault detection for complex and large chemical process (Zumoffen and Basualdo, 2008a,b) are developed on the bases of the approach proposed here.

2. Case study: Tennessee Eastman process

Essentially, the TE plant generates two products from four reactants and has five principal units of operation: the reactor, the product condenser, a vapor-liquid separator, a compressor and a product stripper. The stripper underflow contains the key components G and H. In this work, it is assumed that the plant is operating under conditions considered as *base case*, that is G/H mass ratio of 50/50 and a production rate of 7038 kgG/h and 7038 kgH/h. The plant dynamic model used here was taken from Ricker's control department of the Washington University (`http://depts.washington.edu/control`).

3. Initial considerations

The TE open loop response is unstable, so, the first step is to stabilize the plant through three level controllers at the reactor, separator and the stripper units respectively. Then, the remaining variables to be considered are: 9 inputs and 12 outputs which are detailed at Table 1. In this work the nomenclature used is the same that the suggested one by Downs and Vogel (1993). The system identification (SI) procedure is performed to obtain a linear model in discrete state-space format. The SI begins with the plant inputs excitation using an uniformly distributed random magnitude between $\pm$ 2 % around the base case operation point and a period of 100 hr.. The recorded data of this process are collected using a sampling time of 5 min..

4. Optimal sensor location using GA

According to the TE dimensionality the linear model has $m = 12$ potential measurements and $n = 8$ potential manipulated variables because $XMV(12)$ has no impact on any of the considered output. Therefore, for detecting the less interactive pairing, the development given by Chang and Yu (1990) is accounted. They proposed a pairing selection supported by RGA for non square systems. It involves consider the plant model as $G = [G_s, G_r]^T$ displayed in the Fig. 1 of $m \times n$ dimension, which is partitioned into G_s of $n \times n$ and a G_r

Table 1: Input-output variables and base case values

Inputs	Variable	Base values	Outputs	Variable	Base values
D feed	$XMV(1)$	63.053 %	Recycle flow	$XME(5)$	26.902 kscmh
A feed	$XMV(3)$	24.644 %	Reactor feed	$XME(6)$	42.339 kscmh
A and C feed	$XMV(4)$	61.302 %	Reactor pressure.	$XME(7)$	2705.0 kPa gauge
Comp. rec. valve	$XMV(5)$	22.210 %	Reactor temp.	$XME(9)$	120.40 ^{o}C
Purge valve	$XMV(6)$	40.064 %	Separator temp.	$XME(11)$	80.109 ^{o}C
Stripper steam valve	$XMV(9)$	47.446 %	Separator pressure	$XME(13)$	2633.7 kPa gauge
RCWO t. set point	$XME(21)sp$	94.599 ^{o}C	Stripper pressure	$XME(16)$	3102.2 kPa gauge
CCW flow	$XMV(11)$	18.114 %	Production rate	$XME(17)$	22.949 m^3/h
Agitator speed	$XMV(12)$	250 rpm	Stripper temp.	$XME(18)$	65.731 ^{o}C
			Compressor work	$XME(20)$	341.43 kW
			B comp. in purge	$XME(30)$	13.823 mol%
			G/H ratio	$XME_{G/H}$	1.2257

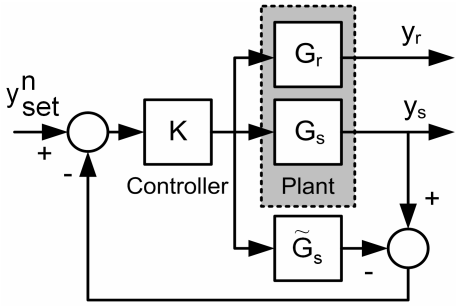


Figure 1: IMC for non-square plant

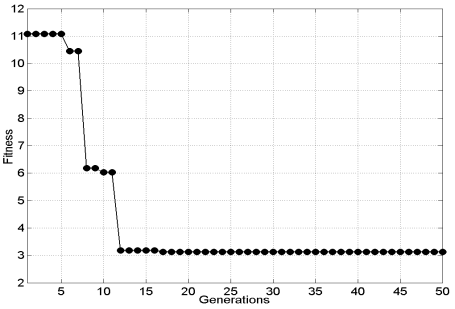


Figure 2: Fitness function

of $(m-n)\times n$. Then,

$$y = \begin{bmatrix} y_s \\ y_r \end{bmatrix} = \begin{bmatrix} G_s \\ G_r \end{bmatrix} u \tag{1}$$

where y_s and y_r are $n\times 1$ and $(m-n)\times 1$ respectively. Considering the IMC theory, a square controller can be designed as $K = \tilde{G}_s^{-1}$, so $u = G_s^{-1} y_{set}$, in addition, perfect plant-model $G_s = \tilde{G}_s$ is assumed and output disturbances are neglected. The closed loop error e is defined as,

$$e = y_{set}^m - y = y_{set}^m - GG_s^{-1} y_{set}^n \tag{2}$$

where $y_{set}^m = [y_{set}^n, y_{set}^{m-n}]^T$, $y_{set}^n = [y_{set}^1, \ldots, y_{set}^n]^T$. Considering $y_{set}^{m-n} = [0, \ldots, 0]^T$ then,

$$e = \left[\begin{bmatrix} I_{n\times n} \\ 0_{(m-n)\times n} \end{bmatrix} - GG_s^{-1} \right] y_{set}^n = A y_{set}^n, \quad \text{with} \quad A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \tag{3}$$

being e $m\times 1$ and considering all the system outputs. Then, eq. 3 can be calculated for the rest of set point changes so as $e(i) = A y_{set}^n(i)$. Here $e(i)$ makes reference to the error on all variables when an unitary set point change in the variable i has been occurred, being $y_{set}^n(i)$ a vector with unit value in the i entry and zero for the rest of the elements. The SSE index computes the sum of the square errors for the overall plant variables accounting each set point variation, that is:

$$SSE = \sum_{i=1}^{n} ||e(i)||_2^2 = \sum_{i=1}^{n} ||A y_{set}^n(i)||_2^2 = \mathrm{tr}\left(A^T A\right) \tag{4}$$

The control objective is to minimize the SSE. Clearly a particular selection of G_s affects the *SSE* performance. Additionally, according to the number of involved variables , the combinatorial problem presents $m!/(n!(m-n)!) = 495$ possible solutions. The optimal solution is found by using the genetic algorithm (GA) approach (Chipperfield et al., 1994). It is a stochastic global search method that mimics the metaphor of natural biological evolution. In this case, the chromosome length can be associated to the possible sensor set dimension and the sensor network structure with the chromosome genes. Thus, $I_j = [g_1, g_2, \ldots, g_{N_c}]$ is the chromosome, where $j = 1, \ldots, N_i$ and N_i the initial population dimension. The genes parameters g_i, with $i = 1, \ldots, N_c$, belong to a binary code $g_i = \{1, 0\}$ indicating the presence or not of a sensor in the i location, and N_c the potential sensor network length. The problem to be solved can be represented as

$$\min_{I_j} [SSE(I_j)] \quad \text{subject to: } ||I_j|| = n \tag{5}$$

where the condition $||I_j|| = n$ guarantees the square selection of the sub-process G_s. Solving eq. 5 by GA the optimal sensor location, I_{op}, can be obtained. In the Fig. 2 can be observed the fitness criteria profile for the best individual selected in less than 50 generations. The sensors net for the optimal solution I_{op} is displayed at Table 2, in addition a suboptimal I_{sop} one is also included . The zero values in these chromosomes represent no sensors in these locations, and the opposite for the unitary values. The optimal solution generates the following optimal sensor network structure: *recycle flow, separator temperature, separator pressure, production rate, stripper temperature, compressor work, B composition in purge and G/H ratio*.

5. Control structure

In this section the control structure is defined by using RGA analysis (Bristol, 1966) carried out for the optimal solution. Computing the RGA from Table 2 results can be obtained the relative gain matrices displayed at Table 3 for optimal Λ_{op} case, where $\Lambda_{op} = G_s(I_{op}) \otimes (G_s(I_{op})^{-1})^T$. The best RGA presents lower interaction between the process variables than the suboptimal one. The final control topology is displayed at Fig. 3.

6. Results

The optimal sensor location and control structure selection presented in previous sections are compared and analyzed. By one side it is evaluated if the process operation conditions stated by Downs and Vogel (1993) are met. By other side both control performance and hardware requirements are compared with the classical control structure presented in previous work (McAvoy and Ye, 1994). The TE process presents four possible set point modifications (production rate, product mix, reactor pressure and B purge composition) from its base case value and different possibilities to analyze disturbance effects (IDV1 to IDV20). Due to space limitation only some results are displayed here. The Fig. 4 considers the IDV1 disturbance effects on both control structures and displays the reactor pressure profile. The simulations are done once the controllers tuning, based on the IMC theory (Rivera,

Table 2: Different solutions from GA

Individual	Sensor network												Fitness
I_{op}	1	0	0	0	1	1	0	1	1	1	1	1	3.1280
I_{sop}	0	0	1	1	1	0	0	1	1	1	1	1	11.0751

Table 3: Optimal solution RGA

I_{op}	$XMV(1)$	$XMV(3)$	$XMV(4)$	$XMV(5)$	$XMV(6)$	$XMV(9)$	$XME(21)sp$	$XMV(11)$
$XME(5)$	0	0	0	0	0.17	0.02	1.38	0.331
$XME(11)$	0	0	0.10	0.47	0.31	0	1.45	0.61
$XME(13)$	0.19	1.07	0.11	0.19	0	0	0	0.18
$XME(17)$	0	0.10	1.07	0.03	0.01	0.03	0	0
$XME(18)$	0.06	0.35	0	0	0	1.92	0	0
$XME(20)$	0.04	0.15	0	1.17	0.13	0.03	0	0.02
$XME(30)$	0.08	0.03	0	0	0.68	0.02	0.05	0.17
$XME_{G/H}$	1.05	0.07	0	0	0	0	0	0.01

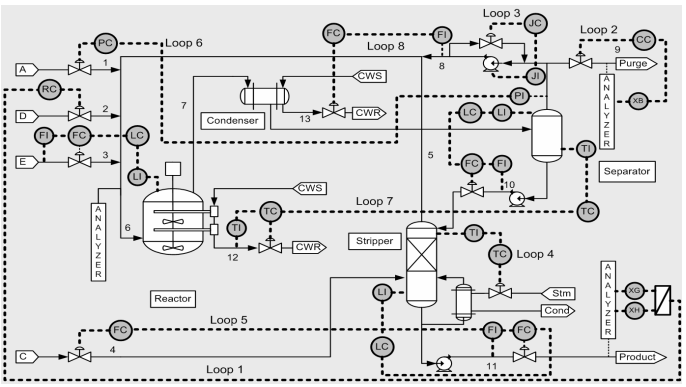


Figure 3: Final control structure

2007) implemented with the dynamic linear model identified previously. All the controllers implemented in this work are discrete PI versions obtained from the continuos case. The control policy obtained from the optimal sensor network clearly shows the good regulatory properties in this case. By other hand, the control structure for the suboptimal sensor network presents problems for rejecting the disturbance. An incorrect pressure control loop produces the plant shut down. The Fig. 5 displays the reactor pressure time evolutions when a production rate change occurs. The suboptimal control structure cannot maintain the correct operation of the process and the shut down (labelled with sd) takes place due to reactor pressure violation. The alternative control structure proposed here fulfills with the specifications showing a similar performance as that presented in McAvoy and Ye (1994). The advantage of the approach presented here is mainly given by a lower number of sensors needed as can be observed at Table 4.

7. Conclusions

In this work has been presented an alternative optimal sensor location method integrated to control structure design for large-scale processes. The final result displays a systematic approach that allows arriving to an efficient decentralized control structure, implemented with the minimum number of measurements. Additionally, this methodology is practically independent of any heuristic information. The simulation results obtained here has been compared with previous works showing clear advantages in performance and investments. As future work, it will be analyzed which is the potentiality for improving expert systems design for fault diagnosis in large chemical plants.

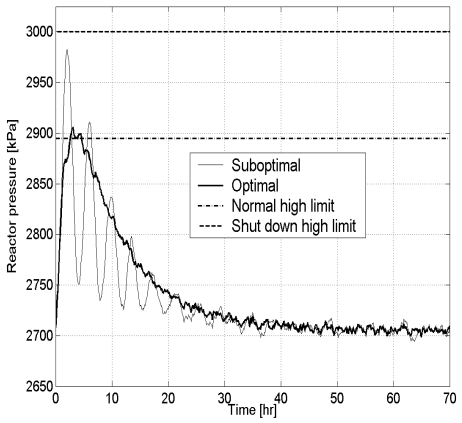


Figure 4: IDV1 disturbance

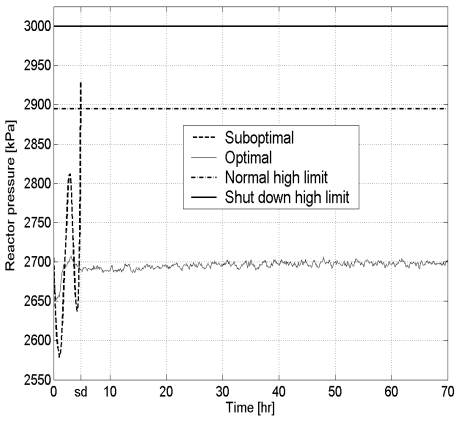


Figure 5: Production rate change

Table 4: Hardware requirements

	McAvoy and Ye (1994)	Ricker (1996)	Larsson et al. (2001)	Proposed here
Measurements	22	16	22	15
Control loops	22	19	17	15

References

E. Bristol, 1966. On a new measure of interaction for multivariable process control. *IEEE Transactions on Automatic Control*.

J. Chang and C. Yu, 1990. The relative gain for non-square multivariable systems. *Chemical Engineering Science*. *45,* 1309-1323.

A. Chipperfield, P. Fleming, H. Pohlheim and C. Fonseca, 1994. Genetic algorithm toolbox. For use with Matlab. *University of Sheffield. Department of automatic control and systems engineering*. `http://www.shef.ac.uk/acse/research/ecrg`.

J. Downs and E. Vogel, 1993. A plant-wide industrial process control problem. *Computers Chem. Engng.. 17,* 245-255.

S. Kadu, M. Bhushan and R. Gudi, 2008. Optimal sensor network design for multirate system. *Journal of Process Control*. *18,* 594-609.

T. Larsson, K. Hestetun, E. Hovland and S. Skogestad, 2001. Self-optimizing control of a large-scale plant: The Tennessee Eastman process. *Ind. Eng. Chem. Res.. 40,* 4889–4901.

T. McAvoy and N. Ye, 1994. Base control for the Tennessee Eastman process. *Computers Chem. Engng.. 18,* 383-413.

E. Musulin, C. Benqlilou, M. Bagajewicz and L. Puigjaner, 2005. Instrumentation design based on optimal Kalman filtering. *Journal of Process Control*. *15,* 629-638.

N. Ricker, 1996. Decentralized control of the tennessee eastman challenge process. *Journal of Process Control*. *6,* 205-221.

D. Rivera, 2007. Una metodología para la identificacion integrada con el diseño de controladores IMC-PID. *RIAI (Rev. Iberoamericana de Autom. e Inf. Ind.). 4,* 5-18.

A. Singh and J. Hahn, 2005. Determining optimal sensor location for state and parameter estimation for stable nonlinear systems. *Ind. Eng. Chem. Res.. 24,* 5645-5659.

D. Zumoffen and M. Basualdo, 2008a. Improvements in fault tolerance characteristics for large chemical plants. Part I: Waste water treatment plant with decentralized control. *Ind. Eng. Chem. Res.. 47,* 5464-5481.

D. Zumoffen and M. Basualdo, 2008b. Improvements in fault tolerance characteristics for large chemical plants. Part II:Pulp mill process with model predictive control. *Ind. Eng. Chem. Res.. 47,* 5482-5500.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
©2009 Elsevier B.V. All rights reserved.

A new systematic approach to find plantwide control structures

Gonzalo Molina[a,c], David Zumoffen[a,b] and Marta Basualdo[a,b]

[a]*GIAIP – Centro Franco-Argentino de Ciencias de la Información y de Sistemas (CIFASIS-CONICET-UNR-UPCAM III). 27 de Febrero 210 bis, S2000EZP Rosario, Argentina*
[b]*Universidad Tecnológica Nacional (UTN) – FRRo. Rosario, Argentina.*
[c]*FCEIyA-Universidad Nacional de Rosario (UNR). Rosario, Argentina.*

Abstract

A new systematic approach for addressing the problem of the plant-wide control structure, supported by steady-state information is presented. It is tested in a reactor/separator with recycle plant. Several authors have presented different control alternatives for this kind of plant. However, most of them mainly focused on the optimum energy consumption or in the regulator problem only. In both cases the decision of the final control structure was adopted based on several heuristic concepts. In this work both objectives are considered avoiding any heuristic concepts. The approach consists meanly on three steps: optimization, stabilization, and final pairing between manipulated and controlled variables focused on rejecting the most critical disturbances. In addition, the reactor design is subject to ensure quality product specification, cost investment together with the overall plant controllability. Finally, the dynamic model of the plant is used for testing the potentiality of the proposed control structure. Hence, a multivariable tuning procedure, based on internal model control (IMC), is performed for evaluating the dynamic closed loop responses under the worst disturbances scenarios.

Keywords: plant-wide control, optimum energy consumption, disturbance rejection, controllability.

1. Introduction

A large number of contributions on plant-wide control structure synthesis can be classified into the process-based experience (engineering judgment) and mathematical-based criteria. Among the most notable contributors on the first category are McAvoy (1998) and Luyben and co-workers (Belanger and Luyben, 1996, 1997). This methodology is systematic in nature and addresses many of the major issues involved in the plant-wide control problem, such as the effects of recycles and energy integration. A nine step approach is developed by Luyben and co-workers based on the process experience of the group. The main result from this philosophy is an alternative way to achieve a decentralized plant-wide control structure. The second category relies on a rigorous mathematical framework of dynamic theory, constrained optimization and systems analysis. Skogestad and co-workers (Skogestad (2000), Larsson (2000), Alstad and Skogestad (2007)) obtained interesting results using these tools. They focused on the selection of self-optimizing control variables that give the smallest loss in profit. The mathematical approaches given by these authors, although rigorous, are often very difficult to formulate for large scale systems, in many cases they turn computationally very intensive. In addition, the model assumptions can affect the convergence and solution. Hence, in this work a kind of consensus between both categories mentioned above is presented. In order to achieve this purpose some useful modifications are proposed to generate a new systematic approach able to deal with the most critical problems such as size, complexity, and multiple objectives commonly found in the majority of chemical plants.
A representative example in the plant-wide control community is studied, which is very useful for obtaining deep knowledge and comparing with other previous results. In the next section a

brief description of the study case is given. In section 3 the complete approach is detailed step by step. Particularly in the third step the extension of the null space criteria is introduced. It allows to perform the variable pairings selection with less interaction effects. In section 4 the closed loop dynamic responses are shown. An acceptable controllers tuning are given for achieving a good regulation under the most critical disturbances. At this point, the reactor design can be also evaluated and modified if it is necessary in case that the product quality,cost investment and controllability aspects do not meet specification.

2. Case Study Description

The studied plant consists on a reactor and a distillation column interconnected whose data has been taken from Wu. and Yu (1996). It is assumed that an irreversible first order reaction $\mathbb{A} \rightarrow \mathbb{B}$ occurs in the continuous stirred tank reactor (CSTR) being the reaction rate k a function of the temperature given by Arrhenius expression $k = k_0.e^{-E/R.T_R}$.
This is an exothermic reaction. The reaction rate is considered constant ($k \approx 0.340843$ hr^{-1}). Some of the reactant $\mathbb{A}$ is consumed in the CSTR and the effluent of the reactor, a mixture of $\mathbb{A}$ and $\mathbb{B}$, is fed into the tray 12 of the 20-tray distillation column. The product $\mathbb{B}$ is obtained from the bottom of the column, while the distillate of the column is recycled to the CSTR. Constant relative volatility $\alpha = 2$ between $\mathbb{A}$ and $\mathbb{B}$ is assumed in the column model. The plant is shown in Fig. 1. The final control structure is also included in this figure. It will be explained in the following sections.

3. The proposed systematic approach

In this section the three main steps of the approach are described.

3.1. Process optimization

Table 1: Case Study Parameters

Parameters, Variables	Base Case	Optimum	Units
Reactor			
Feed Flow Rate (F_0)	460	460	(lb mol/hr)
Feed Flow Composition (z_0)	0.9	0.9	
Recycle Flow Rate (D)	500.15	651.89	(lb mol/hr)
Residence Time	2.5	2.7	(hr)
Column			
Column Feed Flow Rate(F)	960.15	1111.89	(lb mol/hr)
Column Feed Composition (z)	0.5	0.475	(lb mol/hr)
Reflux Flow Rate (L)	1100	862.28	(lb mol/hr)
Distillate Flow Rate (D)	500.15	651.89	(lb mol/hr)
Product Flow Rate (B)	460	460	(lb mol)
Product Composition (x_B)	0.0105	0.0105	
Reboiler Hold Up (N_{reb})	275	275	(lb mol)
Condenser Hold Up (N_{cond})	185	185	(lb mol)
Vapor Boil-Up (VB)	1600	1514	(lb mol/hr)

In this stage a steady state optimization is done in order to minimize the vapor boil up VB of the plant, which is considered the objective function J (Larsson, 2000). Hence, the problem to be solved is $\min_{x,u} J(x,u,d)$ subject to the constraints $f(x,u,d) = 0$, $g(x,u,d) \leq 0$, $y = f_y(x,u,d)$; where $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$ and $y \in \mathbb{R}^{n_y}$ are the states, inputs and measurements, respectively; f is the set of equality constraints corresponding to the model equation; g is the set of inequality constraints that limit the operation and n_x, n_y and n_u are the number of states, measurements and manipulated variables respectively.
This optimization problem is solved by using a steady-state model created in GAMS (software for modeling and optimization). The assumed constraints are: the product quality $x_B \leq 0.0105$, the reactor hold-up $V_R \leq 2400$, and flow rates can be at least two times of the initial value (see Wu. and Yu, 1996). In table 1 can be seen both, the obtained optimal magnitudes and base case values where the disturbances z_0 and F_0 are kept constant. From this step it is straight forward the variables which need to be controlled (Skogestad, 2000)are:

- levels without steady state effect (N_{cond}; N_{reb})

- variables that are active constrain in the optimum (x_B, V_R)

3.2. Plant stabilization

The second step proposes the plant stabilization, which drives to a typical inventory control (McAvoy, 1998). Then most suitable paring between the levels as controlled variables with the manipulated variable is performed looking for the less interactive ones supported by the RGA. Since in all cases they are unstable controlled variables (pure integrator function), the RGA calculations were done considering derivative output with respect to time.

In this case only two possible level control structures are possible. The pairings are: the reboiler level N_{reb} with B and bottom level V_R with F. The control of N_{cond} can be performed with D or with L. In this study, the first one was chosen because changes in D produce less impacts on the reactor level because of the relative size between N_{cond} and V_R. The final inventory control structure for this process is shown in Fig. 1 with grey background.

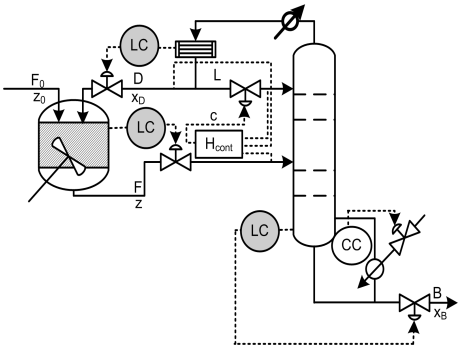


Figure 1: Final Control Structure

Table 2: Measurements Combinations

N°	Variables	H	RGA
1	$[D,F,L]$	-0.1920 -0.2030 0.7359	0.997
2	$[D,F,z]$	-0.0004 -0.0000 1.0000	0.995
4	$[D,L,z]$	-0.0004 0.0001 1.0000	0.995
7	$[F,L,z]$	-0.0009 0.0002 1.0000	0.995
11	$[B,L,z]$	0.0008 0.0003 1.0000	0.995

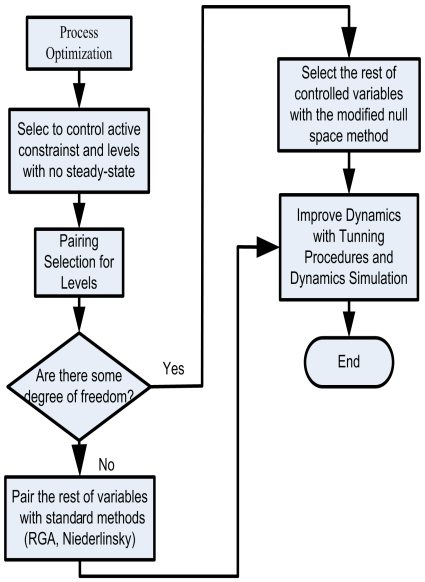


Figure 2: Proposed Procedure

3.3. Pairing Variables Selection for Regulatory Control and Economic Optimum

In this stage the main objective is to keep the plant near the optimal steady-state operation without re-optimizing the process when disturbances affect the plant. This is done based on the main concepts of self-optimizing control and null space method introduced by Alstad and Skogestad (2007). It consists on finding the controlled variables c as a linear combination of measurements y. The requirement is to have at least as many measurements as the unconstrained degrees of freedom, including disturbances. This technique was successfully used for controlling a Petlyuk distillation through a choice of a proper combination of temperature measurements. However, for more complex cases, such as the plant studied here, the selected combination of variables was strongly interacted. Because of this, it is proposed performing the selection focused on those combinations that provide a best decentralized control structure supported by RGA calculations.

The linear relationship between c and y is given by $c = Hy$, where H calculation is done through the optimal sensitivity matrix F whose elements f_{ij} are given by $\frac{\partial y_i^{opt}}{\partial d_j}$, such that:

$$HF = 0 \tag{1}$$

Then, $H \in \mathscr{N}(F^T)$ where $\mathscr{N}(F^T)$ is the null space of F^T. The size of F is $n_y \times n_d$ and its rank is n_d since $n_d \leq n_y$. In addition, the size of H is $n_u \times n_y$ and its rank is n_u. It can be proved that to ensure the existence of H n_y must be chosen such that $n_y \geq n_u + n_d$. In this work 2 disturbances and 5 manipulated variables were considered. Three degrees of freedom were consumed in the previous step for stabilizing the plant and one must be consumed for keeping the product quality under specification (from step 1). Therefore, the minimum n_y is 3 ($n_u = 1$ and $n_d = 2$), which is the number of outputs to be combined for the modified null space application. In other words, in this step must be chosen two loops, one for the quality control and the other for the combined measurements responsable of avoiding the disturbance effects. At this stage, the RGA is useful to detect which combination of measurements is more suitable for control with a decentralized structure (see Bristol, 1966). In this example the paring of the remaining two manipulated variables must be done, so only one element of the RGA is necessary to be calculated. In table 2 the RGA values, corresponding to the pairing $L - c$ are shown, where c is the controlled variable obtained from the lineal combination mentioned above. Then, five combinations are good candidates for implementing decentralized control structure. The chosen combination in this work was the first one for two reasons:

- the best RGA was obtained for this combination,
- all the variables that are included in this combination are easy to be measured.

The resulting control structure is shown in Fig. 1. The block H_{cont} represents the calculation of the linear combination of the variables and the PI controller that manipulates L. In Fig. 2 a flow chart is shown with all the steps included in the approach to obtain the final control structure.

4. Results

In this section the control structure developed above in steady state is tested on the dynamic model. Here, the plant under the worst disturbances scenario is used to tune the controllers guaranteeing the process specifications.

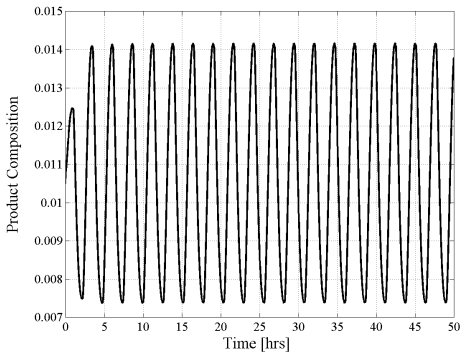


Figure 3: LATV with disturbance F_0

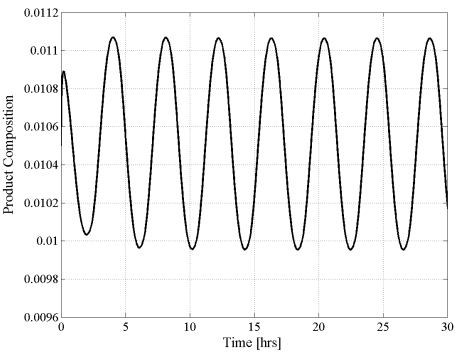


Figure 4: LATV with disturbance z_0

4.1. The Worst Case Disturbances Scenario

A worst case of dynamic disturbances was obtained by using the test described by Belanger and Luyben (1996) because of its simplicity and robustness. It consists in an Auto Tune Variation Test whose cycle tends to maximize the system's output. It is carried out connecting the process output to a device that approximates the derivative of the signal. It results in a signal that passes through zero when the process output reaches an extreme value. This signal is sent to a relay that switches the load variable when its input crosses zero magnitude. In this case the bottom product composition is one of the most important process output.

This test was done accounting the most typical disturbances: F_0 and z_0. The dynamic effects of those perturbations are presented in Fig. 3 and Fig. 4 respectively. Hence, it is possible to

evaluate which disturbances are most problematic because of their effects on the quality product composition. For the first one the frequency regimen is established at 2.46 rad/hr. It will be most critical frequency for the feed flow rate with great impact on the product quality. For the feed composition z_0 the peak of frequency is 1.54 rad/hr. Then, the control system has to be designed for rejecting these disturbances entering at those critical frequencies assumed as the worst cases.

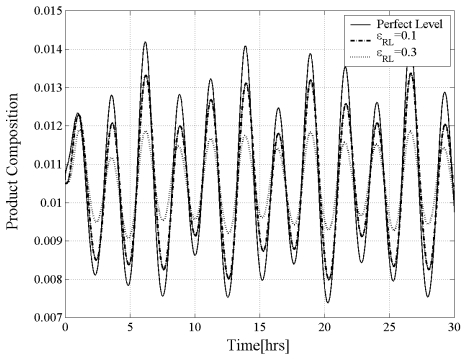


Figure 5: x_B Worst Case Response

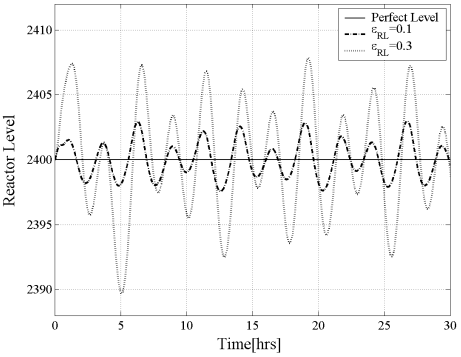
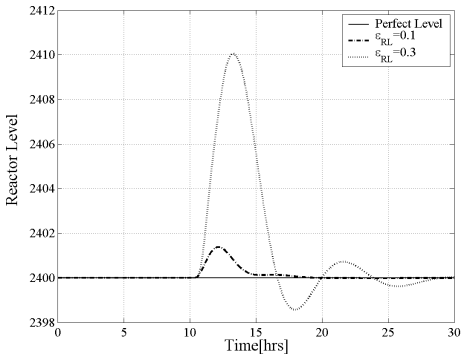


Figure 6: V_R Worst Case Response

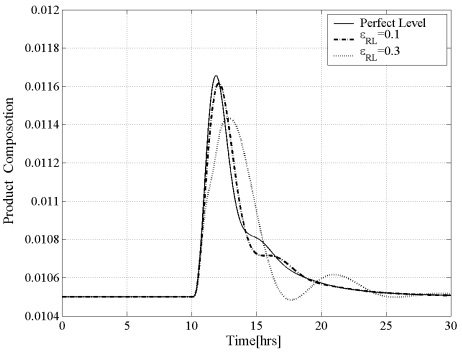
4.2. Reactor Inventory Control Tuning

In this stage a relaxed inventory control in the reactor, as that proposed in Belanger and Luyben (1997), is adopted. The objective is to reduce the product composition variability for the disturbances analyzed here. In this work, the control tuning was done using IMC because of its simplicity on adjusting only the filter parameter (see Rivera, 2007). In the case of the reactor level the parameter is named ε_{RL}. In Fig. 5 and 6 the simulations with both disturbances affecting the plant with different values of ε_{RL} are shown . Taking into account that the composition must be $0.009 \leq x_{AB} \leq 0.012$, a part of the product will be out of specifications if $\varepsilon_{RL} = 0.1$. Otherwise, for $\varepsilon_{RL} = 0.3$ the product composition is always within the limits of the quality requirements even though it requires a larger reactor size. Hence, $\varepsilon_{RL} = 0.3$ is adopted because it gives a good trade off between controllability and product quality though higher investment on the reactor size is necessary.

In addition, step changes in the feed flow rate of about $\Delta F_0 = \pm 10\%$ and feed composition such as $\Delta z_0 = \pm 10\%$ are evaluated.



(a) Reactor Level Response



(b) Composition Product

Figure 7: Responses to Step Change in the Feed Flow Rate

In Fig. 7 the reactor level and product composition responses are shown when a step change in the feed flow rate is applied to the plant. A good regulatory behavior is noticed for the product

composition specially for the lower value of ε_{RL}. However, the opposite situation is found in the reactor level. Therefore, the same conclusion is derived from these last responses.

5. Conclusions

A new systematic methodology for synthesizing plantwide control structures has been presented. The method is based on rescuing the most useful ideas, given by the two main sources: the process-based experience and engineering judgment and a rigorous mathematical framework. This approach achieves a good decentralized structure, considering the most typical objectives for any plant, which are optimum energy management and good regulatory behavior. Even though this decision-making technique is based mainly on steady-state information, good results are achieved as it was demostrated through the simulations performed with the dynamic model. In addition, a systematic multivariable tuning method, thought to be enough robust for hadling well the most crucial disturbances, was very useful . Furthermore, this approach is helpful for deciding whether the equipment sizing is suitable for guaranteeing plant wide controllability. The satisfactory disturbance rejection was demonstrated on the integrated plant tested in the context of the worst cases. Future work will include results of the efficacy of this approach on a more complex flowsheet where a great number of possible combination for pairing variables are denoted. The solution of this large scale problem is planned to be performed with the help of data mining tools.

References

Alstad, V. and S. Skogestad (2007). Null space method for selecting optimal measurement combinations as controlled variables. *Industrial Engeenering Chemical Research* **46**, 846–853.

Belanger, P. W. and W. L. Luyben (1996). A new test for evaluation of the regulatory performance controlled variables. *Industrial Engeenering Chemical Research* **35**, 3447–3457.

Belanger, P. W. and W. L. Luyben (1997). Inventory control in processes with recycle. *Industrial Engeenering Chemical Research* **36**, 706–716.

Bristol, E. H. (1966). On a measure of interaction for multivaraible process control. *IEEE Trans. Autom. Control* **11**, 133–134.

Larsson, T. (2000). Studies On Plantwide Control. PhD thesis. Norwegain University of Science and Technology.

McAvoy, T.J. (1998). A methodology for screening level control structures in plantwide control systems. *Computers and Chemical Engineering* **22**(11), 1543–1552.

Rivera, D. E. (2007). Una metodología para la identificación integrada con el diseño de controladores imc-pid. *Revista Iberoamericana de Atutomática e Informática Industrial* **4**(4), 5–18.

Skogestad, S. (2000). Plantwide control: the search for self-optimizing control structure. *Journal of Process Control* **10**, 487–503.

Wu., K.L. and C.C. Yu (1996). Reactor / separator processes with recycles -1.candidate control structures for operability. *Comput. Chem. Eng* **20**, 1291–1316.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Control Strategy for a *Zymomonas mobilis* Bioreactor Used in Ethanol Production

Fabio C. Diehl and Jorge O. Trierweiler

Departament of Chemical Engineering - Federal University of Rio Grande do Sul, Rua Luiz Englert s/n, CEP: 900040-040 – Porto Alegre, RS, Brazil

Abstract

High oil prices, increasing focus on renewable carbohydrate-based feedstocks for fuels and chemicals have provided continuing stimulus for studies on *Zymomonas mobilis* bacteria. The ethanol production with this bacterium shows interesting nonlinear dynamic behaviors. In the continuous fermentation, there are equilibrium multiplicity, Hopf points and stability changes. Various models have been proposed in the literature (Daugulis *et al*., 1997; Jarsebski, 1992; Ghommidh *et al*., 1989; Jöbses *et al*., 1986) for to describe such phenomena. In special, the Jöbses' model has captured these behaviors and revealed an attractive operating range with high substrate to product yield. This article presents a multivariable control strategy for the continuous ethanol synthesis through *Z. mobilis*, in a high yield solution branch inside the steady-state multiplicity region.
Keywords: RPN methodology, MIMO control, equilibrium multiplicity, bifurcation.

1. Introduction

Zymomonas mobilis has attracted considerable interest over the past decades as a consequence of its unique metabolism and ability to rapidly and efficiently produce ethanol from simple sugars. However, despite its apparent advantages of higher yields and faster specific rates when compared to yeasts, no commercial scale fermentations currently exist which use *Z. mobilis* for the manufacture of fuel ethanol. In the literature, the bacteria *Zymomonas mobilis* has been proposed as a more promising microorganism than conventional yeast Saccharomyces cerevisiae for industrial production of ethanol (Rogers *et al*., 2007). Continuos glucose fermentation, with *Z. mobilis*, exhibit oscillatory behavior (Daugulis *et al*., 1997; Jarsebski, 1992; Bruce *et al*., 1991), and others nonlinear characteristics reported by Elnashaine *et al*. (2006), Maheca-Botero *et al*. (2006), e Garhayan e Elnashaine (2004). All these publications are based on the Jöbses's model, shortly describe in the section 2. In this paper is used the Jöbses's model to carry out the process analysis and studies. Based on this model, a suitable control strategy to ethanol fermentation by *Z. mobilis* is proposed and simulated.

2. Model Description and Analysis

Several models are reported in the literature to describe the *Z. mobilis* fermentation:

- Daugulis *et al*. (1997) – an unstructured and unsegregated model that proposes a macroscopic approximation, incorporating a dynamic specific growing rate (consider the recent process inhibition historic). This model was validated by McMellan *et al*. (1999) at low dilution rates.
- Jarsebski (1992) – consider product inhibitions and substrate limitations. It is an extension of Ghommidh *et al*. (1989) structured model, that divide the biomass in viable, unviable and dead cells.

- Jöbses *et al*. (1986) – an unsegregated-structured two-compartment representation. Divide the biomass in compartments containing specific groupings of macromolecules. It incorporates the product inhibition.

In this work was used the Jöbses' model to all analysis and projects. This model can describe several nonlinear behaviors and have a good acceptance. It is validated for low dilution rate. As a high yield operating branch appears in the steady-state multiplicity at high dilution rates, our contribution will assume that this extrapolation is acceptable. Here we are proposing a control strategy suggestion to maintain the system working at this more profitable operating region.

2.1. Jöbses's model

Since the Jöbses's model can predict a branch with higher ethanol production, which has been experimentally confirmed (at least for low dilution rates) by Elnashaie *et al.* (2006), was decided to analyze the control problem of a continuous bioreactor with the Jöbses *et al.* (1986) kinetic model, which is shortly described in this subsection.

The Jöbses's model consists of the following four differential equations:

$$\frac{dC_S}{dt} = -\left(\frac{C_S C_e}{Y_{Sx}\left(K_S + C_S\right)}\right) - m_S C_x + D\left(C_{So} - C_S\right) \tag{1}$$

$$\frac{dC_x}{dt} = \left(\frac{C_S C_e}{K_S + C_S}\right) + D\left(C_{xo} - C_x\right) \tag{2}$$

$$\frac{dC_e}{dt} = \left(k_1 - k_2 C_P + k_3 C_P{}^2\right)\left(\frac{C_S C_e}{K_S + C_S}\right) + D\left(C_{eo} - C_e\right) \tag{3}$$

$$\frac{dC_P}{dt} = \left(\frac{C_S C_e}{Y_{Px}\left(K_S + C_S\right)}\right) - m_P C_x + D\left(C_{Po} - C_P\right) \tag{4}$$

where C_S is the substrate concentration (glucose), C_x is the biomass (*Z. mobilis*), C_P is the product (ethanol), and C_e is an auxiliary variable used to lag the effect of the ethanol concentration in the kinetic model. The variables D and C_{So} represent the dilution rate and the feed substrate concentration, respectively. The C_e and the (k_1-k_2C_P+$k_3C_P{}^2$) terms make possible the model to depict the oscillatory behavior of the Hopf bifurcations. More details about the model origin can see found in Jöbses *et al*. (1986). The model parameters (K_S, k_1, k_2, k_3, m_S, m_P, Y_{Sx}, Y_{Px}) are shown in Jöbses *et al*. (1986).

2.2. Operational optimum

The three dimensional static/dynamic system characterization was performed using the Matcont toolbox, implemented on Matlab®, where the analyzed bifurcation parameters were D and C_{So}. The Figure 1(a) shows the results for C_P state.

The bifurcation diagram reveals triple multiplicity under some operational range. Two solutions groups are stable branches and between them, there is an unstable region. One of the stable equilibrium set present high ethanol production (region *A*) in comparison of the other stable region (region *C*). It is clear that to operate the reactor on the superior region is more profitable because the yield is bigger. The multiplicity unknowledge will lead to a trivial low yield operation. In the C_P codimension-2 analysis a Hopf bifurcation set is represented by points (•), at low dilution rates, in the lower yield field. The unstable solutions (region *B*) are linked to the superior stable branch for a saddle bifurcations line. This limit can see view in the Figure 1(b). Besides, this frontier divides the system in two partitions: the space *M*, where the multiplicity exists; and the

space *S*, where there is only a single solution. To reach the optimal profit operation, the process must be leaded in the space *M*, at the superior C_P region. As the productivity is directly dependent of *D*, the optimum operating point is close to the saddle point frontier formed by the intersection between the above stable operating range and the unstable one.

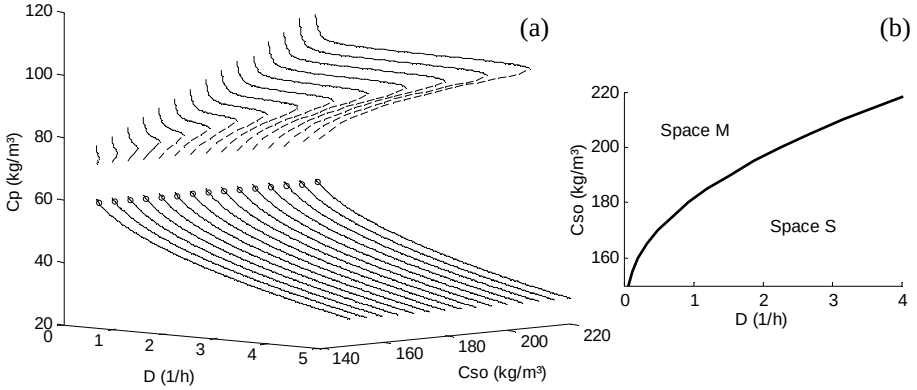


Figure 1 – Codimension-2 bifurcation (a); stability and multiplicity frontier (b).

3. Operational Controllability Analysis

Ethanol and substrate concentrations are intuitive process control variable while the natural manipulated variables are the *D* and C_{So}. It was considered the pairing $D \rightarrow C_P$ and $C_{So} \rightarrow C_S$. The operational controllability has been analyzed through the RPN methodology. The RPN (Robust Performance Number) index was introduced by Trierweiler and Engell, (1997), and the *relative* RPN (Trierweiler, 2002) indicates how potentially difficult it is for a given system to achieve the desired performance robustly. The performance with 2, 6, and 12 times faster than the open loop response and 10% overshoot have been analyzed for the high ethanol concentration branch. The RPN indicates that is possible and easy to design a robust controller with the desired performance. Additionally the nonlinearity degree has also been analyzed by the nonlinear RPN in (Trierweiler and Diehl, 2009). This analysis, reveal a predominance of static nonlinearity to region A and dynamic nonlinearity to region C. The global plant nonlinearity, region *A* and *C* together, present static and dynamic contributions. However high performance (i.e. rise time $t_R = 0.10\ h$) becomes the global nonlinearity essentially static. Moreover, for this performance it is expected that a linear controller will be able to control the system in both stable operating regions. More details and discussions about these controllability analyses are explained in (Trierweiler and Diehl, 2009).

4. Control System

A linear PI MIMO controller was designed, based on RPN methodology tuning. The transfer matrix of the PI controller is given by

$$PI = \begin{bmatrix} -0.16\left(1+\dfrac{1}{0.113s}\right) & 1.70\left(1+\dfrac{1}{0.006s}\right) \\ 12.29\left(1+\dfrac{1}{0.629s}\right) & 6.42\left(1+\dfrac{1}{0.361s}\right) \end{bmatrix} \tag{5}$$

This quite simple controller can control all linearized models with a good performance as it is shown in Figure 2. Ten linearized models (five belonging to region *A* and other five belonging to region *C*) are controlled by the PI MIMO, proving that it is possible to control the global plant with a linear controller. The control system is tested in a nonlinear Jöbses's model for different set points at region *A*. The results are showed in Figure 3. The performance is very good, but in a real process application the C_{So} difficultly would be prepared instantly. For this reason a first order transfer function was inserted before the C_{So} bias sum (representing a mixer for example). This function adds 10 minutes time delay distributed in a first order curve. The Figure 3 shows the control system robustness characteristic, because no new design was performed.

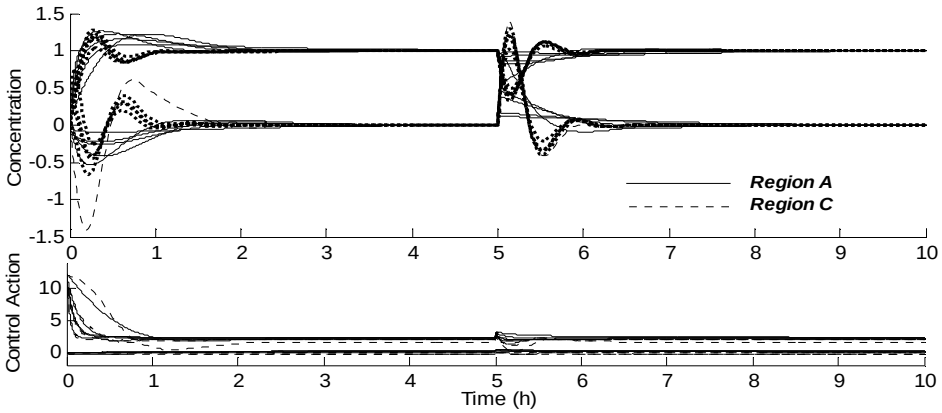


Figure 2 – PI MIMO: global plant control.

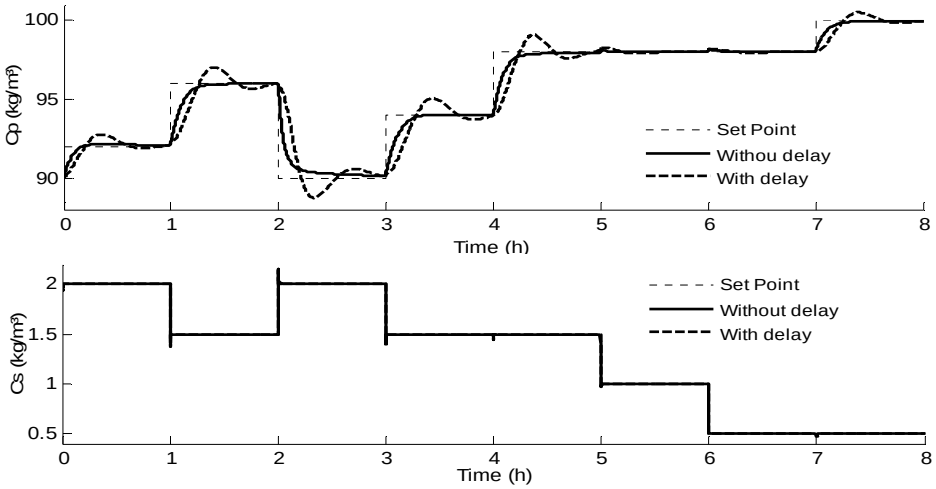
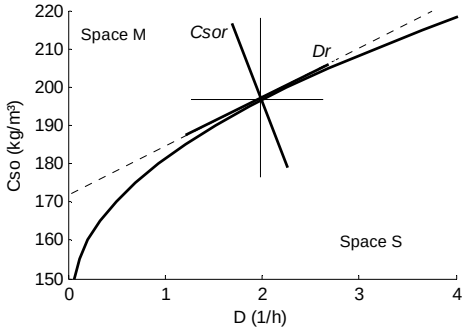


Figure 3 – Multivariable linear controller in the nonlinear model.

According to previously discussed, the operational optimum is the stability frontier. It is desirable to maintain the process at the space *M*, in the high yield field (region *A*). However immeasurable disturbance and/or unfeasible set points can lead the reactor convergence to the region *C*. To avoid this situation the control system actuation requires an operational limitation. The Figure 4 illustrates the restriction strategy. First the *D* and C_{So} axes are transposed to a new origin point at the stability frontier. Through a rotation matrix *R* the actions are recalculated to a rotated coordinates. The space *M* actuation is guaranteed since the C_{so} variable is positive. In the rotation matrix *R* the β angle is the angular coefficient of the tangent line to a stability limit point. The successful strategy results can be viewed in the Figure 5, for an unfeasible C_P set point.



$$R = \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix}$$

Figure 4 – Limit actuation problem solution.

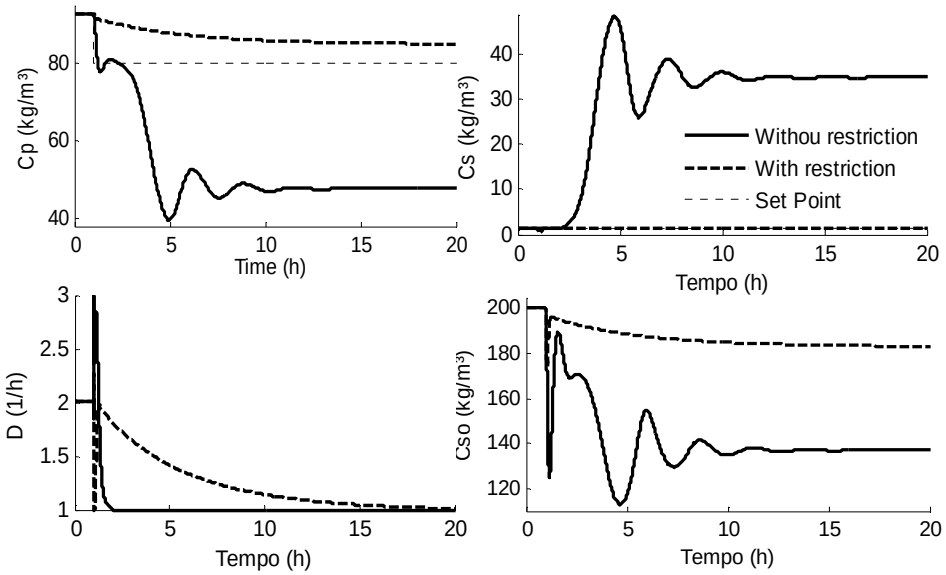


Figure 5 – Control variables without and with restrictions.

5. Conclusions

This contribution proposes a control strategy to high ethanol yield in an equilibrium multiplicity region. All analyses were based in the Jöbses's model. However, few experimental results have been confirmed the predicted results at high dilution rates so that more studies should be performed to the model validation. The codimension-2

bifurcation analysis reveals the multiplicity extension behaviour at Jöbses's model. A high stable yield operating region appears on the triple operational regions. This branch is linked with an unstable branch and the stability limit curve formed represents the operational optimum. So a control system was designed to maintain the process on this optimum.

A MIMO control structure was proposed, with the pairing $D \rightarrow C_P$ and $C_{So} \rightarrow C_S$. Based on the RPN methodology the robust performance is easily attainable. The fast controller performance minimizes the dynamic nonlinearity, becoming the global nonlinearity essentially static. The nonlinear analysis revels in a possibility to control the global process with a linear controller.

A multivariable PI controller was designed based on the RPN methodology. This controller could control the system in both stable operational regions (global control). Yet, it was capable to hold the operation on the region A, when the nonlinear model is simulated.

It was presented a safe strategy to maintain the operation at high yield region A. This strategy simply restricts the variable control actuation area. For this a rotation matrix change the action coordinates for satisfy a limit criterion. When Cso_R is bigger than zero the system is on the multiplicity space *M*.

References

A. Maheca-Botero, P. Garhyan, S.S.E.H. Elnashaine (2006) Non-linear characteristics of a fermentor for ethanol production and their implications. *Nonlinear Analysis: Real World Applications* 7, 432-457.

A.B. Jarzebski (1992) Modelling of oscillatory behaviour in continuous ethanol fermentation. *Biotechnology Letters* 2, 137-142.

A.J. Daugulis, P.J. Mclellan, J. Li (1997) Experimental investigation and modeling of oscillatory behavior in the continuous culture of *Zymomonas mobilis*. *Biotechnology and Bioengineering,* 56**,** 99-105.

C. Ghommidh, J. Vaija, S. Bolarinwa, J.M. Navarro (1989) Oscillatory behaviour of Zymomonas in contínuos cultures: a simple stochastic model. *Biotechnology Letters* **9**, 659-664.

I.M.L. Jöbses, G.T.C. Egberts, K.C.A.M. Luyben, J.A. Roels, (1986) Fermentation kinetics of *Zymomonas mobilis* at high ethanol concentrations: Oscillations in continuous cultures. *Biotechnology and Bioengineering,* 28**,** 868-877.

J.O. Trierweiler (2002) Application of the RPN Methodology for Quantification of the Operability of the Quadruple-Tank Process. *Brazilian Journal of Chemical Engineering,* 19**,** 195-206.

J.O. Trierweiler, F.C. Diehl (2009) Analysis, control, and operational optimization of a *Zymomonas mobilis* reactor with equilibrium multiplicity. *Submitted to ADCHEM 2009*

J.O. Trierweiler, S. Engell (1997) The Robust Performance Number: A New Tool for Control Structure Design. *Computer Chemical Engineering* 21**,** 237-243.

L.J. Bruce, D.B. Axford, B. Ciszek, A.J. Daugulis (1991) Extractive fermentation by *Zymomonas mobilis* and the control of oscillatory behavior. *Biotechnology Letters* 13, 291-296.

P. Garhyan, S.S.E.H. Elnashaie (2004) Utilization of mathematical models to investigate the bifurcation and chaotic behavior of ethanol fermentors. *Mathematical and Computer Modelling* 39, 381-427.

P. Rogers, Y. Jeon, K. Lee, H. Lawford (2007) *Zymomonas mobilis* for Fuel Ethanol and Higher Value Products. *Biofuels.* Springer Berlin / Heidelberg.

P.J. McLellan, A.J. Daugulis, J. Li (1999) The incidence of oscillatory behavior in the continuous fermentation of *Zymomonas mobilis*. *Biotechnology Progress* 15, 667-680.

S.S.E.H. Elnashaie, Z. Chen, P. Garhyan, P. Prasad, A. Mahecha-Botero (2006) Practical Implications of Bifurcation and Chaos in Chemical and Biological Reaction Engineering. *International Journal of Chemical Reactor Engineering 4*.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Robust output-feedback nonlinear model predictive control using high-gain observers

Rui Huang[a], Sachin C. Patwardhan[b], Lorenz T. Biegler[a]

[a] *Department of Chemical Engineering, CMU, Pittsburgh, PA, USA 15213.*
[b] *Department of Chemical Engineering, IIT Bombay, Mumbai, India, 400076.*

Abstract

A robust output-feedback nonlinear model predictive control (NMPC) system is proposed by combining state-feedback NMPC with a certain type of high-gain observer. We show that robust stability of this formulation deteriorates from its state-feedback NMPC counterpart because of observation error. Then we demonstrate this idea through simulation studies of a CSTR example and offset free regulatory behavior is obtained because the estimated error is used to correct model-mismatch in the controller.
Keywords: Robust controller, Output-feedback system, High-gain observer.

1. Introduction

Nonlinear model predictive control (NMPC) has received considerable attention over the past years. Schemes to guarantee the stability of the closed-loop systems with state-feedback NMPC have been widely studied (Limon et al, 2008 & Mayne et al, 2000). However, the assumption that the plant is available for measurement usually does not hold in practice. Hence, the output-feedback NMPC controller is required by integrating a state-feedback NMPC controller with a state observer. For linear systems, the well-known separation principle guarantees that the closed-loop system is nominally stable if both the controller and the state observer are stable. For nonlinear systems, the issue of stability of the closed-loop systems is still open due to the lack of general valid separation principles. In order to analyze the nominal stability for certain nonlinear output-feedback systems, two possibilities have been explored: 1) developing a certain equivalent separation principle and designing observers accordingly (Atassi & Khalil, 1999, Findeisen, et al, 2003); 2) considering the observer error in the NMPC controller (Michalska & Mayne, 1995, Magni et al, 2004).

The above formulations focus on the nominal closed-loop stability or stability in the presence of vanishing perturbations. In this work, we propose a robust output-feedback NMPC based on a high-gain observer. We show the robust stability analysis for the proposed formulation with bounded observation errors. Section 2 starts with the introduction of the output-feedback system and stability analysis. Section 3 presents simulation examples of continuous stirred tank reactor (CSTR). We also demonstrate that the proposed formulation generates offset-free regulatory and servo behavior under moderate perturbations in model parameters. Section 4 concludes the paper.

2. Problem Formulation

2.1. State-feedback NMPC and stability

In this work, we consider the dynamic model of a plant with output,

$$x_{k+1} = f(x_k, u_k, \theta_k) \qquad y_k = Cx_k, \quad k \geq 0 \tag{1}$$

where $x_k \in \Re^{n_x}$, $u_k \in \Re^{n_u}$, $y_k \in \Re^{n_y}$ and $\theta_k \in \Re^{n_\theta}$ are the plant states, controls, outputs and uncertainty parameters respectively, defined at time steps t_k with integers $k \geq 0$. Without losing generality, we assume that the given plant (1) has an equilibrium point at the origin, that is $f(0,0,0) = 0$.

Given $x(k)$, the current state value at time step t_k, the state-feedback NMPC formulation can be described in the following discretized form:

$$V(x_k) := \min \sum_{j=0}^{N-1} l(z_{k+j}, u_{k+j}) + F(z_{k+N}) \tag{2a}$$

$$\text{s.t} \quad z_{k+j+1} = f(z_{k+j}, u_{k+j}, 0), \quad j = 0, \ldots, N-1 \tag{2b}$$

$$z_k = x_k, z_{k+j} \in \mathcal{X}, z_{k+N} \in \mathcal{X}_f, u_{k+j} \in \mathcal{U} \tag{2c}$$

where N is the finite time horizon, z_{k+j} is the sequence of the predicted state variables, and u_k is the calculated control action based on the plant state x_k at time step t_k. Note the uncertainty parameter is *0* in the controller, introducing the plant-model mismatch. The calculated state-feedback control law from (2) can be written as $u_k = h(x_k)$, and the plant state at the next time step t_{k+1} can be expressed as $f(x_k, h(x_k), \theta)$.

The closed-loop stability with the state-feedback control law has been widely studied. Here we summarize the key definitions and analysis developed in Limon et al 2008. Here, $|\cdot|$ denotes the Euclidean vector norm in $\Re^n$ or the associated matrix norm. For a given sequence w_k, we denote $\|w\| \triangleq \sup_{k>0}\{|w_k|\}$. For two $\mathcal{K}$ functions τ_1 and τ_2, we define $\tau_1 \circ \tau_2(s) \triangleq \tau_1(\tau_2(s))$.

Lemma 1: *If f(x,y) is a uniformly continuous function in both $x \in A$ and $y \in B$, then there exist $\mathcal{K}$ functions τ_1 and τ_2, such that $|f(x_1, y_1) - f(x_2, y_2)| \leq \tau_1(|x_1 - x_2|) + \tau_2(|y_1 - y_2|)$.*

Theorem 1: (Limon et al 2008) *Assume function $f(x, h(x), \theta)$ is uniformly continuous in θ, and the plant (1) is nominally asymptotically stable. If there exists a uniformly continuous Lyapunov function $V(x)$ for plant (1), then the closed-loop system with state-feedback controller is Input-to-State stable (ISS).*

2.2. Output-feedback NMPC and stability

Consider now a state observer for plant (1),

$$\hat{x}_{k+1} = g(\hat{x}_k, y_k, h(\hat{x}_k)), \quad k \geq 0 \tag{3}$$

where $\hat{x}_k$ is the state estimated from the outputs at time step t_k. Hence, the initial condition in the NMPC formulation is modified as $z_k = \hat{x}_k$, instead of $z_k = x_k$ in (2c). The calculated output-feedback control law is $u_k = h(\hat{x}_k)$, and the plant state at t_{k+1} can be modified as $f(x_k, h(\hat{x}_k), \theta)$.

In order to establish the robust stability of the output-feedback NMPC formulation, we make use the following assumptions.

Assumption 1 (Observer assumption)

a) the initial observer error is bounded by a positive constant e, i.e. $|\hat{x} - x| \leq e$;

b) the future observer error satisfies $|\hat{x}_k - x_k| \leq \rho|e|\eta^{-k} + \beta(\|\theta\|)$, where $\rho \geq 0, \eta > 1$ are constants, β is a $\mathcal{K}$ function.

c) the observer is a high-gain observer.

Lemma 2: *There exists a constant $\epsilon \geq 0$ such that $\rho|e|\eta^{-k} \leq \epsilon$. As a result $|\hat{x}_k - x_k| \leq \epsilon + \beta(\|\theta\|)$.*

Lemma 3: *If $|\hat{x}-x| \le e$, then $|x|-e \le |\hat{x}| \le |x|+e$. For a $\mathcal{K}$ function $\alpha(\cdot)$, there are other $\mathcal{K}$ functions $\alpha_L(\cdot)$, $\alpha_U(\cdot)$ and positive constant c_1, c_2, such that $\alpha_L(|x|)-c_1 \le \alpha(|x|-e) \le \alpha(|\hat{x}|) \le \alpha(|x|+e) \le \alpha_U(|x|)+c_2$. Similarly if $|\hat{x}-x| \le \epsilon+\beta(\|\theta\|)$, we can find other $\mathcal{K}$ functions $\hat{\alpha}_L(\cdot)$, $\hat{\alpha}_U(\cdot)$, $\beta_L(\cdot)$, $\beta_U(\cdot)$ and positive constants c_3, c_4, such that $\hat{\alpha}_L(|x|)-\beta_L(\|\theta\|)-c_3 \le \alpha(|\hat{x}|) \le \hat{\alpha}_U(|x|)+\beta_U(\|\theta\|)+c_4$.*

Robust stability of this output-feedback NMPC can be established by the following:

Theorem 2: *Assume function $f(x,h(\hat{x}),\theta)$ is uniformly continuous in θ and x, and system (1) is nominally asymptotically stable and the observer satisfies Assumption 1. If there exists a uniformly continuous Lyapunov function $V(x)$, then the closed-loop system with output-feedback controller is Input-to-state practical stable (ISpS).*

Proof: The analysis is similar as the proof for Theorem 2 in Limon et al 2008, but we need to consider the observer error. From the continuity of $V(x)$ and $f(x,h(\hat{x}),\theta)$, there exist $\mathcal{K}$ functions α, σ_V, σ_f and σ_θ such that $V(f(\hat{x},h(\hat{x}),0))-V(\hat{x}) \le -\alpha(|\hat{x}|)$, $V(\hat{x})-V(x) \le \sigma_V(|\hat{x}-x|)$, and $V(f(x,h(\hat{x}),\theta))-V(f(\hat{x},h(\hat{x}),0)) \le \sigma_V(|f(x,h(\hat{x}),\theta)-f(\hat{x},h(\hat{x}),0)|) \le \sigma_V \circ (\sigma_f(|\hat{x}-x|)+\sigma_\theta(|\theta|))$. As a result:

$$\begin{aligned} V(f(x,h(\hat{x}),\theta))-V(x) &= V(f(\hat{x},h(\hat{x}),0))-V(\hat{x})+V(\hat{x})-V(x) \\ &\quad +V(f(x,h(\hat{x}),\theta))-V(f(\hat{x},h(\hat{x}),0)) \\ &\le -\alpha(|\hat{x}|)+\sigma_V(|\hat{x}-x|)+\sigma_V \circ (\sigma_f(|\hat{x}-x|)+\sigma_\theta(|\theta|)) \end{aligned} \tag{4}$$

Now, we need to ensure that the closed-loop system is robustly stable for both the initial stage of the observer when $|\hat{x}-x| \le e$ and later stages when the observer error is corrupted by the uncertainties, i.e. $|\hat{x}-x| \le \epsilon+\beta(\|\theta\|)$.

If $|\hat{x}-x| \le e$, from Lemma 3, we have $-\alpha(|\hat{x}|) \le -\alpha_L(|x|)+c_1$. In addition, we can find a constant $c_5 \ge 0$ such that $\sigma_V(e)+\sigma_V \circ \sigma_f(e)+c_1 \le c_5$ and a $\mathcal{K}$ function σ_2 such that $\sigma_V \circ \sigma_\theta(|\theta|)) \le \sigma_2(|\theta|)$. Consequently equation (4) leads to:

$$\begin{aligned} V(f(x,h(\hat{x}),\theta))-V(x) &\le -\alpha(|\hat{x}|)+\sigma_V(|\hat{x}-x|)+\sigma_V \circ (\sigma_f(|\hat{x}-x|)+\sigma_\theta(|\theta|)) \\ &\le -\alpha_L(|x|)+\sigma_2(|\theta|)+c_5 \end{aligned} \tag{5}$$

Then the closed-loop system is ISpS stable for the initial observer error.

If $|\hat{x}-x| \le \epsilon+\beta(\|\theta\|)$, we have $-\alpha(|\hat{x}|) \le -\hat{\alpha}_L(|x|)+\beta_L(\|\theta\|)+c_3$, then we can pick a constant $c_6 \ge 0$ and a $\mathcal{K}$ function σ_3 such that $\beta_L(\|\theta\|)+c_3+\sigma_V(|\hat{x}-x|)+\sigma_V \circ \sigma_f(|\hat{x}-x|)+\sigma_V \circ \sigma_\theta(|\theta|) \le \sigma_3(\|\theta\|)+c_6$. Then equation (4) leads to:

$$\begin{aligned} V(f(x,h(\hat{x}),\theta))-V(x) &\le -\alpha(|\hat{x}|)+\sigma_V(|\hat{x}-x|)+\sigma_V \circ (\sigma_f(|\hat{x}-x|)+\sigma_\theta(|\theta|)) \\ &\le -\hat{\alpha}_L(|x|)+\beta_L(\|\theta\|)+c_3+\sigma_V(|\hat{x}-x|)+\sigma_V \circ \sigma_f(|\hat{x}-x|)+\sigma_V \circ \sigma_\theta(|\theta|) \\ &\le -\hat{\alpha}_L(|x|)+\sigma_3(\|\theta\|)+c_6 \end{aligned} \tag{6}$$

Then the closed-loop system is ISpS stable for the later stage when observer error is corrupted by the uncertainties. ?

Remark: From the analysis, we see that the ISS stability of state-feedback system deteriorates to ISpS stability of output-feedback system, because of the corruption of observer error. For the output-feedback system, all the past uncertainty signals affect the closed-loop stability, while only the current uncertainty signal plays a role in the

robust stability for the state-feedback system. In addition, stronger assumption is required to guarantee the robust stability of the output-feedback system.

2.3. EKF as the High-gain observer stability

In this section we consider an extended Kalman filter (EKF) developed by Reif & Unbehauen (1999) as the state observer, satisfying Assumption 1.

$$x_{k+1}^- = f(\hat{x}_k, u_k, 0) \quad \hat{x}_k = x^- + K_k[y_k - Cx_k^-] \tag{7}$$

where x_k^- and $\hat{x}_k$ are called the prior and posterior estimate. K_k is the Kalman gain calculated from:

$$K_k = P_k^- C^T [CP_k^- C^T + R]^{-1} \quad P_{k+1}^- = \alpha^2 A_k P_k^+ A_k^T + Q \tag{8a}$$

$$P_k^+ = [I - K_k C] P_k^- \tag{8b}$$

where $A_k = \dfrac{\partial f}{\partial x}(\hat{x}_k, u_k, 0)$, $G_k = \dfrac{\partial f}{\partial q}(\hat{x}_k, u_k, 0)$ are the linearizations of the model. Q and R are symmetric positive definite matrices. $\alpha \geq 1$ is a exponential data weight to control the convergence rate of the EKF. Hence this EKF formulation is a high-gain observer. If $\alpha = 1$, this formulation reduces to the conventional EKF. Let $\varphi(\cdot, \cdot, \cdot)$ be the higher order term, the residual of the EKF can be defined by the following equations:

$$f(x_k, u_k, \theta) - f(\hat{x}_k, u_k, 0) = A_k[x_k - \hat{x}_k] + G_k\theta_k + \varphi(x_k, \hat{x}_k, u_k, \theta_k) \tag{9}$$

Defining estimation error as $\zeta_k = x_k - \hat{x}_k$, we have from (1), (7), (8) and (9) ,

$$\zeta_{k+1} = (I - K_{k+1}C)[A_k\zeta_k + r_k] \tag{10}$$

where $r_k = \varphi(x_k, \hat{x}_k, u_k, \theta_k) + G_k\theta_k$

Theorem 2: *Consider the EKF as stated by equations (7) and (8) and let the following assumptions hold:*

- *a) There are positive numbers $\bar{a}, \underline{p}, \bar{g}$ and $\bar{p}$ such that $|A_k| \leq \bar{a}$, $|G_k| \leq \bar{g}$, $\underline{p}I \leq P_k^- \leq \bar{p}I$ and $\underline{p}I \leq P_k^+ \leq \bar{p}I$, $\forall k \geq 0$.*
- *b) A_k is nonsingular for $k \geq 0$.*
- *c) $\theta_k \in \Omega_\theta \subset \Re^{n_\theta}$ where Ω_θ is a compact set.*
- *d) There are positive real numbers κ, γ_θ, $\mathcal{K}$ functions $\delta(\|\theta\|)$ and $\mu(\|\theta\|)$, such that $|\varphi(x_k, \hat{x}_k, u_k, \theta_k)| \leq \kappa |x - \hat{x}_k|^2 + \delta(\|\theta\|)$ whenever $|x - \hat{x}_k| \leq \mu(\|\theta\|) + \gamma_\theta \leq \bar{\varepsilon}$.*

Then with Assumption 1.a, there exist $\rho \geq 0, \eta > 1$ and $c_e \geq 0$ are constants, β is a $\mathcal{K}$ function, such that the error sequence ζ_k defined in equation (10) behaves according to $|\zeta_k| \leq \rho|\zeta_0|\eta^{-k} + \beta(\|\theta\|)$.

The proof follows the similar line from the proof in Reif & Unbehauen (1999). With this theorem, we can see Assumption 1.b and 1.c are true for this EKF.

2.4. Output-feedback NMPC with EKF

To achieve offset free closed-loop behavior, we propose to carry out multi-step predictions by explicitly using observer errors for future predictions as follows:

$$z_{k+j+1} = f(z_{k+j}, u_{k+j}, 0) + K_k\beta_{k+j}, \; y_{k+j} = Cz_{k+j} + \eta_{k+j}, \; j = 0, \ldots, N-1 \tag{11a}$$

$$\beta_{k+j+1} = \beta_{k+j}, \eta_{k+j+1} = \eta_{k+j}, z_k = \hat{x}_k, \; \beta_k = y_k - Cx_k^-, \; \eta_k = y_k - C\hat{x}_k \tag{11b}$$

The predictive controller is then formulated as

$$\min_u \ \sum_{j=1}^{N-1}[E_{k+j}^T W_E E_{k+j}] + \sum_{i=0}^{q-1} \Delta u_{k+i}^T W_{\Delta u} u_{k+i} + E_{k+N}^T W_\infty E_{k+N}$$

$$s.t.\ z_{k+j} \in \mathcal{X}, z_{k+N} \in \mathcal{X}_f, u_{k+i} \in \mathcal{U}, j = 0, \ldots, N-1,\ i = 0, \ldots, q-1, \quad (11a,b)$$

$$u_{k+j} = u_{k+t_i} \text{ for } t_i \leq j < t_{i+1},\ t_0 = 0 < t_1 < t_2 < \ldots < t_{q-1} = N-1. \quad (12)$$

where $E_{k+j} \triangleq (y_{k+j} - y_r)$, $W_E, W_{\Delta u}$ and W_∞ are positive semi-definite matrices.

3. Simulation Studies

We consider a simulated NMPC scenario with a nonlinear CSTR model represented by the following differential equations:

$$\frac{dz_c}{dt} = (z_c - 1)/u_1 + k_0 z_c \exp(-E_a/z_T) \quad (13a)$$

$$\frac{dz_T}{dt} = (z_T - z_T^f)/u_1 + k_0 z_c \exp(-E_a/z_T) + \nu u_2 (z_T - z_T^{cw}) \quad (13b)$$

This system involves two states $z = [z_c, z_T]$ corresponding to dimensionless concentration and temperature, and two manipulated inputs, corresponding to the inverse of dilution rate (u_1) and cooling water flow rate (u_2). The model parameters are $z_T^{cw} = 0.38$, $z_T^f = 0.395$, $E_a = 5$, $\nu = 1.95 \times 10^4$ and k_0 is an uncertainty parameter in the plant with nominal value $k_0 = 300$ in the model. The system is operated at a stable steady state $z_c = 0.1247$ and $z_T = 0.74070$ corresponding to $u_1 = 20$ and $u_2 = 378$. The NMPC is formulated using $W_E = W_\infty = \text{diag}[1 \times 10^6\ 1 \times 10^5]$, prediction horizon N = 20, control horizon q =5, with input blocking and each block equals to 4 samples, and sampling time is 1. $W_{?u}$ is chosen as a null matrix. The EKF is tuned with $Q = \left[\frac{\partial f}{\partial u}|_{z_{ss}, u_{ss}}\right] \bar{Q} \left[\frac{\partial f}{\partial u}|_{z_{ss}, u_{ss}}\right]^T$, where $\bar{Q}$ is chosen to be the possible variations in the inputs $\text{diag}[6.25, 0.04]$, and $R = \text{diag}[1 \times 10^{-6}, 1 \times 10^{-6}]$ which is the possible covariance of the outputs. To ensure that EKF works as a high gain observer, α in equation (8a) is chosen equal to 2.5.

Figure 1 presents the variation of controlled output and state estimation errors generated in response to a sequence of +/- 40 % step changes in model parameter k_0 (see Figure 2) for two different scenarios, 1) without output noises, 2) outputs are corrupted with white noise with standard deviation of $[1 \times 10^{-3}, 1 \times 10^{-3}]^T$. The corresponding variation of manipulated inputs is presented in Figure 2. It is clear from Figure 1 that, each time after a step change is introduced in k_0, the state estimator errors generated by the high gain observer settle to a constant value in a few samples. Moreover, the high gain observer's performance is hardly influenced by the tuning of the controller. It may be noted that the proposed output-feedback NMPC formulation eliminates the offset in both the cases.

4. Concluding Remarks

The purpose of this paper is to point out that although there are no general separation principles for nonlinear systems with plant-model mismatches, the closed-loop stability can still be ensured by using certain types of observers, and the controller performance can be improved by considering the observer error and tuning the observer. In future, this technique will be implemented on large scale applications and other types of observers will also be studied.

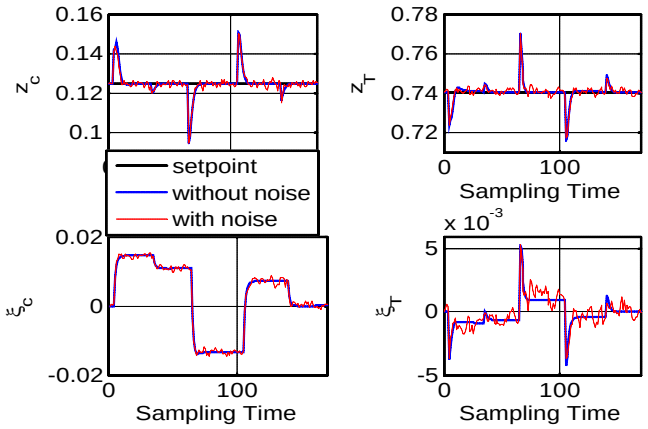


Figure 1: Variation of controlled outputs and state estimation errors

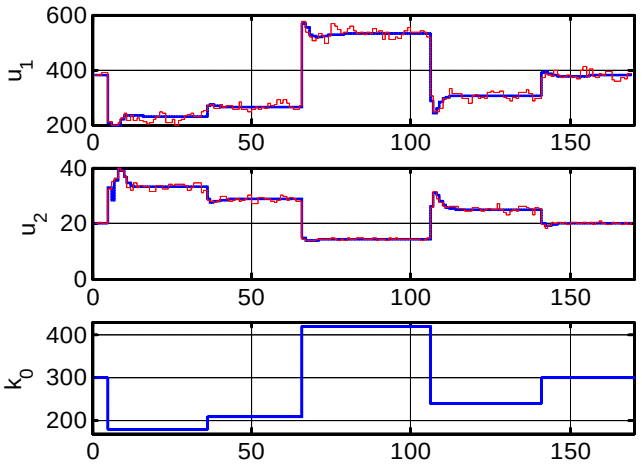


Figure 2: Variation of manipulated input and parameter disturbance

References

A. Atassi & H. Khalil, 1999. "A separation principle for the stabilization of a class of nonlinear systems", IEEE Trans. on Auto. Contr, 44: 1672.

R. Findeisen, L. Imsland, F. Allgöwer & B.A. Foss, 2003. "State and output feedback nonlinear model predictive control: an overview", European J. of Control, 9: 2.

D. Limon, T. Alamo, D. Raimondo, D.M. Peña, J. Bravo & E. Camacho, 2008. "Input-to-state stability: a unifying framework for robust model predictive control", In Assessment and Future Directions of Nonlinear Model Predictive Control, Springer, in press.

L. Magni, G.De Nicolao & R. Scattolini, 2004. "On the stabilization of nonlinear discrete-time systems with output feedback", International J. of Robust and Nonlinear Control. 14:1379.

D. Mayne, J. Rawlings, C. Rao & P.O.M. Scokaert, 2000. "Constrained model predictive control: stability and optimality", Automatica, 36: 789.

H. Michalska & D. Mayne, 1995. "Moving horizon observers and observer-based control", IEEE Trans. on Auto. Contr, 40: 995.

K. Reif & R. Unbehauen, 1999. "The extended Kalman filter as an exponential observer for nonlinear systems", IEEE Trans. on Auto. Contr, 47: 2324.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Scilab/Scicos: An alternative tool for real-time monitoring and advanced control of Fieldbus industrial systems

Thiago V. Costa, Ana M. F. Fileti, Flávio V. Silva

LCAP/DESQ - Department of Chemical Systems Engineering, School of Chemical Engineering, University of Campinas, Av. Albert Einstein, 500, CP 6066, CEP13083-970, Campinas-SP, Brazil

Abstract

The present work deals with real-time data acquisition and advanced control evaluation utilizing the open source scientific platform Scilab. Implementation and visualization of online data with the Scicos toolbox and utilization of OPC technology were discussed. The feasibility and effectiveness of the proposed methodology was shown by means of application of fuzzy controllers to a bromelain enzyme precipitation process instrumented with Foundation Fieldbus devices. Results confirmed Scilab/Scicos suitable for HMI and control systems applied in small industrial applications.

Keywords: real-time control, OPC connection, precipitation process, fuzzy logic.

1. Introduction

Personal computers (PC's) have become a reliable option to process control, as they are practical, have low cost and high computational performance. While dedicate machines have advantages over personal computers like optimized operation and specific hardware system for control, PC's have a greater range of programming languages (*C, Java* etc.) and development tools (Seborg et al, 2004). The greater availability of computer tools can improve the diversification of current industrial control loops while the real-time implementation and evaluation of the control algorithms in a PC control application can be performed with minimal effort through an OPC communication. Based on Microsoft *OLE/COM* technology, OPC (ole for process control) consists of standardized protocols to easy information flow and interoperability among automation, control loops and software systems. These common specifications eliminate the need of proprietary hardware vendor drivers, allowing the development of control systems, human machine interfaces (*HMI*) and supervisory control and data acquisition (*SCADA*) systems in scientific softwares, such as Scilab, making use of its mathematical toolboxes in the formulation of real-time advanced control algorithms, as proposed by Peng and Ma (2006).

Scilab is a free open source scientific platform for numerical computations developed in 90's at *INRIA* (*"Institut National de Recherche en Informatique et Automatique")*, maintained since 2003 by a consortium composed by companies and academic institutions (www.scilab.org). Scilab is not properly intended for real-time control; it's a simulation and analysis program. Even though, it is fast enough to be considered for soft real-time applications, which means that a suitable control loop could be evaluated in not so fast sampling times (more than 1s for example). This fortunately covers most of industrial chemical processes. For the implementation and

visualization of online data this work proposes the utilization of the toolbox Scicos, a block diagram editor primarily intended for dynamic, discrete and hybrid systems simulation (Najafi et al, 2003; Nikoukhah and Steer, 2006). Drag and drop and interconnection of user defined and pre-programmed function blocks allow algorithm implementation in an easier and straightforward manner. Interface and interoperability between the process data and the user is done in the same fashion with display boxes and scale widgets resulting in clear and functional supervisory screens.

To show experimentally the benefits of this methodology, a bromelain precipitation plant was real time monitored and controlled by Scilab/Scicos. Temperature control of the precipitation tank was performed by advanced controllers based on PI, PI+D and PI+PD forms of feedback digital fuzzy logic controllers using Scilab's mathematical routines. Online closed loop tests evaluating the fuzzy logic controllers in comparison with classical PID controller were developed and accompanied by performance indexes like *ITAE* (integral time average absolute error criteria), energy consumption and system behavior aspects.

2. Scilab/Scicos OPC Communication

The Scilab toolbox for OPC connection was developed by Peng and Ma (2006). The functions were compiled in a dynamic linking library (*DLL*) and linked with Scilab, using its interface to *C* and *FORTRAN* codes. Table 2.1 shows the toolbox functions available from the *OPC.dll* file.

Table 2.1 OPC toolbox functions (adapted from Peng and Ma, 2006)

opc_connect	connect OPC server
opc_disconnect	disconnect OPC server
opc_add_group	add a group to OPC server
opc_add_item	add OPC items to OPC group
opc_item_write	write OPC items
opc_item_read	read OPC items

The OPC elements arrangement is shown in Figure 2.1. After the connection with the OPC server, a *Group* responsible for *Items* organization (tags or memory addresses) is added. *Items* are placed in a string vector and called by their position.

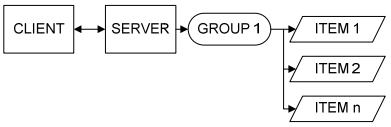


Figure 2.1 *OPC* client/server connection.

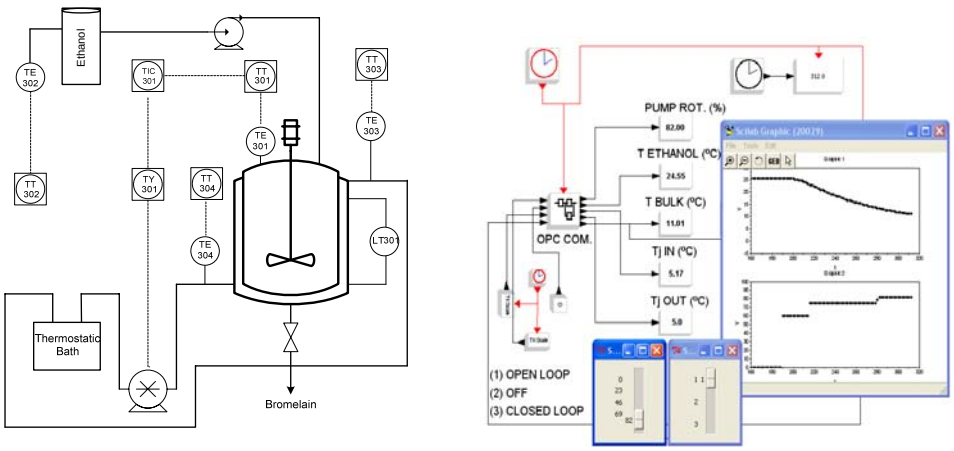
Scicos interface with field instruments is carried out in a quite easy procedure. First, the diagram has to be configured for *real-time scaling* in *Simulation/Setup* option. Then, the instructions for *connection, group* and *items* adding are placed in *Diagram/Context*. These instructions will run every time the diagram is loaded. Reading and writing of process data is performed by Scilab code blocks *(scifuncs)* which are placed in the diagram loop.

3. Bromelain Precipitation Process

The bromelain precipitation process (see Figure 3.1a) consists in a jacketed fed-batch stirred tank, instrumented with Fieldbus Foundation intelligent devices, where bromelain enzyme is recovered from pineapple pulp by precipitation with alcohol at low temperatures (5°C).

The tank (1000ml of nominal capacity) is initially charged with 150mL of pineapple extract. Precipitating agent (ethanol, 99.5%) at room temperature is then constantly added at a rate of 0.6mL/s until the final volume of 750mL is reached (Leite et al, 2008). The bulk temperature, represented by TE301, is controlled in order to prevent the enzyme denaturation during the precipitation process.

Setpoint is fixed at 5°C and control action is performed manipulating the coolant flow rate (propylene glycol) by means of pump rotation with the help of a frequency converter. Instrumentation consists of Fieldbus Foundation intelligent devices like temperature and level transmitters. Interface with the system is done by the Distributed Field Interface (DFI302) configured by *Smar Syscon*® software and connected to Scilab through the OPC communication. Process data is monitored with the utilization of display boxes and graphic windows as shown in Figure 3.1 (b).



(a) (b)

Figure 3.1 Instrumentation diagram and Scicos control application for the bromelain precipitation process.

4. Fuzzy Control

Fuzzy control algorithm is briefly described in the next steps: (i) the inputs error *(E)* and changing of the error *(ΔE)* are weighted in a *fuzzification* process by means of the activation level obtained from their respective membership functions; (ii) influence of inputs on the output is inferred by a set of rules (*IF … THEN… ELSE*) consisting of the linguistic model of the fuzzy controller; (iv) active rules are logically combined and an output fuzzy set is generated; (v) control action *(U)* for the PD (proportional and derivative) algorithm and the changing in control action *(ΔU)* for the PI (proportional and integral) algorithm are obtained from the *defuzzification* of the previously generated output fuzzy set. For a more complete description of fuzzy control from classical PID

approaches see Espinosa and Vandewalle, (1999); Li and Gatland (1995) and Pivoñka (2002).

PI and PD fuzzy control diagrams are shown in Figure 4.1 (a) and (b) respectively. The *FLS* (fuzzy logic structure) block evaluates the fuzzy logic linguistic model. The inputs are scaled before entering the *FLS* by *GE* and *GCE* gains. The outputs are scaled back to crisp values by *GCU* and *GU,* respectively.

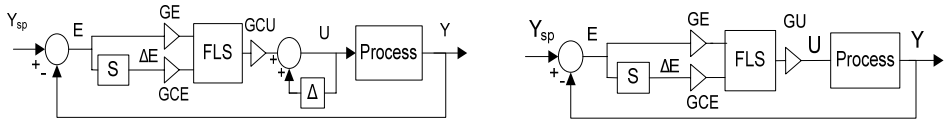


Figure 4.1 Fuzzy logic control diagrams.

The scale gains constitute tuning parameters for the fuzzy controller. For their computation, Pivoñka (2002) methodology was utilized. It relates classical PID parameters *(Kc, Ti, Td)* to the scale gains through a coefficient *M.* This coefficient maps the rule base normalizing the inputs within the limits of [-1, 1]. *Equation (1)* shows the scale gains: τ corresponds to *Ti,* for the PI structure, and *Td* for PD structure.

$$GE = \frac{1}{\tau}; GCE = \frac{\tau}{M}; GU = \frac{KcMTs}{Ti}; GCU = KcM \tag{1}$$

In this work, inputs *E* and *ΔE* and output *ΔU* consist of seven fuzzy sets (see Figure 4.2): the adjectives are *NB* (negative big), *NM* (negative medium), *NS* (negative small), *ZE* (zero), PS (positive small), *PM* (positive medium) and *PB* (positive big). The 7x7 linear rule base for the fuzzy control structures are presented in Table 4.1. Inputs are logically combined by *AND* (product) operator and rules evaluated by *ELSE* (max) operator. The *defuzzification* corresponds to the fuzzy centroid algorithm. PI+PD fuzzy controller corresponds to the parallel combination of PI and PD fuzzy controllers while PI+D fuzzy controller corresponds to the PI fuzzy controller with crisp derivative action.

Table 4.1 Rule base for fuzzy PI/PD controller

e \ Δe	*NB*	*NM*	*NS*	*ZE*	*PS*	*PM*	*PB*
NB	NB	NB	NB	NM	NS	NS	ZE
NM	NB	NM	NM	NM	NS	ZE	PS
NS	NB	NM	NS	NS	ZE	PS	PM
ZE	NB	NM	NS	ZE	PS	PM	PB
PS	NM	NS	ZE	PS	PS	PM	PB
PM	NS	ZE	PS	PM	PS	PM	PB
PB	ZE	PS	PS	PM	PB	PB	PB

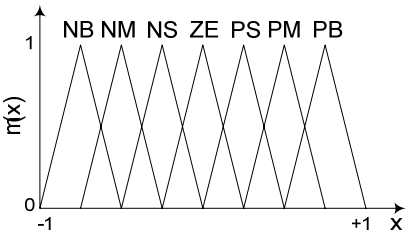


Figure 4.2 Fuzzy set for *E, ΔE* and *ΔU*.

5. Results

The online experimental tests with the proposed control application showed efficient performance for the bromelain precipitation process. Although the sample time errors have not been directly measured, no abnormal situation or errors concerning the system time response were observed. Furthermore, the achieved results are consistent with previous works (see Leite, 2007) performed with the same process.

Figure 5.1 shows the PID controller compared to the open loop response using a fixed rotation of 40% in the coolant pump. The system behavior under PI, PI+D and PI+PD fuzzy controllers is shown in Figure 5.2. Performance indexes are given in Table 5.1. Tuning parameters were investigated by Leite (2007) and consists of *Kc=37%/°C* and *Ti=30s* for the *PI* controller and *Kc=35%/°C*, *Ti=18s* and *Td=7s* for the PID controller. A sampling time of *2s* was utilized. Fuzzy controllers were tuned with the same parameters as the classical PI and PID controllers with a scale coefficient *M=5*.

The large initial overshoot noted in both Figures 5.1 and 5.2 results from the small effective surface area of the wall between the early volumes of precipitating mixture and coolant liquid, heat provided by the dilution of the alcohol in water and by the precipitation process as well.

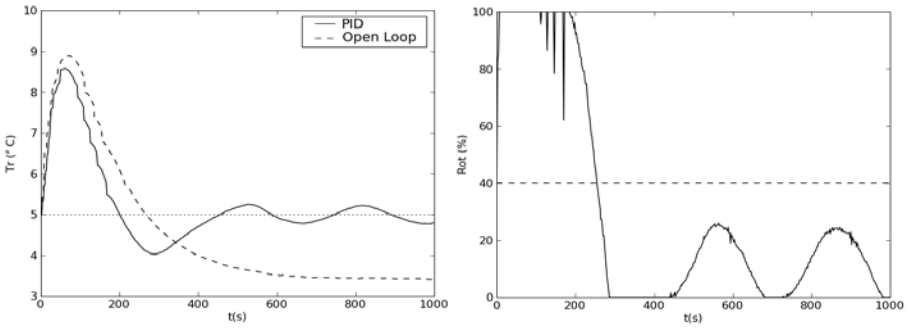


Figure 5.1 System behaviors under classical PID control.

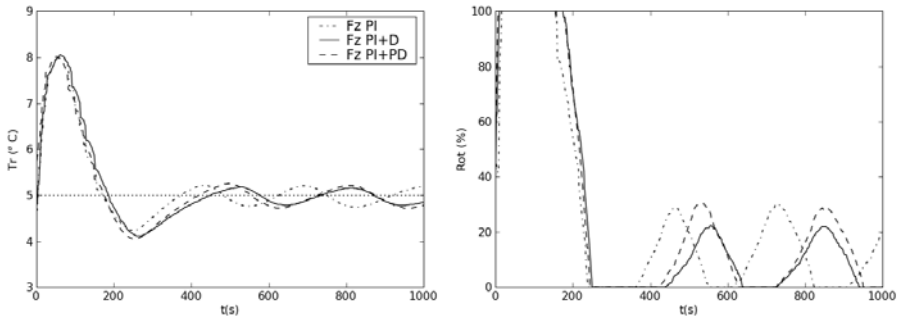


Figure 5.2 System behavior under fuzzy controllers.

As Table 5.1 and Figures 5.1 and 5.2 shows, fuzzy controllers performed better when compared with the classical PID controller. Lower overshoot is a necessary condition for the final product quality. The energy saving, 14% with Fuzzy PI when compared with the PID controller, and the small manipulated variable saturation, required for maintenance cost reduction, are known as key prerequisites for a cost saving production. These characteristics in addition with controller performance index *(ITAE)* indicate the fuzzy PI and the fuzzy PI+D as better control algorithms for the present process.

Table 5.1 Control performance indexes.

Controller	*ITAE*	Overshoot (°C)	MV Saturation (s)	Energy Consumption (kJ)
PID	130424	3.57	184	228
Fuzzy PI	104114	3.05	138	196
Fuzzy PI+D	125430	3.04	162	194
Fuzzy PI+PD	113034	3.01	166	207

6. Conclusions

Through the application of fuzzy control in the bromelain precipitation process, implementation and real-time utilization of advanced controllers in the open source software Scilab were discussed. Results showed better performance of the fuzzy controllers when compared with the classical PID controller. Tracking of control time evaluation confirms Scilab/Scicos suitable for standalone *HMI* and control system in small applications. Future experiments with different control algorithms, such as model predictive control and neural network approaches should be employed in order to promote their utilization not only in the academia, but in industrial applications as well.

References

D. Seborg, T. Edgar & D. Mellichamp, 2004, Process Dynamics and Control, 2nd ed, Wiley.

H. Li & H. Gatland, 1995, A new methodology for designing a fuzzy logic controller, IEEE transactions on systems, man, and cybernetics., 25, 505-512.

J. Espinosa & J. Vandewalle, 1999, Fuzzy Control and PID, Lecture Notes of the Studienamiddaf, Tentoonstelling PID-regelaars: moderene afstellingsmethodes.

M. Leite, 2007, Estudo Comparativo do Desempenho de Controladores Fuzzy e Convencional Aplicados em um Bioprocesso, Universidade Estadual de Campinas, São Paulo - Brasil.

M. Leite, T. Fujiki, F. Silva & A. Fileti, 2008, Determinação de Condições Operacionais de um Processo de Precipitação da Bromelina via Planejamento Experimental, XVII Congresso Brasileiro de Engenharia Química.

M. Najafi, A. Azil & R. Nikoukhah, 2003, Implementation of Continuous-Time Dynamics in Scicos, Proceedings of 15th European Simulation Symposium.

P. Pivoňka, 2002, Comparative Analysis of Fuzzy PI/PD/PID Controller Based on Classical PID Controller Approach, Proceedings of the 2002 IEEE International Conference on Computational Intelligence, 1, 541-546.

R. Nikoukhah & S. Steer, 2006, SCICOS A Dynamic System Builder and Simulator User's Guide, Version 1.0, Institut National de Recherche en Informatique et Automatique.

Y. Shimanuki, 1999, OLE for Process Control (OPC) for New Industrial Automation Systems. IEEE International Conference on Systems, Man, and Cybernetics,6, 1048-1050.

Z. Peng & L. Ma, 2006, The realization of SCADA based on Scilab, Institute of Industrial Control of Zhejiang University.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Flexible and Configurable Steel Plant Scheduling System

Iiro Harjunkoski and Margret Bauer

ABB Corporate Research, Wallstadter Str. 59, 68526 Ladenburg, Germany
E-mail address: iiro.harjunkoski@de.abb.com

Abstract

Steel plant scheduling is very complex, particularly because of the multitude of products and a large variety of possible steel plant configurations and production rules. In this paper, we will present how the steel plant scheduling model based on the one reported in Harjunkoski and Grossmann (2001) has been generalized and adapted to be applied to a real production environment. Besides minor adjustments to the mathematical model, the work involves advanced software development and interface definition as well as implementation and testing.

Keywords: Steel plant, scheduling, connectivity, ISA-95.

1. Introduction

In the steel industry, a strong demand exists for systems that can generate daily production schedules with the assistance of mathematical optimization methods. The demand for such a system stems from two needs: 1) to take the workload off the planning personnel by automatically generating a daily schedule and 2) to potentially increase the production by finding better schedules than is manually possible in the short available time. Depending on the steel plant in question, one need may be more important than the other. In both cases, it is necessary to produce a schedule that reflects the actual operation of the plant and considers all physical production rules. The task is very challenging and numerous, also recent contributions can be found in the literature, for instance Atighehchian et al. (2009) and Pacciarelli and Pranzo (2004).

A steel plant scheduling model describes a multi-stage production process that can schedule different steel grades and families. Considerable effort has been spent on the adaptability of the number and order of stages to allow for all possible plant configurations, see e.g. Harjunkoski and Sand (2008). The model is formulated as a mixed-integer linear programming (MILP) problem and decomposes the overall problem into several sub-problems to achieve fast solution times.

However, even the most carefully constructed production planning method can never see the daylight in a real production environment, unless it can be tied to the real-world situation. This means that the production schedule needs to be generated on-demand and not as a once-off computation. The input data, that is, the list of products to be scheduled, reflects directly the order book and can vary considerable depending on the customer in particular and the market situation in general. Also, plant specific rules may have to be considered that could not be anticipated in a general case, for example, one machine is preferred over another simply because it is newer and more efficient.

In this paper, we will describe the steps and challenges of building a scheduling system. Only the implementation in a real production environment will reveal the validity of any

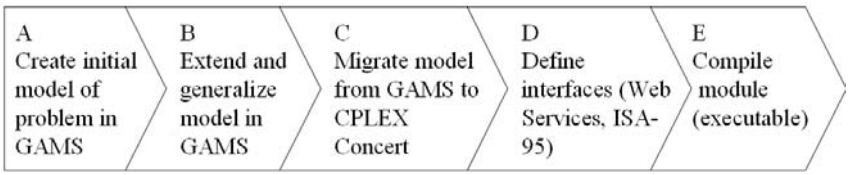


Figure 1. Development steps of a scheduling optimization solution.

model. There is always a certain measure of doubt involved whether the model will work and it is vital to minimize the doubt by generalizing the model for easy adaptation when constructing it. The purpose of this paper is to show the aspects that had to be considered in a real-world implementation and the lessons to be learned for future model developments. These aspects are arguably as important as the derivation of the model. In particular, we will focus on the following:

- *Configurability*: The model has to be adapted to the specific plant configurations which may change throughout the lifetime of the scheduling system.
- *Robustness*: The system must always return an output, at minimum a statement indicating that no result could be obtained.
- *Connectivity*: The system must be able to be integrated into the surrounding computational environment and/or users.
- *User interface*: For industrial application, the critical sets and parameters must be changeable through user friendly interfaces.

In the following sections, these aspects will be discussed in detail. Fig. 1 shows the major development steps that were followed in this work. These steps together with test results from a real-life implementation will be discussed.

2. Implementation aspects

Production in a steel plant has been described in detail in previously (Harjunkoski and Grossmann, 2001). For clarity, the process is briefly introduced in Fig. 2 showing the four stages EAF (electric arc furnace), AOD (argon oxygen decarburization unit), LF (ladle furnace) and CC (continuous caster). As seen in Fig. 2, there can also be more stages and several parallel equipment at each stage and the products (heats) are passed between stages on defined routes.

2.1. Configurability

Ideally, the same model should be applied to different steel plants. A good model should therefore not only focus on one single problem instance, but should be adaptable to various cases. The most important modifications in a steel plant apply to:

- Plant configuration, routing and transportation times
- Certain production rules that may change (setup, clean-up times)
- Maintenance operations

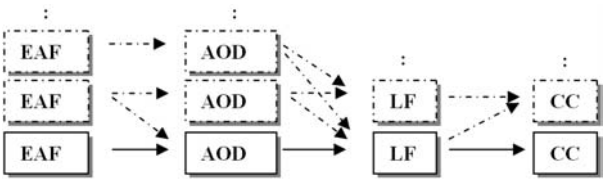


Figure 2. Plant layout and routing between equipment

- Current status of the plant
- Customer / order priority handling (due/release dates, forbidden equipment)
- Objective function components and penalties (lateness, hold-up times, etc)

Also, a steel plant is never static as new process equipment are installed or removed every few years. For example, a new EAF might be added to overcome a bottleneck of steel melting. At the same time, a second equipment may drop out for a long-term maintenance. If the lifespan of the scheduling system should last more than a few months, it is mandatory that it can be easily adapted to incorporate new equipment. If every minor change requires an action of a high-level expert, the engineering costs will be immense. This poses a challenge to the modeling which must be general and adaptable to different plant layouts by changing as few parameters as possible. In Harjunkoski and Sand (2008), the issue of generalizing the scheduling model of a disaggregated solution was discussed. Here, we need to consider all the modeling and solving steps.

Ideally, all issues relating directly to specific production data can simply be handled as parameters in the system input. The model must thus be flexible enough to be able to support various plant layouts, situations in which an equipment is inactive for a given time, changes in transportations times between equipment due to the situation at the site or breakdown of transportation means. In particular, the plant configuration should be handled as model input parameters specifying the number of stages and equipment.

Basic production data can be entered as standard model parameters, e.g. production times and due dates, which are typically represented in any scheduling model. This becomes more complex when production times vary. Production times then have to be expressed as a variable in the model with lower and upper bounds. This may make some production constraints non-linear and complicates significantly the use of a discrete-time representation. Therefore, this work is restricted to a continuous-time approach.

2.2. Robustness

During implementation, different use cases of the scheduling system have to be defined, that is, generation of for example daily or weekly schedules that reveal different characteristics of the production environment. Some products may trigger certain rules, others not and some may not even be captured by the implementation. Scheduling twenty products of the same type is certainly a different task than scheduling twenty different products. The use cases thus result in different problem size and complexity.

When supporting the actual running of a production, the scheduling system should always produce a result – and if not possible this should be clearly communicated to the user with some hints of a possible problem source. Also, the solution time has to be within reasonable limits for all the use cases. It is usually not acceptable to let the user

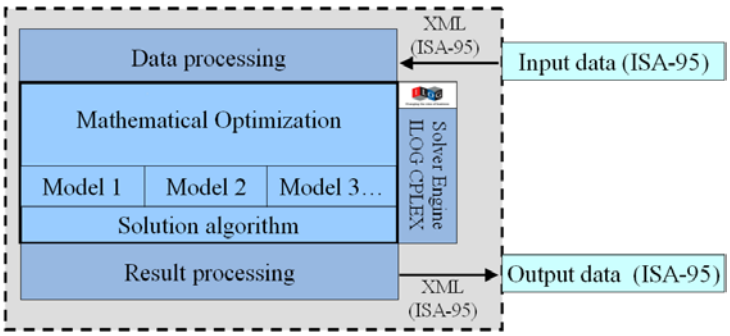
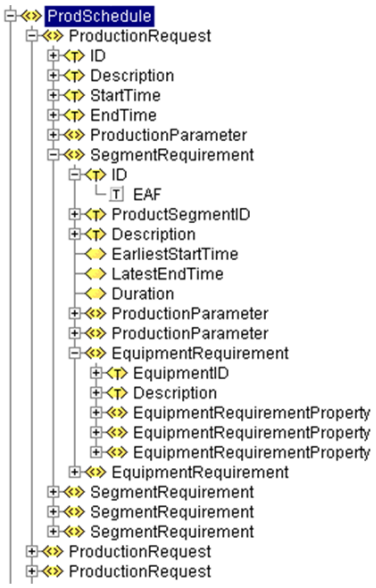


Figure 3. System architectural overview.



Production Request

Fieldname	Fieldname desription
ID	Unique ID of production request (order)
Description	Description of the production request, for example steel grade
StartTime	Scheduled start time of production request
EndTime	Scheduled end time of production request
ProductionParameter	Parameter related to the order
SegmentRequirement	Requirements of the segments or stages (e.g. EAF) to which the production requests may be assigned
EquipmentRequirement	Requirements of the equipments (e.g. EAF 1) to which the production requests may be assigned

Fig 4. ISA-95 XML structure highlighting a production schedule and description of fieldnames.

wait for the result for more than a couple of minutes. The optimization routine should be aborted after an agreed time, returning the message to possibly reduce the number of products. Developing a system that can return an intelligent result in all (foreseeable) situations requires actions on all levels: from the modeling to the software development. The approach here varies strongly from the commonly applied scheduling systems in that respect that we aim to optimality or close-to-optimality, i.e. utilize MILP methods as far as possible, supported by intelligent heuristics that makes the solution several magnitudes faster without losing the strive to optimality.

For instance the disaggregation process in Harjunkoski and Grossmann (2001) builds upon the idea of splitting and merging large optimization problems by the use of optimization methods. The various models can easily be covered by the solution structure, as seen in Fig. 3. Other possibilities includes using time limits and if a solution should be available at the given time the optimization should simply wait until a first feasible solution has been found or abort the task and inform about the limiting time constraint.

A common challenge for real production problems is that as long as a system can handle a certain amount of order efficiently, the user naturally wants to increase the number of orders. This leads to a continuously growing problem size and the fact of combinatorial explosion must be considered in the core of the solution. It is therefore necessary to agree on the outset on the time window, that is, daily or weekly schedule, and thus the maximum number of orders that can be scheduled.

2.3. Connectivity

The scheduling system should be embedded in the IT architecture of the steel plant. This concerns the inputs and the outputs as well as the parameterization of the system. The input data comprises the production orders as well as the plant configuration parameters. As a result, the model is decoupled from the graphical user interface and the visualization development work can be carried out separately from the model adaptations. In order to enable a highly flexible and well defined connection, ISA-95

compliant interfaces through XML files were defined. The ISA-95 standard describes models and terminology for the information exchange between enterprise resource planning (ERP) and, among others, scheduling system. Fig. 4 shows an example of an ISA-95 structure. Both production orders and production history are normally entered and stored in the ERP. ISA-95 as a standard does not provide all the needed elements for such a specific industry and problem. However, these have been added in a way that is enabled by the standard.

Fig. 3 shows that the optimization problem itself is implemented in a very flexible way. The required optimization models can be added or modified and are coordinated by a solution algorithm. The solution algorithm calls the optimization models as needed, using the underlying MILP/LP solver ILOG CPLEX Concert Technology.

Data pre-processing is mainly required for consistency checking of the input data and for converting it into the solution memory data structures. For example, it is easy to capture the current status incorrectly and assign two products to a machine at the start. This has to be checked before starting the optimization and a message with an explanatory error message should be returned.

The output XML file contains the start and end time of each product on each equipment and is also ISA-95 compliant as it may be stored in the ERP system. The result post-processing is necessary to generate the output file from the optimization results consistently with the input data.

2.4. User Interface

The user interface has to accommodate for two main tasks as highlighted in Table 1: Generate a production schedule and configure the production parameters. Configuration is a less frequent task that should be only carried out by more senior users with an overview insight of the complete production process. The standard functionality is to select an input file, start the optimization and then display the results. The input file is a list of products specifying the product ID, type, due date and current production status. When generating a rolling schedule, some parts of the products are already in the process and it is necessary to capture which product is already at what stage on which machine. The output is a list of these products specifying the product ID, start time and end time on each equipment. The output is best displayed in a Gantt chart.

Arguably the biggest challenge when designing a scheduling system is to accommodate for the needs of the person who operates that system. Any intelligent person operating the system will only trust and then use it if she a) understands why the schedule has been generated in the way it has, b) tune the behaviour and c) if she has the possibility to manually change the automatically generated schedule. The latter reason is a much debated issue.

Task description	Actions	User	Frequency
Generate production schedule	• Select input file for new schedule • Start optimization • Display the generated schedule	Scheduler	Several times a day
Configure production parameters	• Specify the production / operational parameter • Upload changes	Plant operating manager	Infrequent

Table 1. Description of use cases of the graphical user interface.

Fig 5. A Gantt Chart of an implemented test case solution

3. System Development Steps and Test Cases

The more the implementation aspects of configurability, user interaction, robustness and connectivity have been considered in conceptual modeling phase, the easier, the shorter the system development will be. Steps for the model adaptation are as follows:

1. Collect production rules in interviews
2. Define use case scenarios and desired results
3. Iteratively
    a. Test the use case scenarios
    b. Discuss results
    c. Adapt model to consider rules

It is important to realize that the model cannot be adapted in one single shot and the recursion has to be repeated several times. Also, it is quite likely that the use cases will not cover all situations that may later occur when the system is in operation over a number of years. Thus, exceptions have to be anticipated in all variables and parameters.

The system can generate acceptable solution to normal two to three days scheduling problems within a few seconds / minutes. An example of a resulting Gantt Chart with around 50 heats and nine equipment is shown in Fig. 5.

4. Conclusions

In this paper, we have discussed the challenges that occur when developing a scheduling solution from an existing scheduling model. Only in very rare, possibly never heard of, cases can the model be used without any changes. If you have constructed a scheduling model then it is most likely that it requires adaptation to the specific plant configuration. The more generic and flexible the model has been formulated with an industrial environment in mind, the easier the development of the scheduling system.

References

Atighehchian A., Bijari M. and Tarkesh H. (2009). A novel hybrid algorithm for scheduling steel-making continuous casting production, Computers & Operations Research, Volume 36, Issue 8, August 2009, pp. 2450-2461

Harjunkoski I. and Grossmann I.E. (2001). A Decomposition Approach for the Scheduling of a Steel Plant Production, Computers chem. Engng, 25, pp. 1647-1660

Harjunkoski I. and Sand G. (2008). Flexible and Configurable MILP-Models for Meltshop Scheduling Optimization. Proceedings, 18th European Symposium on Computer Aided Process Engineering, Lyon, June 1-4, 2008

Pacciarelli D. and Pranzo M. (2004). Production scheduling in a steelmaking-continuous casting plant. Computers & Chemical Engineering, Volume 28, Issue 12, pp. 2823-2835

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Monitoring penicillin G acylase (PGA) production using principal component analysis (PCA)

Edson R. Nucci, Antonio J.G. Cruz, Raquel L.C. Giordano and Roberto C. Giordano

LaDABIO – Laboratory for Development and Automation of Bioprocesses, Department of Chemical Engineering, Federal University of São Carlos, P.O. Box 676, 13565–905, São Carlos – SP, Brazil. E-mail address: roberto@ufscar.br

Abstract

The complexity of biological processes turns infeasible the development of detailed, structured phenomenological models of the cultivation of microorganisms in bioreactors. Therefore, cause-effect relations between on-line measurements and the state variables that are important for the optimal operation of industrial fermenters are sometimes hard to ascertain. In this context, data pre-treatment techniques are useful for control and fault detection. Among them, principal component analysis (PCA) has an important role. This work presents a case study of the application of this technique during real experiments, where the enzyme penicillin G acylase (PGA) was being produced by *Bacillus megaterium*. PGA hydrolyzes penicillin G to yield 6-aminopenicilanic acid (6-APA) and phenyl acetic acid. 6-APA is an important substrate, used to produce semi-synthetic β-lactam antibiotics. The algorithm was implemented for on-line detection of deviations from the desired process behavior. The experiments were carried out in a 2-liter bioreactor (Applikon®) operated in batch and fed-batch mode in different experimental conditions. A static PCA model was initially developed for the process, and its results are presented here. Hotteling's T^2 was the discrimination criterion employed in this multivariable problem, and the method showed a high sensibility for fault detection in all real cases that were studied.

Keywords: Multivariate statistical process control (MSPC), fault detection, *Bacillus megaterium* cultivation, bioreactor monitoring

1. Introduction

Process monitoring and fault diagnosis are the determinant for the successful operation of bioreactors and for quality control. In batch and fed-batch processes, small changes in the operating conditions may impact severely the quality of the desired product, frequently a bio-molecule which has its concentration analyzed off-line. If the product does not complain to the quality standards, the entire batch is lost. The industrial demand for reliable methodologies for on-line fault detection is therefore evident. The implementation of early fault detection algorithms would allow the operator to take corrective measures before the batch process is accomplished (Xu et al., 2006).

A large amount of process data can be collect and stored in modern biotechnology industries, whose processes have computer-based monitoring and control. Therefore, multivariate statistical process control (MSPC) has been widely used for fermentation processes (Hu and Yuan, 2008; Gnoth *et al.*, 2008). In MSPC the correlation between the original variables is considered, thus decreasing the number of false alarms. Statistical tools such as principal component analysis (PCA) are commonly used to explore historical data (Lopes and Menezes, 2004). Most often, PCA is employed in

statistical control to get a quick overview and to detect deviations from the desired process behavior (Chiang *et al.*, 2006; Al-Alawi *et al.*, 2008; Gnoth *et al.*, 2008).
This work presents a real case study of the on-line application of PCA in order to recognize fluctuations and their respective effects on faults that occurred during the cultication of *B. megaterium* for the production of the enzyme penicillin G acylase (PGA). Penicillin G acylase (E.C.3.5.1.11) is an enzyme of great importance for the pharmaceutical industry, used in the production of 6-aminopenicilanic acid (6-APA), a key component for the synthesis of β-lactam antibiotics.

2. Material and Methods

The microorganism employed was *Bacillus megaterium* ATCC 14945, from Fundação Tropical (Campinas, SP, Brazil). The experiments were carried out in a 2-liter bioreactor (Applikon®). Dissolved oxygen concentration was controlled by manipulating stirrer speed, temperature was kept at 30°C and the aeration rate at 2.0 $L.min^{-1}$ (1 atm and 21°C). The bioreactor was coupled to a data acquisition system (National Instruments®) and on-line variables were stored each 10 seconds. Further details about the composition, procedure and laboratory analyses can be found in Silva et al. (2006) and Nucci et al. (2007).

3. Methodology

PCA is used to capture the major normal statistical correlation between measurements of process variables, from residuals that quantify their main variances. Some kind of statistics of these residuals can then be used to detect and diagnose faults (Xu et al., 2007).
PCA involves finding the eigenvalues of the sample covariance matrix, which are the variances of the principal components. For a normalized (mean centered, variance scaled) sample Matrix *X[n,m]* with *n* samples and *m* variables, PCA will find *m* uncorrelated new variables, the variance of which decreases from first to last (Albert and Kinley, 2001). Let the new variables be represented by t_i for a particular sample *i* as follows:

$$t_i = \sum_{j=1}^{m} X_j \times p_{ji} \tag{1}$$

The first principal component t_1 is found by selecting p_i, so that t_1 has the largest possible variance subject to the condition shown by Equation 2:

$$\sum_{i=1}^{m} p_i^2 = 1 \tag{2}$$

In covariance matrix, c_{ij} is the covariance between variables X_i and X_j. Of course, the diagonal element c_{ii} is the variance of X_i. The variances of the individual principal components are the eigenvalues of the matrix *C*, and the sum of the eigenvalues is equal to the sum of the variances of the original variables. For *m* input variables there will be *m* principal components, some of which might be negligible, if the original variables were either correlated or collinear. By retaining only the first *r* principal components, the *X* matrix is approximated by the following equation:

$$\hat{X} = \sum_{i=1}^{r} t_i \times p_i^T + E \tag{3}$$

In equation 3, *E* is the residual matrix, *p[m,n]* are the loading and *t[n,r]* are the scores.
This transformation results in several desirable mathematical and statistical properties that are associated with the transformed data (scores), enabling the derivation of statistical confidence limits. If the original variables are correlated, a reduced number of control charts

can be achieved ($r<m$) (Albert and Kinley, 2001). The ability to adequately represent an m variable data set in $r<m$ dimensions is one the main advantages of PCA, that is, data compression. More details about T^2 Hotelling, control limits and contribution plot can be found in AlGhazzawi and Lennox (2007); Hu and Yuan (2008).

4. Results and Discussion

A data set of seven runs of the PGA production process was used to test the procedure, with 7291 samples in normal operation conditions. 4448 samples were used for testing and validation. The monitored variables are listed in Table 1. The average length of cultivation is 24 h, and the sampling interval for laboratorial analyses is 2 h

Table 1 Variables monitored during penicillin G acylase production

No.	*Process variables*
1	agitation speed, rpm
2	aeration rate, $L.min^{-1}$
3	concentration of dissolved oxygen in the fermentation media, % saturation
4	pH
5	temperature, °C
6	fraction molar of CO_2, %

Fig. 1 shows T^2 analysis of a faultless experiment. The chart shows no violation of the control limits during all experiment, except for the initial samples, corresponding to the start-up of the process.

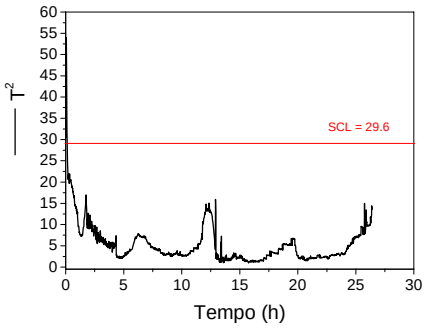
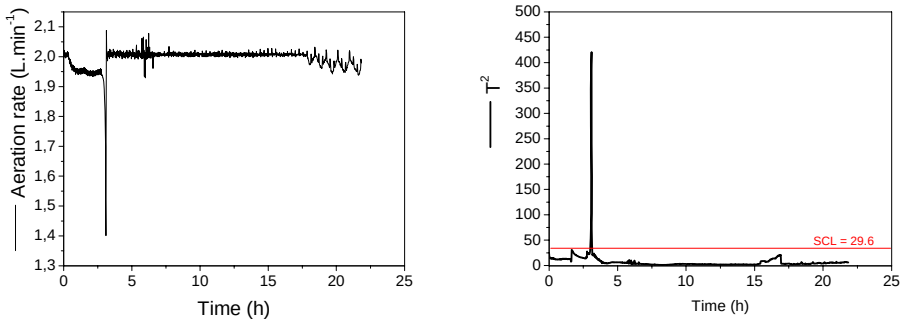


Figure 1 – On-line monitoring chart without faults. SCL = superior control limit.

Figure 2 shows a run where a fault occurred, the aeration rate (corresponding to variable 2) decreased from 2 $L.min^{-1}$ (the set-point) to 1.4 $L.min^{-1}$ approximately after 3 h of cultivation, see Figure 2(a). Figure 2(b) shows the sensitivity of the algorithm. Around 2 h the threshold of alarm (SCL = 29.6) was almost reached, and at 3 h there is a clear detection of the fault. The aeration rate influences the supplement of oxygen to the strictly aerobic microorganism: the decrease of oxygen mass transfer from the gas phase to the liquid medium, due to the smaller velocity of the air stream, will diminish the dissolved oxygen concentration thus restraining the biomass growth. It should be stressed that the operator might well have missed the deterioration of the air flow control in a normal situation, and so the chart (Fig. 2b) proves to be an important tool, indicating the occurrence of a fault.

Fig. 3 shows the contributions to the fault of each process variable. As expected, variable 2 provided the largest contribution to the deviation of the process from the normal operation region. Hence, the aeration rate is identified to be the source of the fault. Earlier detection of the faults in batch process and determination of its root cause will guide the operator to prompt corrective measures.



(a) (b)

Figure 2 – (a) On-line monitoring chart for fault 1 (aeration); (b) T^2 on-line monitoring charts using PCA for fault 1

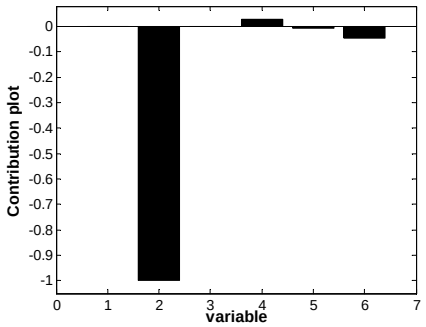


Figure 3 – Contribution plot for the aeration rate fault. Variable 1.Agitation speed; 2. Aeration rate; 3.Dissolved oxygen; 4.pH; 5. Temperature; 6.Fraction molar of CO_2.

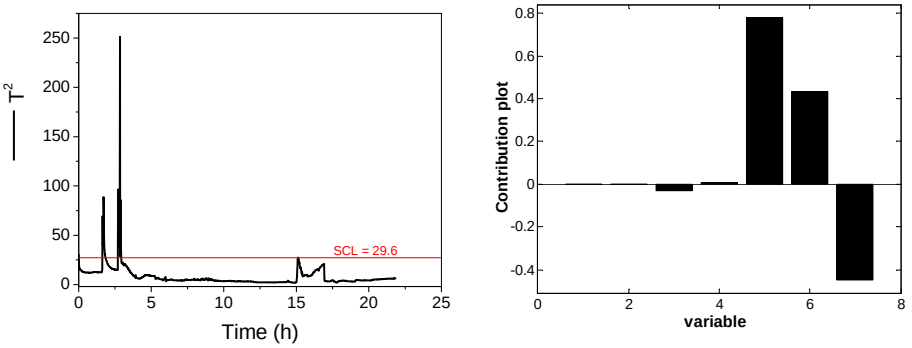
Robustness is an important characteristic of monitoring and fault detection. In almost every real industrial process, measured variables inherently contain noise that originates from various sources such as measurement devices and electrical equipment. Besides, outliers usually exist in sample data. If the signal-to-noise ratio is small, one may encounter misleading or biased results at the stages of data processing (Hu and Yuan, 2008).

In the previous fault, the input variable that caused the abnormal operation was directly monitored. In order to test the robustness of the method, this variable was removed from the samples data set, and the method was once again applied, now off-line, with the purpose of testing its robustness. Two other variables were introduced in this test. These variables were not used formerly, though being available in the data acquisition system. They are: the molar fraction of O_2 at the exhaust gas and the respiratory quotient (RQ), which is the ratio between produced CO_2 and consumed O_2 (calculated from the overall mass balance).

In this way, a new situation was emulated: instead of directly measuring the input variable responsible for the fault (the air flow rate), an output operational variable was sampled (molar fraction of oxygen at the gas outlet), and a second one was calculated (the RQ). Of course, there is a cause-effect correlation between the inlet air flow, the metabolic state of the culture and, consequently, the O_2 in the effluent gas and the RQ. Figure 4 shows that the method was able to identify the fault correctly, also in this case.

It should be noticed that this is a common situation in real processes, i.e., a non-monitored input variable reaching an abnormal level and causing a fault. In our case, the output measurements were able to provide enough information for the PCA algorithm to identify the fault. Fig 4a shows violations of the control limit between 2 and 3 h. The

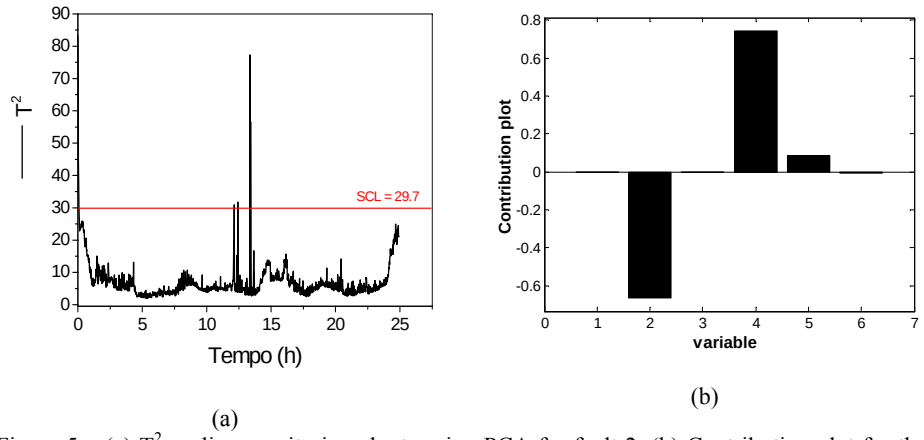
contribution plot (Fig. 4b) indicated a problem with variables 5, 6 and 7 (molar fraction of CO_2, of O_2 and RQ, respectively), i.e., these variable are strongly correlated. These results illustrate the robustness and efficiency of the method. Indeed, these variables are correlated, and directly affected by the reduction of the inlet flow rate of air.



(a) (b)

Figure 4– (a) T^2 on-line monitoring charts using PCA for fault 3; (b) Contribution plot for the pH fault. Variable 1.Agitation speed; 2.Dissolved oxygen; 3.pH; 4. Temperature; 5.Fraction molar of CO_2; 6.Fraction molar of O_2; 7. Respiratory quotient (RQ)

For the run shown in Figure 5, two faults occurred. The pH decreased until 6 h (after this time it started to increase from 6.95 to 7.30). And a problem appeared during the aeration rate data acquisition. The pH influences the growth of the microorganism, and consequently the enzyme activity because the growth is associated with the production of the enzyme. Figure 5a shows the T^2 on-line monitoring chart. It can be seen that between 12 and 14 h there is a violation of the control limit (SCL = 29.7). Fig. 5b shows the contributions of each of the process variables. This chart correctly highlights two variables as the source of the problem, variables 2 and 4



(a) (b)

Figure 5 – (a) T^2 on-line monitoring charts using PCA for fault 2; (b) Contribution plot for the pH and aeration rate fault. Variable 1.Agitation speed; 2. Aeration rate; 3.Dissolved oxygen; 4.pH; 5. Temperature; 6. Molar fraction of CO_2

In this case, besides an input variable, the air flow rate, a state variable, the pH, was also at abnormal values. These two variables were removed from the data set and the PCA was re-calculated with two new process variables, i.e., molar fraction of O_2 and RQ. T^2 chart was

calculated and similar results were obtained as Fig. 5a. On the other hand, the contribution plot, indicates those variables 4, 5 and 6 (fraction molar of CO_2, fraction molar of O_2 and RQ, respectively) provide the largest contributions to the deviation of the process from the normal operational region. Once again, the variables more closely related to the metabolic state of the culture had the higher contribution. It should be stressed that the term contribution here is within a statistical framework, and does not reflect a cause-effect relation. Anyway, the operator should translate these results as an indication that the deviation from the acceptable operational region is causing a metabolic stress to the microorganism, and thus seek for the causes of this unexpected behavior, but already excluding the variables with almost null contribution. Hu and Yuan (2008) obtained similar results using PCA during a cephalosporin fed-batch fermentation process. The authors indentified the faults in two abnormal experiments (decrease in aeration rate and agitation speed).

5. Conclusion

This article applied PCA (*principal component analysis*) to a bioreactor producing an enzyme, PGA, growing the strictly aerobic bacterium *B. megaterium* in an agitated and aerated tank. The method was able to filter normal random oscillations, and could detect faults without false positives. The obtained results demonstrate the power and advantages of the proposed approach. It was efficient and robust, identifying the abnormal data during the real experiments of validation in all cases. Even when the variable that caused the fault was not monitored, the results were good enough to support decision-making and the intervention of the operator. Therefore, this method may be very useful to ensure that each production run is a close as possible to pre-established optimal process trajectories.

Acknowledgments

The authors thank FAPESP, CNPq and CAPES for the financial support

References

Albert, S., Kinley, R.D., 2001, Multivariate statistical monitoring of batch processes: an industrial case study of fermentation supervision. TRENDS in Biotechnology, 19(2), 53-62.

AlGazzawi, A., Lennox, B., 2008, Monitoring a complex refining process using multivariate statistics. Control Engineering Practice,.16, 294-307.

Al-Alawi, S.M., Abdul-Wahab, S.A., Bakheit, C.S., 2008, Combining principal component regression and artificial neural networks for more accurate predictions of ground-level ozone. Environmental Modelling & Software. 23, 396-403.

Chiang, L.H., Leardi, R., Pell, R.J., Seasholtz, M.B., 2006, Industrial experiences with multivariate statistical analysis of batch process data. Chemometrics and intelligent laboratory systems. 81, 109-119.

Gonth, S., Jenzsch, M., Simutis, R., Lubbert, A., 2008, Control of cultivation processes for recombinant protein production: a review. Bioprocess and Biosystem Engineering. 31, 21-39.

Hu, K., Yuan, J., 2008, Statistical monitoring of fed-batch process using dynamic multiway neighborhood preserving embedding. . Chemometrics and intelligent laboratory systems. 90, 195-203.

Lopes, J.A., Menezes, J.C., 2004, Multivariate monitoring of fermentation process with non-linear modelling methods. Analytica Chimica Acta. 515, 101-108.

Nucci, E.R., Silva, R.G., Souza, V.R., Giordano, R.L.C., Giordano, R.C., Cruz, A.J.G., 2007, Comparing the performance of multi layer perceptrons networks and neuro-fuzzy systems for on-line inference of *Bacillus megaterium* cellular concentrations. Bioprocess and Biosystem Engineering, 30, 429-438.

Silva, R.G., Souza, V.R. Nucci, E.R., Pinotti, L.M., Cruz, A.J.G, Giordano, R.C., Giordano, R.L.C., 2006, Using a medium of free amino acids to produce penicillin G acylase in fed-batch cultivations of *Bacillus megaterium* ATCC 14945. Brazilian Journal of Chemical Engineering, 23, 37-43.

Xu, Z., Weiwu, Y., Huihe, S., 2006, Monitoring and Fault Diagnosis for Batch Process in Feature Extract in Fisher Subspace. Chinese J. chem. Eng., 14(6), 759-764.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Stochastic and Deterministic Performance Assessment of PID and MPC Controllers: Application to a Hydrotreater Reactor

Alain C. Carelli, Mauricio B. de Souza Jr.

Department of Chemical Engineering – Federal University of Rio de Janeiro
Av. Athos da Silveira Ramos, n. 149 - CT - Bloco E, Sala E-209, CEP 21941-909,
Rio de Janeiro – RJ, Brazil

Abstract

Hydrotreating processes are very important in petroleum refineries due to environmental reasons. However, the hydrotreater operational costs are very high, being increased when there is not an optimization on the catalyst lifetime, which is harmed by an inefficient control. At the same time, a control system tends to lose performance over time if its response is not monitored and thus there is no support information on how to adjust it. Enhanced controller projects have complementary systems to identify and diagnose performance reduction and also to implement predetermined solutions vis-à-vis the desirable type of output. When selecting the controller design and strategy, the thorough simulation of the underlying process is important in order to draw a greater number of possible alternatives of control and their application on this simulator, thus conferring more information for the decision. The goal of this paper was to investigate the response of feedback controllers in contrast with model predictive ones applied to the first bed of a diesel hydrotreating reactor, which was simulated using a phenomenological model, and to asset their deterministic and stochastic performances. The deterministic performance was evaluated through the integral of time multiplied by the absolute error (ITAE) and the quadratic variation of the control action; the stochastic performance was analyzed based on the normalized Harris index and on the variability and quadratic variation of the control actions. The predictive control algorithm used in this paper was the Generalized Predictive Control (GPC) one, developed by Clarke *et al.* in 1987. Besides comparing different controller designs, this work also analyzed the impact of the changes in the tuning parameters. The feedback controllers were tuned by different heuristic and model based methods. The relation between predictive control performance and variations in reference trajectory and internal model was investigated. This study adds background information on deterministic and stochastic performances of process controllers, through the application of different designs and strategies of controllers to a hydrotreating complex reactional process.
Keywords: process control, model predictive control, controller audit, petroleum refinery.

1. Introduction

Global environmental concern has pushed the contaminant contents in fuel oil specifications to lower levels, stimulating conscious investments in petroleum refineries with the goal of guaranteeing larger removal of contaminants, such as sulfur and nitrogen. Hydrotreating units are being improved and developed with new process control technologies, allowing cleaner output of fuel oils through reactions as

hydrodesulfurization, hydrodenitrogenation, hydrodeoxygenation, hydrodearomatization, hydrodemetallization and hydrodeasphaltenization.
Modern industries present many automatic control loops making the use of assessment tools mandatory in order to properly monitor all of them. MITCHELL and SHOOK (p. 3 – 4, 2005) related their experience in Model Predictive Control, showing four paths that MPC control applications may follow, depending on the level of support and monitoring they receive: 25 % fail within just months of commissioning; 40 % fail approximately 3 years of commissioning with inadequate support and no monitoring; 25 % continue to deliver benefits – albeit at a reduced level – with good support and no monitoring and 10 % will deliver increased benefits over time with adequate support and online monitoring.
The main aim of this paper is to compare the deterministic and stochastic performances of feedback and predictive controllers applied to a diesel hydrotreating reactor which was simulated using a phenomenological model.

2. Conceptual Aspects

2.1. Diesel Hydrotreating
The hydrotreating (HDT) unit considered in this work employs two trickle bed reactors (TBR) in series, each composed by two fixed beds, as showed in Figure 1. The oil feed is combined with makeup and recycle hydrogen and heated to the reactor inlet temperature. Heat is provided by heat exchange with the reactor effluent and by a pre-heater furnace. The reaction of hydrogen and oil occurs in the reactors in the presence of the catalyst. Quench gas (cold recycled hydrogen gas) is added between reactors and between catalyst beds of the reactors in order to maintain reactor temperatures in the desired range. The 2nd reactor effluent exchanges heat with the reactor feed in order to recover the heat released from the reactions. After cooling, the reactor effluent is flashed in the hot, high-pressure separator (HHPS) to recover hydrogen and to make a rough split between light and heavy reaction products. The liquid from HHPS has its pressure lowered, than it is sent to the low-pressure separators, and on to the product fractionator. The HHPS vapor is cooled and water is injected to absorb hydrogen sulfide and ammonia produced in the reactors. The mixture is further cooled to condense naphtha and gas oil and is flashed in the cold, high-pressure separator (CHPS). The CHPS separates vapor, liquid water, and liquid light hydrocarbons. The pressure of the liquid hydrocarbon is lowered and it is sent to the low-pressure separators. The water is sent to a sour water recovery unit for removal of the hydrogen sulfide and ammonia. The hydrogen-rich gas from the CHPS flows to the H_2S absorber. The purified gas flows to the recycle compressor where it is increased in pressure so that it can be used as quench gas and recombined with the feed oil. Liquid from the low-pressure separators is fed to the atmospheric fractionator, which splits the hydroprocessed oil from the reactors into the desired final products.
The model adopted here to represent the HDT reactor was introduced by CARNEIRO (1992) and applies the concept proposed by HLAVÁCEK (1982) to represent a fixed bed reactor through multiple CSTR-CELLs in series. The CSTR-CELL model considers mass and heat axial dispersion in the bed, mass diffusion and heat transportation between fluid and solid phases. The following assumptions were made: only one 1st order – with respect to the mean concentration of a pseudo-reagent A in the solid phase porous – reaction occurs and the reaction rate can be described by the Arrhenius equation; there is no volume variation in the reactor; the reactors are adiabatic; there is only one liquid and one solid phase with constant physical-chemical properties; there is

only longitudinal transport phenomena; and, there are non-linear interactions between kinetic and thermal processes.

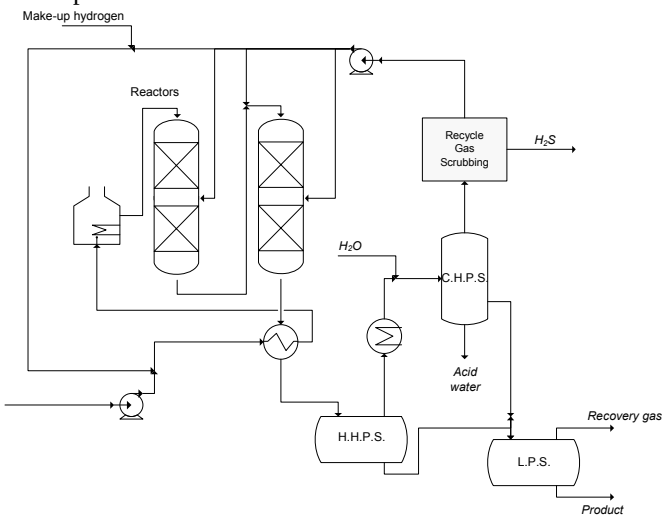


Figure 1 – Diesel hydrotreating process

CARNEIRO's (1992) model was employed here for being both able to represent the main process dynamics and simple, as it is composed only by ordinary differential equations.

2.2. Model Predictive Control – Generalized Predictive Control

The Generalized Predictive Control (GPC) algorithm covers a large variety of control goals in contrast to other methods, so that some of them can even be considered as GPC specific cases (CLARKE *et al.*, 1987).

With the premise that all process natural disturbances can be characterized by a stochastic disturbance, the principle of the superposition can be used to represent all disturbances as a unique influence in the output. Then, the process can be described by the following CARIMA (controlled auto-regressive and integrated moving average) model:

$$A(q^{-1})y(t) = q^{-d}B(q^{-1})u(t-1) + C(q^{-1})e(t)/\Delta(q^{-1}) \quad (1)$$

where A, B and *C* are polynomial in the backward shift operator q^{-1}, *d* is the dead-time, *e(t)* is an uncorrelated random sequence and $\Delta(q^{-1})$ is the differencing operator $1 - q^{-1}$.

The CARIMA model may be considered to be the most appropriated model for many industrial applications with non-stationary disturbances. In practice, it has two main types of disturbance: occurrence of random steps in random interval (e.g. changing of the product quality) and Brownian movement which is met in plants that depend on the energy balance (CLARKE, 1988).

The following cost function is assumed:

$$J = (Gu + f - w)^T (Gu + f - w) + \lambda u^T u \quad (2)$$

where, $G(q^{-1}) = E(q^{-1})B(q^{-1})$; $E(q^{-1})$ and f are provided by the Diophantine Equation; *w* is the reference trajectory or *set-point*; and, λ is a weighting sequence.

Assuming that there are no constraints in the control actions, the minimum of *J* can be met by equating to zero the gradient of the cost function with respect to the control actions. Therefore, the following result is used in order to obtain the control actions:

$$\Delta u = \left(G^T G + \lambda I\right)^{-1} G^T (w - f) \quad (3)$$

3. Methodology

The application of the controllers was not proceeded for all the control loops in the hydrotreating unit because the goal here was the comparison among several controller performances. The controller considered in this work was the master one that controls the outlet temperature of the first bed in the first reactor while manipulating the set-point for the slave controller. The slave controller controls the inlet temperature of this bed by manipulating the fuel flow that enters in the pre-heater furnace. However, the slave controller was disregarded of this work due to the insignificant time constant of the slave loop in contrast to the master loop.

3.1. Tuning Feedback Control

The feedback controller in the referred control loop was tuned by four different methods: Ziegler & Nichols (*ZN*), Cohen & Coon (*CC*), Error Integral (*EI*) and Internal Model Controller (*IMC*). Considering proportional (*P*), proportional-integral (*PI*) and proportional-integral-derivative (*PID*) controllers, ten feedback controllers (P/PI/PID – ZN; P/PI/PID – CC; PI/PID – EI; PI/PID – IMC) were tested. In Table 1, the tuning parameter values of each feedback controller can be observed.

Table1: Tuning feedback parameters

METHODS	*Controller Type*	*Proportional term (K)*	*Integral term (τ_I)*	*Derivative term (τ_D)*
Ziegler & Nichols	P	0.58	-	-
	PI	0.52	131.53	-
	PID	0.70	78.92	19.73
Cohen & Coon	P	0.72	-	-
	PI	0.52	272.75	-
	PID	0.84	953.22	160.96
Error Integral	PI	0.46	776.25	-
	PID	0.72	367.64	193.64
IMC	PI	0.30	557	-
	PID	0.60	811	174.45

3.2. Predictive Control Design

The controller internal model is the function which mimics the process and influences all the control strategy deduction. Two models were investigated here: one – given by Equation (4) – which describes the simulated process data very well and another – represented by Equation (5) – which renders an inferior description of the simulated process data.

$$\frac{y}{u} = \frac{2.06}{557s+1} e^{-500s} \tag{4}$$

$$\frac{y}{u} = \frac{2.06}{300s+1} e^{-600s} \tag{5}$$

The function considered to represent the reference trajectory was a first order model, with an adjustable parameter α. The larger α, the more cautious are the control actions and, inversely, the smaller α, the more energetic they behave. If α is equal to zero, the trajectory is constant and equal to the set-point. Three cases were analyzed by varying the internal model and α: *I.* good internal model (4) with α equal to 0.8; *II.* worse internal model (5) with α equal to 0.8; and, *III.* good internal model (4) with α equal to 0.7. The prediction (*N*) and control (*NU*) horizons were held, respectively, equal to 4 and 3 sampling times for the cases 1 and 3 and equal to 8 and 3 sampling times for the case 2. These parameters were adjusted because of change in the internal model.

3.3. Controller Performance Assessment

The deterministic regulatory performance assessment was accomplished through two indexes: the integral of the product between the time and the absolute error (*ITAE*), to evaluate the output behavior, and the mean of the squared variations in the control actions (*Su*), to evaluate the controller aggressiveness. The *Su* can be calculated through the following equation:

$$Su = \frac{1}{n+1}\sum_{k=0}^{n}\Delta u^2(k) \quad (6)$$

where *n* in the number of available data points. The following disturbance equation was introduced in the controlled variable for the stochastic performance assessment:

$$C(q^{-1})e(t) = e(t) + 0.5e(t-1) + 0.25e(t-2) \quad (7)$$

Three indexes were calculated to assess the stochastic performance of the controllers: the normalized performance index (ANORM), only for the feedback controllers, proposed by DESBOROUGH and HARRIS, 1992; the standard deviation of the controlled variable (*Std(y)*); and the mean of the squared variations in the manipulated variable (*Su*).

4. Results

The control of the first bed in the first reactor of the hydrotreating unit was simulated through block diagrams developed in Simulink/ Matlab 7.0. This simulator was composed by 48 differential equations that represented the bed physical-chemical behavior. In order to evaluate the behavior of the controllers, different scenarios were simulated. For the deterministic study, disturbances were introduced in the feed concentration and in the feed flow rate (± 5% from the respective steady-state values). For the stochastic study, a Gaussian white noise disturbance – e(t) – with variance 1 was introduced in the controlled variable.

Table 2: Comparison between the controllers

Controllers		Deterministic Performance								Stochastic Performance		
		Regulatory										
		Concentration				Flow						
		5%		-5%		5%		-5%				
		ITAE	Su	ITAE	Su	ITAE	Su	ITAE	Su	ANORM	std(y)	Su
GPC	Case 1	11,67	2,09	12,76	1,57	12,72	1,56	11,81	2,60	-	1,13	1,53
	Case 2	10,62	2,74	9,37	1,98	8,94	1,96	11,24	3,42	-	1,56	1,00
	Case 3	8,85	3,82	8,13	2,76	7,80	2,33	9,35	4,77	-	1,00	1,69
ZN	P	114,64	1,00	113,39	1,00	115,19	1,00	112,19	1,00	0,1	1,13	23,61
CC	P	101,13	1,59	100,50	1,63	101,98	1,62	99,08	1,54	0,4	1,15	38,62
	PID	1,00	3,45	1,00	3,41	1,00	3,47	1,00	3,75	0,6	1,37	3623,09
IMC	PI	2,11	1,29	1,91	1,08	1,72	1,06	2,34	1,52	0,3	1,42	15,10
	PID	1,15	2,10	1,16	1,94	1,16	2,03	1,15	2,42	0,4	1,05	938,71
EI	PI	1,65	1,75	1,62	1,55	1,62	1,52	1,73	1,97	0,3	1,44	17,71
	PID	3,48	6,43	3,14	5,79	2,15	5,29	5,24	8,23	0,6	1,76	3697,41

In Table 2, indexes normalized by the smaller value of the column (with exception of the ANORM) can be observed. So, the index equal to 1 in every column represents the smaller value. The cells of the table were colored based on the difference between the respective value and the smallest value of the column: the cells which present values equal to one are green; the cells which present values between one and eight are yellow; and, the cells which present values larger than eight are orange.

The controllers PI and PID – ZN presented unstable behavior, thus they were discarded from this work. The ratio of the pseudo dead time to the process time constant is often referred to as the "uncontrollability" factor (Fc) that is an indication of the quality of control that can be expected (CAMPOS and TEXEIRA, 2006). It is important to realize that the uncontrollability factor of the process is 0.9 and the Ziegler & Nichols method was developed for uncontrollability factor between 0.3 and 0.5 (CAMPOS and

TEXEIRA, 2006). Despite of the fact that the Cohen & Coon method was developed for uncontrollability factors larger than 0.3 (CAMPOS and TEXEIRA, 2006), the controller PI – CC also presented unstable behavior and it was also discarded.

5. Conclusions

The results of this work may also be useful for other chemical engineering applications as the simulated process was identified using a first order plus time delay model, which is a rather typical model in this area. In terms of deterministic performance, some PIs and PIDs presented unstable behavior. The stable ones achieved smaller ITAEs than the ones obtained using GPC. The Case 3 GPC presented an average ITAE nine times larger than the best obtained ITAE. Nevertheless, the Case 3 GPC reduced the controlled variable variability in 41%, requiring an effort in the control action (*Su*) 88% smaller than the smaller effort required among the feedback controllers (IMC-PI), which achieved a reduction of 16% in this variability. The largest achieved reduction in the controlled variable variability among the feedback controllers was 38% (IMC-PID) which required control actions 550 times more intense than the Case 3 GPC. Considering the ANORM index, the feedback controller that achieved a stochastic performance closer of the Minimum Variance Control (MVC) was ZN-P. CARELLI (2008) presents the figures of all the simulations performed. These figures are not reproduced here for space reasons, however Table 2 sums up the main results of the simulations. The performance evaluation for the SISO case studied indicates the GPC controller as the one that provides the best balance, taking into consideration deterministic and stochastic performances, as it is more equipped to deal with random disturbances (due to the CARIMA model) at the same time that provides an adequate deterministic behavior. The superiority of the model predictive controller can also be noticed by the facts that some of the feedback controllers rendered unstable behavior and that the focus of this work was not on finding optimum tuning parameters for the GPC controllers, but parameters that resulted in stable behavior and good regulatory behavior, both in the presence of deterministic and stochastic disturbances.

References

A. C. CARELLI, 2008, Stochastic and deterministic performance assessment of PID and MPC controllers: application to a hydrotreater's reactor – Programa em Tecnologia de Processos Químicos e Bioquímicos, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 111 p.

D. W. CLARKE, C. MOHTADI, P. S. TUFFS, 1987, Generalized predicitive control – part I: the basic algorithm. Automatica, v.23, n. 2, p. 137 – 148, ago.

D. W. CLARKE, 1988, Application of generalized predictive control to industrial processes, IEEE Control Systems Magazine, p. 49 – 55, abr.

H. P. CARNEIRO, 1992, Robust control of a fixed bed chemical reactor, Dissertation of Master Title in Chemical Engineering – Coordenação dos Programas de Pós-Graduação de Engenharia (COPPE), Universidade Federal do Rio de Janeiro, Rio de Janeiro, 142 p.

L. DESBOROUGH, T. J. HARRIS, 1992, Performance assessment measures for univariate feedback control, The Canadian Journal of Chemical Engineering, v. 70, p. 1186 – 1196, dez.

M. C. M. M. CAMPOS, H. C. G. TEIXEIRA, 2006, Controles típicos de equipamentos e processos industriais. Brasil: Edgard Blücher. 396 p. (In portuguese).

V. HLAVÁCEK, 1982, Fixed bed reactors, flow and chemical reaction, Alemanha: Verlag Chemic, p. 103 – 111.

W. MITCHELL, D. SHOOK, 2005, Industrial application of multivariable controller analysis and monitoring techniques, Canadá: Matrikon Inc., 23 p.

ACKNOWLEDGMENTS – This work was sponsored by 'Brazilian Research and Projects Financing Agency' (FINEP) and PETROBRAS (Grant 01.04.0902.00).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Model-based fault diagnosis using a hybrid dynamic simulator: Application to a chemical process

Nelly OLIVIER-MAGET[a,c], Gilles HETREUX[a,c], Jean-Marc LE LANN[a,c], Marie-Véronique LE LANN[b,c]

[a]*LGC-CNRS, Université de Toulouse, INPT-ENSIACET, 118, route de Narbonne F-31077 Toulouse Cedex 04, France*
[b]*LAAS-CNRS, Université de Toulouse, 7 avenue du Colonel Roche, F-31077 Toulouse, France*
[c]*Université de Toulouse ; INPT, INSA*

Abstract

This work presents a fault detection and isolation methodology for the monitoring of Hybrid Dynamic Systems. This methodology rests on a mixed approach, which combines a model-based method for the fault detection and an approach based on data (pattern matching) for the identification of fault(s). This methodology is integrated within the simulation platform *PrODHyS*, through the development of the module *PrODHySAEM*. The goal of this paper is to underline the potentialities of our approach for the diagnosis of the system. This methodology is illustrated by the studies of diagnosis problems in the field of Chemical Process System Engineering.

Keywords: Fault detection and diagnosis, Hybrid Dynamic Systems, generation of non binary signatures, Manhattan distance, extended Kalman filter, object differential Petri nets.

1. Introduction

Thanks to their large application field, numerous works on Hybrid Dynamic Systems (HDS) deal with modelling, stability and control (Zaytoon, 2001). Among them, some research works focus on the monitoring of these systems and many methods have been developed for Fault Detection and Isolation (FDI). The literature quotes as many fault detection and diagnosis methods as many domains of application (Venkatasubramanian, et al., 2003) and the techniques are generally classified as: methods without models – such as quantitative process history based methods or qualitative process history based methods– and model-based methods which are composed of quantitative and qualitative model-based methods. In our case, a model-based approach has been developed. It exploits the extended Kalman Filter to a hybrid dynamic system. The main idea is to reconstruct the outputs of the system from the measurement using observers or Kalman filters and using the residuals for fault detection (Mehra and Peschon, 1971; Welch and Bishop, 1995; Simani and Fantuzzi, 2006). The purpose is to detect the presence of a fault and to locate the occurrence time. The estimations are compared to the normal variable values and so, deviations are interpreted as faults. Next, the problem is similar to a pattern recognition problem. This paper is organized as follows. The first part of this communication presents the proposed model-based methodology. Next, its implementation is underlined and the main fundamental concepts of the simulation library *PrODHyS* are described. This is followed by a presentation of a modelling

within *PrODHyS*. These concepts and so our approach are exploited through the simulation of the monitoring of a didactic example. Finally, section 6 summarizes the contributions and achievements of the paper and some future research works are suggested.

2. *PrODHyS* Environment

The research works performed for several years within the PSE research department (LGC) on process modelling and simulation have led to the development of PrODHyS (Jourda *et al.*, 1996; Sargousse, 1999; Perret *et al.*, 2004; Olivier-Maget *et al.*, 2008). This environment provides a library of classes dedicated to the dynamic hybrid simulation of processes. Based on object concepts, PrODHyS offers extensible and reusable software components allowing a rigorous and systematic modeling of processes. The primal contribution of these works consisted in determining and designing the foundation buildings classes.

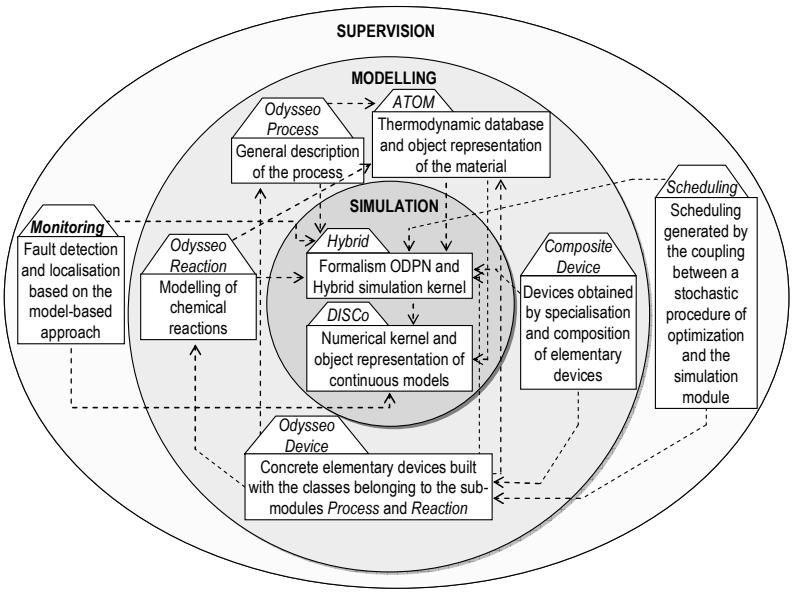


Figure 1. *PrODHyS* Environment

Currently, this library is made up of more than one thousand classes distributed into three independent functional layers (simulation / modelling /supervision) and nine modules (Figure 1). Developed during these research works, the Monitoring module (*PrODHySAEM*) manages the studies of the process monitoring. It is based on our methodology. Thus, each fundamental element of our approach is described by an object class. Moreover the high sequential aspect of the considered systems justifies the use of Petri nets model. This is why the Object Differential Petri Nets (*ODPN*) formalism is used to describe the simulation model associated with each component. A detailed description of this formalism can be found in (Perret *et al.*, 2004; Olivier-Maget, 2007).

3. Supervision module

For this purpose, the simulation model of *PrODHyS* is used as a reference model to implement the functions of detection and diagnosis. The global principle of this system is shown in Figure 1, where the sequence of the different operations is underlined.

Moreover, a distinction between the on-line and off-line operations is made. Our approach is composed of three parts: the generation of the residuals, the generation of the signatures and the generation of the fault indicators.

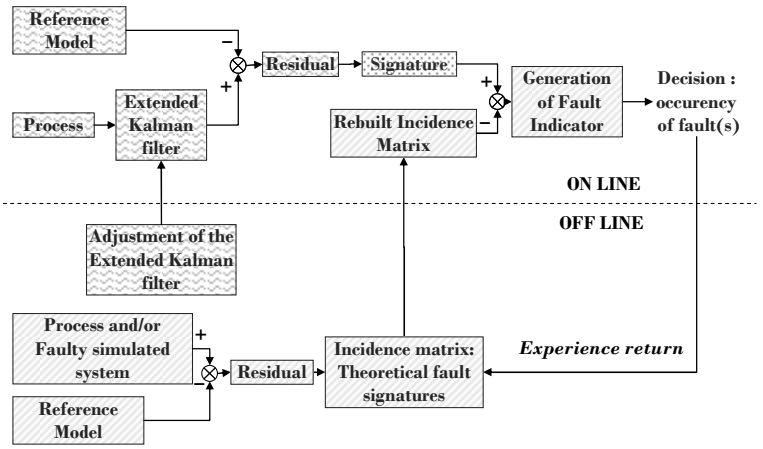


Figure 2. Supervision Architecture

3.1. Generation of the residuals

The first part concerns the generation of the residuals (waved pattern in the Figure 2). In order to obtain an observer of the physical system, a real-time simulation is done in parallel. So, a complete state of the system will be available at any time. Thus, it is based on the comparison between the predicted behavior obtained thanks to the simulation of the reference model (values of state variables) and the real observed behavior (measurements from the process correlated thanks to the Extended Kalman Filter). A description of the extended Kalman filter can be found in (Olivier-Maget *et al.*, 2007). Besides the residual is defined according to the following equation:

$$r_i^r(t) = \frac{\hat{X}_i(t) - X_i(t)}{X_i(t)} \quad with \quad i \in \{1, n\} \qquad \text{(Eqn. 1.)}$$

where X_i is the state variable, $\hat{X}_i$ is the estimated state variable with the extended Kalman Filter and n is the number of state variables. Note that the generated residual $r_i^r(t)$ is relative. As a matter of fact, this allows the comparison of a residual of a variable with a residual of an other one, since the residual become independent of the physical size of the variable.

3.2. Generation of the signatures

The second part is the generation of the signatures (doted pattern in the Figure 2). This is the detection stage. It determinates the presence or not of a fault. This is made by a simple threshold $\varepsilon_i(t)$. The generated structure $S_i^{rN}(t)$ is denoted by the following equation 2.

$$S_i^{rN}(t) = \frac{Max\left[\left(\left|r_i^r(t)\right| - \varepsilon'_i(t)\right); 0\right]}{\sum_{k=1}^{n} Max\left[\left(\left|r_k^r(t)\right| - \varepsilon'_k(t)\right); 0\right]} \quad with \quad i \in \{1, n\} \qquad \text{(Eqn. 2.)}$$

with $\varepsilon'_i(t) = \frac{\varepsilon_i(t)}{X_i(t)}$, where ε_i is the detection threshold. The value of ε_i is chosen according to the model error covariance matrix of the Extended Kalman Filter.

3.3. Generation of the fault indicators

The last part deals with the diagnosis of the fault (hatched pattern in the Figure 2). The signature obtained in the previous part is compared with the theoretical fault signatures by means of distance. A theoretical signature $T_{\bullet j}$ of a particular fault j is obtained by experience or in our case, by simulations of the process with different occurency dates of this fault. Then, a fault indicator is generated. For this, we define two distances: the relative Manhattan distance and the improved Manhattan distance. The first distance is denoted by the following expression 3 and the second distance, which allows the diagnosis of many simultaneous faults, is denoted by the expression 4:

$$D_j^{Mr}(t) = \frac{\sum_{i=1}^{n} \left| S_i^{rN}(t) - T_{ij} \right|}{n} \quad \text{(Eqn. 3.)} \qquad D_j^{Ma}(t) = \frac{\sum_{i=1}^{n} \left| S_i^{rN}(t) \times m' - T_{ij} \times n' \right| \cdot T_{ij}}{n'} \quad \text{(Eqn. 4.)}$$

where n' is the number of non-zero elements of the theoretical fault signature $T_{\bullet j}$ and m' is the number of non-zero elements of the fault signature $S^{rN}(t)$.

4. Application

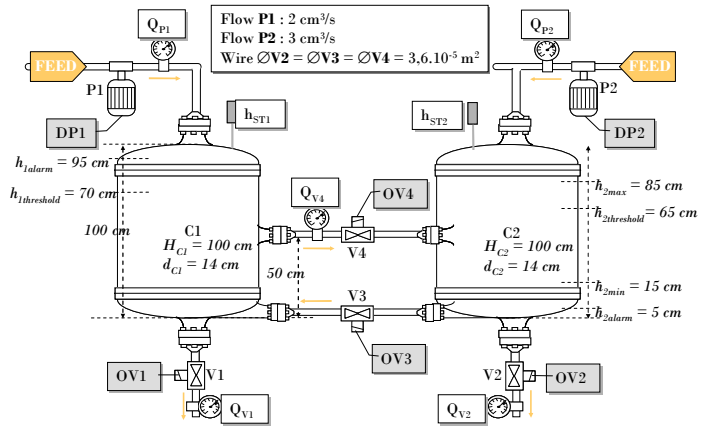


Figure 3. Flowsheet of the benchmark

The considered hydraulic system (Figure 3) is inspired by a benchmark defined by the *AS193* "Diagnosis of the hybrid systems" (cf. www.univ-lille1.fr/lail/AS193/).This system consists of two cylindrical tanks **C1** and **C2**, connected by two pipes with "on/off" valves **V3** and **V4**. The feed of the tanks is maintained by the "on/off" pumps **P1** and **P2**. The tank **C2** can be drained through the "on/off" valve **V2**. The instrumentation of the process is composed of 5 flow sensors and 2 level sensors. The goal of the control device consists in maintaining the liquid level h_2 in **C2** between the heights h_{2min} and h_{2max} by controlling the valve **V4**. The valve **V3** is opened only when the level in **C2** is such $h_2 \leq h_{2alarm}$. The Petri net associated with the control level is presented on figure 4.

4.1. Results

The monitoring of this process is simulated thanks to *PrODHySAEM*. This process is a system based on hydraulic phenomena.

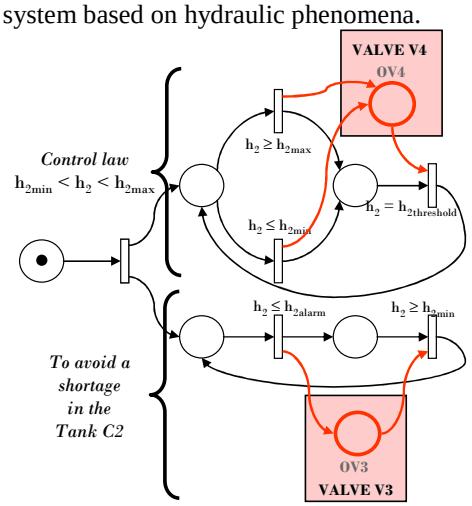


Figure 4. Control Petri net

Two simultaneous faults are introduced: a sensor fault (level sensor of the tank C1, fault 1) and an actuator fault (valve V1, fault 2).

4.1.1. Detection results

We remind that the thresholds for the detection correspond to the model uncertainties obtained by the adjustment of the Extended Kalman filter. Both faults are introduced at t = 1550 seconds. The valve V1 remains opened (fault 1) and the level sensor of the valve V1 returns none value (fault 2).

Figure 5 shows the detection stage. It illustrates the evolution of the residual linked to the liquid level in the tank C1.

From t = 1570 seconds, the residual value of underlines the abnormal behaviour of the process. The diagnosis is launched at t = 1580 seconds.

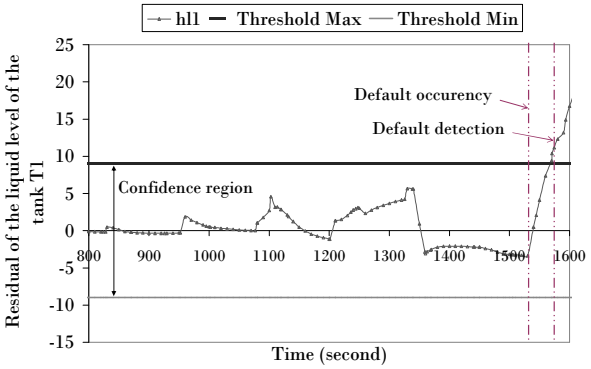


Figure 3. The evolutions of the composition residuals during the evaporation stage

Signature	Liquid level h_1	0.01788
	Liquid level h_2	0
	Flow rate in V1	0
	Flow rate in V2	0
	Flow rate in V3	0
	Flow rate in V4	0.98212
	Flow rate in P1	
	Flow rate in P2	0

Table 3. The instantaneous fault signatures

4.1.2. Diagnosis results

The residual is then estimated and we obtain the corresponding instantaneous fault signature (Table 1). Notice that the exploited signature in this approach is non binary, in order to quantify the deviation due to the fault. The construction of the theoretical fault signatures is based on numerous simulations, in which one of the faults exposed in the Table 2 is generated.

We compare the instantaneous fault signature (Table 1) with the theoretical fault signatures, by calculating the relative and improved Manhattan distances (Eqn. 3. and 4.). Then, the fault indicators are generated (Table 2). They correspond to the complement to 1 of these distances.

Fault	Description	Manhattan relative indicator	Manhattan improved indicator
1	**The level sensor of the tank C1 returns none value.**	**0.99676**	**0.99676**
2	**The valve V1 remains opened.**	**0.99954**	**0.99954**
3	The valve V1 remains closed.	0.76170	0.76190
4	The valve V4 remains opened.	0.75447	0.75551
5	The valve V4 remains closed.	0.86747	0.86958
6	The valve V2 remains opened.	0.81459	0.81575
7	The valve V2 remains closed.	0.78329	0.78389
8	The pump P1 provides continuously material.	0.99553	0.87947
9	The pump P1 no more provides any material.	0.77239	0.77279
10	The pump P2 provides continuously material.	0.77444	0.71020
11	The level sensor of the tank C2 returns none value.	0.75447	0.87716

Table 3. The fault indicators of the example

The relative Manhattan indicator points out three faults: the faults 1, 2 and 8 whose indicators are higher than 99%. Nevertheless, any fault is discriminated, since their indicators are higher than 0.68. 0.68 is the fixed criterion, which corresponds to the probability at the standard deviation according to the normal distribution. In the opposite, the results of the improved Manhattan indicator underline the faults 1 and 2 whose are still higher than 99%. Nevertheless, we can not be sure of the presence of either or the both faults. Notice that the simulation of these three scenarios removes the ambiguity. So, this simulation underlines the ability of our monitoring system to detect the simultaneous faults.

5. Conclusion

In this work, the feasibility of using the simulation as a tool for fault detection and diagnosis is demonstrated. Our methodology is based on the hybrid dynamic simulator *PrODHyS*. The fault detection and diagnosis approach, developed here, is a general method for the detection and isolation of the occurency of a fault. Besides, this approach allows the detection of numerous types of fault and has the ability to underline the simultaneous occurency of many faults. The works in progress aim at integrating this simulation model within a model-based supervision system. The goal is to define a recovery solution following the diagnosis of a fault. For this, we exploit the results of signatures in order to generate qualitative information. For example, with these results, we have the ability to distinguish a simple degradation and a failure. Next, we combine our diagnosis approach with an other method, such as classification or case-based reasoning.

References

L. Jourda, X. Joulia and B. Koehret, 1996, Comp & Chem Engineering, 20A, pp. S157-S164

R.K. Mehra and J. Peschon, 1971, Automatica, 5, pp. 637-640

N. Olivier-Maget, 2007, PhD Thesis of the Toulouse University (INSA)

N. Olivier-Maget, G. Hétreux, J.M. Le Lann and M.V. Le Lann, 2008, Chem Eng and Proc, 47(11), pp.1942-1952

J. Perret, G. Hétreux and J.M. LeLann, 2004, Cont Engineering Practice, 12(10), pp. 1211-1223

A. Sargousse, 1999, PhD Thesis of INPT, France

S. Simani and C. Fantuzzi, 2006, Mechatronics, 16, pp. 341-363

V. Venkatasubramanian, R. Rengaswamy, K. Yin and S. N. Kavuri, 2003, Computers & Chemical Engineering, 27, pp. 293-346

G. Welch and G. Bishop, 1995, Technical Report TR 95-041, University of North Carolina

J. Zaytoon, 2001, Systèmes dynamiques hybrides, Hermès Sciences publications

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal Control of Heat Exchanger Networks

Luís F. Novazzi,[a] Roger J. Zemp[b]

[a]*Department of Chemical Engineering, Centro Universitário da FEI, Av. Humberto Castelo Branco 3972, São Bernardo do Campo CEP 09850-901, Brazil*
[b]*Faculty of Chemical Engineering, Universidade Estadual de Campinas, Av. Albert Einstein 500, Campinas CEP 13083-852, Brazil*

Abstract

Energy integration among chemical process streams can lead to quite complex heat exchanger networks (HEN) with difficulties in terms of control. In the HEN, the control system primary objective is to keep outlet stream temperatures in a specified range. This objective can be achieved by manipulating bypasses in the exchangers or heat loads in coolers or heaters, which results in a positive degree of freedom, since the number of possible manipulated variables is greater than the number of controlled variables. Thus, a secondary control objective can be set: minimization of utility consumption. In this work HEN dynamics and steady state aspects were addressed with the aim of proposing a control strategy that minimizes utility consumption and satisfies imposed restrictions. With the tools developed in this work it was proposed a feedforward optimal control strategy, which consisted in the minimization of a steady state objective function, connected to utility consumption and subjected to constraints in outlet stream temperatures. The suggested approach is an interesting one since rigid or flexible control objectives can be set. The solution of the minimization problem resulted in optimal bypasses positions in the HEN, which were dynamically implemented using a ramp function and a step function. It was verified that ramp implementation was better, with tolerable dynamic violations.

Keywords: Heat Exchanger Networks, Optimal Control, Dynamics

1. Introduction

Due to the continuous oil price rise since the seventies and to environmental issues, efficient use of energy in chemical processes is very important. Almost thirty years ago the theoretical foundations of Process Integration for the efficient use of energy were established by Linhoff et al. (1982), when they proposed Pinch Technology, an elegant approach to set energy / cost targets for HEN as well as rules to design such networks. Nowadays these synthesis techniques, including some which are based on mathematical programming, are well established in Process Design and are easily found in many Chemical Engineering textbooks.

The design of a HEN depends on nominal streams supply temperatures and flowrates. However, during plant operation such nominal conditions can change, influencing stream target temperatures and propagating in the network. With the aim of studying such propagation, a dynamic mathematical model of the HEN is required. Some dynamic HEN models were suggested in the literature (Mathisen, 1994) and they were based on the resolution of a partial differential equation system.

Since ideally the disturbances must be restrained within the HEN, its control system has as a primary objective to drive stream target temperatures to the specified nominal

conditions. In this regard, it is possible to manipulate bypasses positions in the heat exchangers or to manipulate heat loads in heaters / coolers. There is a positive degree of freedom in doing so since the number of manipulated variables is greater than the number of controlled variables. Due to this fact, a secondary control objective can be set: optimization of utility consumption.

In the paper presented by Boyaci et al. (1996) a control strategy based on the optimization of a steady state objective function satisfying the primary control objective was suggested, but the secondary control objective was not addressed. Glemmestad et al. (1999) presented an alternative approach to the optimal operation of HEN systems based on on-line optimization of a steady state function and a fixed control structure, chosen offline. Lately, Giovanini and Marchetti (2003) have shown that low-level Distributed Control System is also capable of handling HEN control problems when a flexible control loop structure is provided.

Model predictive control (MPC) is an advanced method of process control based on a dynamic model of the plant. In this approach the future moves of manipulated variables depend on the model and on plant measured output variables, in such a way that an on-line constrained optimization is performed. The control problem in the HEN was also considered in the more recent work of González et al. (2006), where MPC in connection to optimization of a steady state HEN model were used, in a two level control structure.

As a simpler way to control HEN when compared to MPC methods, in the present work an optimal feedforward control strategy is proposed, where the HEN utility consumption is minimized, satisfying the secondary control objective, and the primary control objective is written as a restriction in the optimization formulation.

2. Optimal Control in HEN

Quite as important as efficient use of energy in a HEN is the accomplishment of the primary objective control, i.e., stream target temperatures must be in the specified range. In a HEN the variables can be classified as follows: stream supply temperatures T_i^s and stream flowrates m_i are disturbances; bypasses positions f_j and heat loads in coolers or heaters are manipulated variables and stream target temperatures T_i^t are controlled variables. As a consequence of the positive degree of freedom in the HEN, there is a set of different bypasses positions and of heat loads that satisfies the specified target temperatures. Thus, optimal control can be applied to the problem.

2.1. Optimal Control Strategy satisfying the Primary Control Objective

Boyaci et al. (1996) proposed an optimal control strategy, which consisted in the minimization of a steady state objective function, defined by $\sum\left(\delta T_i^t\right)^2$, where the operator δ indicates the difference between a variable and its nominal value. This optimization problem can be written as:

$$\min_{\mathbf{f}}\left(\delta\mathbf{T}^t\right)'\left(\delta\mathbf{T}^t\right) \text{ subjected to } \mathbf{0} \le \mathbf{f} \le \mathbf{1} \tag{1}$$

where $\mathbf{f}$ is a vector containing bypasses positions, $\mathbf{T}^t$ is a vector composed of target temperatures and the operator ′ indicates transposition of a vector or matrix. The restriction imposed in Equation 1 means that the bypasses could be totally closed ($f_j = 0$), fully open ($f_j = 1$) or could be positioned between these extremes. A schematic representation of the methodology is indicated in Figure 1. According to the scheme, disturbance variables $\mathbf{m}$ (flowrates) and $\mathbf{T}^s$ (supply temperatures) are monitored in discrete time intervals and a steady state optimization is performed.

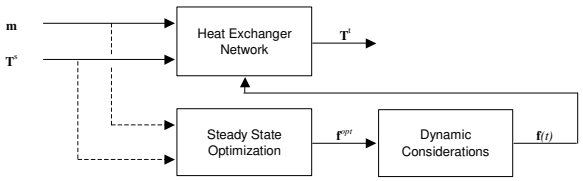


Figure 1 Scheme of the feedforward optimal control strategy

Since the optimization is based on a steady state model, it is necessary to define how bypasses positions change in time. Thus, it was suggested a dynamic implementation (Boyaci et al., 1996) in the form of a step function (Equation 2) and also in the form of a ramp function (Equation 3). In Equations 2 and 3 the bar above variables refers to nominal condition, τ_r is the ramp time duration and t is time.

$$f_i = \begin{cases} \overline{f}_i & \text{for} \quad t < 0 \\ f_i^{opt} & \text{for} \quad t \geq 0 \end{cases} \tag{2}$$

$$f_i = \begin{cases} \overline{f}_i + \left(f_i^{opt} - \overline{f}_i\right)(t/\tau_r) & \text{for} \quad t < \tau_r \\ f_i^{opt} & \text{for} \quad t \geq \tau_r \end{cases} \tag{3}$$

2.2. Proposed Control Strategy

Although the suggested strategy previously described is an interesting one since stream target temperatures are driven to their setpoints, utility minimization is not a concern. In this section an alternative approach is proposed where both primary and secondary control objectives are met.

In a HEN the heaters / coolers are typically positioned in a stream extremity. If δT_i^t assumed a positive value in a hot stream extremity containing a cooler, a higher load in the cooler would be necessary to drive δT_i^t to zero. The same idea applies when a heater is positioned in the extremity of a cold stream. This observation suggests that the optimization control problem can be written to include both primary and secondary control objectives:

$$\min_{\mathbf{f}} \mathbf{p}'\left(\delta \mathbf{T}^t\right) \text{ subjected to } \mathbf{0} \leq \mathbf{f} \leq \mathbf{1} \text{ and } \mathbf{L} \leq \delta \mathbf{T}^t \leq \mathbf{U} \tag{4}$$

where $\mathbf{p}$ is a vector of the same size as $\delta \mathbf{T}^t$ and each $\mathbf{p}$ element is related to a target temperature δT_i^t. The primary control objective, i.e., satisfying temperature setpoints, is represented by $\mathbf{L} \leq \delta \mathbf{T}^t \leq \mathbf{U}$ whereas the minimization function deals with the secondary control objective, i.e., minimizing utility consumption.

Let CP_H and CP_C be the heat capacity for a hot and a cold stream, respectively. The elements of vector $\mathbf{p}$ are filled in as follows: for a cooler in the extremity of a hot stream the element is set to CP_H; for a heater in the extremity of a cold stream the element is set to $-CP_C$ and in the case of absence of a heater or a cooler the element value is set to zero. As a result, the objective function decreases as less hot and cold utility are required. The optimization of the objective function $\mathbf{p}'\left(\delta \mathbf{T}^t\right)$ is subjected to $\mathbf{L} \leq \delta \mathbf{T}^t \leq \mathbf{U}$, where $\mathbf{L}$ and $\mathbf{U}$ are vectors with lower and upper limits for tolerated target temperatures deviations. When there is a rigid control objective, with no tolerable deviations, L_i and U_i are set to zero. On the other hand, when there is a flexible control objective, with relaxed setpoints, L_i and U_i are greater than zero.

It is important to remark that the optimization proposed in Equation 4, which leads to bypasses optimal positions $\mathbf{f}^{opt}$, involves a non linear programming (NLP) problem, with restrictions. The dynamic implementation of $\mathbf{f}^{opt}$ was studied according to Equations 2 and 3, as well as the influence of time τ_r, a very sensitive parameter in the control strategy used in this work.
In terms of implementation, the strategy runs as follows: possible deviations in plant input conditions, like stream supply temperatures and stream flowrates, are monitored and sent to a steady state optimization, based on Equation 4. When this equation is solved and bypasses optimal positions $\mathbf{f}^{opt}$ are found, the signal is sent to the plant and the bypasses will move to their ideal positions. When compared to MPC methods, the proposed strategy herein for HEN control is interesting since it can act as a supervisory system: optimal bypasses positions are fed to local feedback controllers, in a decentralized way.

3. Case Study

With the aim of illustrating the optimal control strategy suggested, it is taken as an example the HEN published by Heggs and Vizcaino (2002), composed by five streams, one heater (H) and four process heat exchangers (HE1, HE2, HE3 and HE4), as represented in Figure 2. In the HEN originally published there was not the presence of bypasses in the exchangers. However, in this work it is supposed that they exist and are positioned in 0,25. As a result, the target temperature vector is equal to: $\mathbf{T}^t = [111{,}3 \quad 133{,}3 \quad 177{,}8 \quad 175{,}5 \quad 204{,}5]'$. Stream properties and heat exchanger characteristics of this HEN can be found in Heggs and Vizcaino (2002).

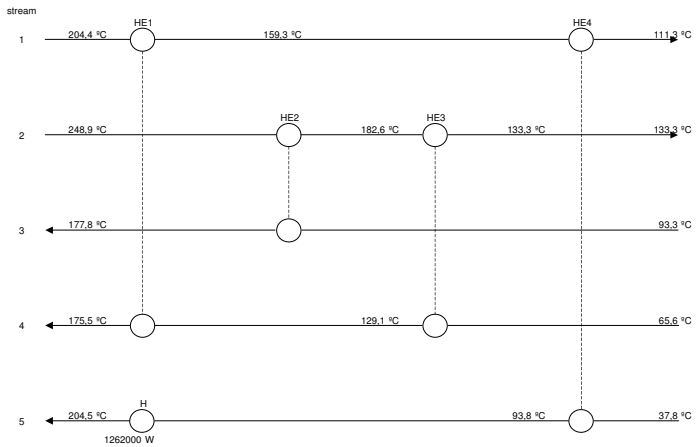


Figure 2 HEN example in grid diagram format

If there is a disturbance in stream 2 supply temperature or flowrate, it will affect the entire network: the heat load in HE2 will change, disturbing stream 3; the heat load in HE3 also will change, interfering in stream 4 and in HE1. Finally the disturbance in HE1 affects HE4, stream 1 and stream 5. To exemplify, let the disturbance be in stream 2 supply temperature, varying from 248,9°C to 258,9ºC, i.e., a δT_2^s equal to 10ºC, in the instant 200 s. The proposed approach was applied and it was considered a rigid control objective, with $\delta \mathbf{T}^t = \mathbf{0}$. By solving Equation 4, with $\mathbf{p} = [0 \quad 0 \quad 0 \quad 0 \quad -11400]'$, one gets the following optimal bypasses positions:

$\mathbf{f}^{opt} = [0{,}4132 \quad 0{,}2975 \quad 0{,}3540 \quad 0{,}2922 \quad 0{,}2131 \quad 0{,}0427 \quad 0{,}0875 \quad 0{,}0000]'$

The constrained optimization was performed in Matlab and a steady state model of the HEN was used (Novazzi, 2007). The first and second elements of $\mathbf{f}^{opt}$ represent bypasses positions of hot and cold sides of HE1, respectively, and so on until HE4. The dynamic implementation of $\mathbf{f}^{opt}$ was firstly studied by using a step function (Equation 2) as soon as the disturbance was detected, in 200 s. The dynamic responses are indicated in Figure 3 and they were obtained in Matlab by solving a system of partial differential equations (Novazzi, 2007).

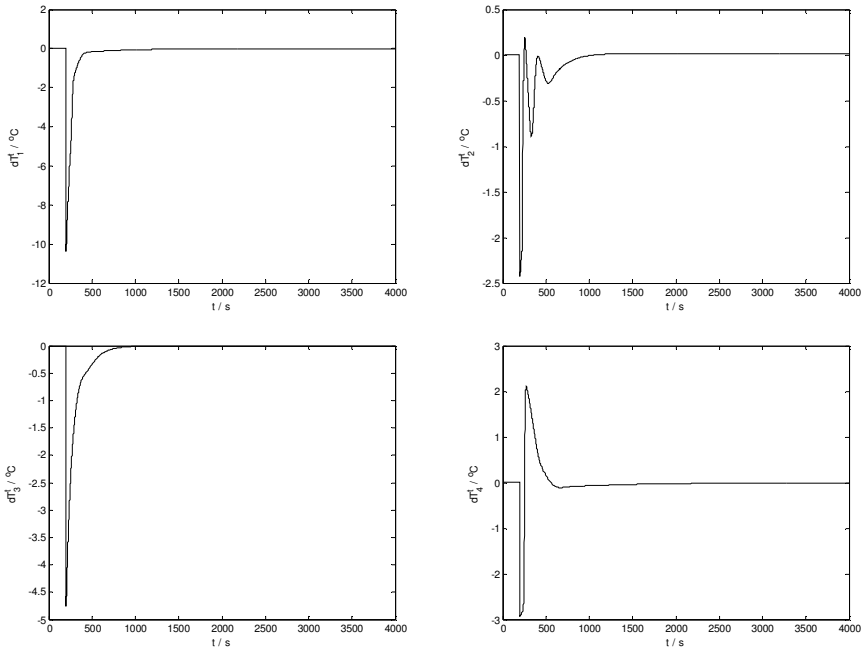


Figure 3 HEN dynamic responses for bypasses opening in a step function

The target temperatures $\delta\mathbf{T}^t$ oscillate before they eventually settle to the former steady state, where $\delta\mathbf{T}^t = \mathbf{0}$, showing that the primary control objective was achieved. Besides, the secondary control objective was also accomplished, since hot utility consumption was reduced from 1,262 kW to 1,096 kW. However, during the transient response severe violations in $\delta\mathbf{T}^t$ happened, mainly in stream 1, with deviations that surpassed 10ºC. Stream 5 was not included in the analysis since it was supposed that any disturbance could be easily met with hot utility in heater H. In spite of the good results in steady state, the responses were very poor dynamically, with severe violations of the restriction $\delta\mathbf{T}^t = \mathbf{0}$.

As a second trial, the bypasses were opened according to a ramp function (Equation 3), with different durations τ_r. The new dynamic responses are indicated in Figure 4, where the solid line represents a τ_r of 500 s, the discontinuous line a τ_r equal to 1000 s and the dashed line a τ_r of 2000 s. The HEN dynamic responses for bypasses opening in a ramp function are clearly superior when compared to bypasses opening in a step function. In stream 1, for example, the dynamic violation of δT_1^t was reduced to only 1,3ºC for a τ_r of 2000 s. It is a significant reduction in the dynamic violation since previously it was greater than 10ºC (Figure 3).

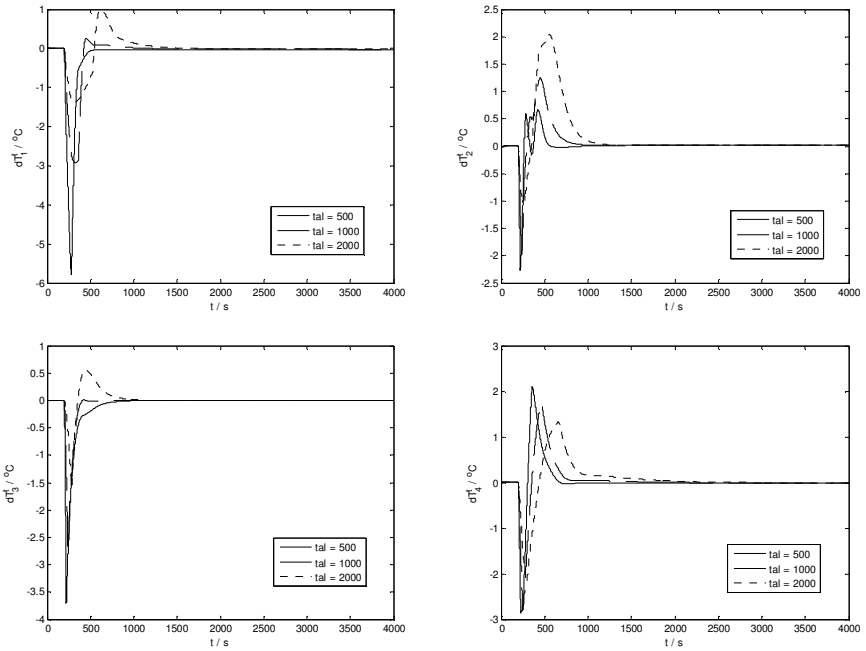


Figure 4 HEN dynamic responses for bypasses opening in a ramp function

4. Conclusions

In this work a feedforward optimal control strategy was suggested and applied in HEN. It consisted in the minimization of utility consumption through bypass manipulation and at the same time stream target temperatures were kept at their setpoints, in a constrained optimization formulation (Equation 4). The proposed approach was successfully applied to a HEN with five streams and the primary and secondary control objectives were met. Dynamic implementation of the optimum bypasses positions showed that step functions led to poor responses, with severe dynamic violations (Figure 3) whereas the bypass opening in a ramp function led to quite satisfactory results (Figure 4).

References

C. Boyaci, D. Uzturk, A.E.S. Konukman, U. Akman, 1996, Dynamics and Optimal Control of Flexible Heat Exchanger Networks, *Computers and Chemical Engineering*, v.20, p.S775-S780

L.L. Giovanini, J.L. Marchetti, 2003, Low-level Flexible Structure Control Applied to Heat Exchanger Networks, *Computers and Chemical Engineering*, v.27, p.1129-1142

B. Glemmestad, S. Skogestad, T. Gundersen, 1999, Optimal Operation of Heat Exchanger Networks, *Computers and Chemical Engineering*, v.23, p.509-522

A.H. González, D. Odloak, J.L. Marchetti, 2006, Predictive Control Applied to Heat Exchanger Networks, *Chemical Engineering and Processing*, v.45, p.661-671

P.J. Heggs, F. Vizcaino, 2002, A Rigorous Model for Evaluation of Disturbance Propagation through Heat Exchanger Networks, *IChemE*, v.80, p.301-308

B. Linnhoff, D.W. Townsend, D. Boland, G.F. Hewitt, B.E.A. Thomas, 1982, *A User Guide on Process Integration for the Efficient Use of Energy*. London: Warwick Printing Company Ltd.

K.W. Mathisen, 1994, *Integrated Design and Control of Heat Exchanger Networks*. University of Trondheim, 195 p., Thesis (D)

L.F. Novazzi, 2007, *Dinâmica e Controle de Redes de Trocadores de Calor*. Universidade Estadual de Campinas, 138 p., Thesis (D)

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Soft-sensor for real-time estimation of ethanol concentration in continuous flash fermentation

Elmer Ccopa Rivera, Daniel Ibraim Pires Atala, Aline Carvalho da Costa, Francisco Maugeri Filho, Rubens Maciel Filho

School of Chemical Engineering, State University of Campinas, P.O. Box 6066, 13081-970, Campinas, SP, Brazil

Abstract

In this work, Artificial Neural Network-based software-sensors (ANN-SS) have been developed for the on-line estimation of ethanol concentration in a continuous flash fermentation process at laboratory scale. The process consists of three interconnected units: fermentor, filter (tangential microfiltration for cell recycling) and vacuum flash vessel (for the continuous separation of ethanol from the broth). The concentration of ethanol in the fermentor and of ethanol condensed from the flash are successfully monitored on-line using ANN-SS calibrated with laboratory experimental data. Results exhibited acceptable predictions for both concentrations of ethanol in the fermentor and condensed ethanol with correlation coefficients of 0.82 and 0.91. The proposed ANN-SS represents an accurate model-based approach which is expected to contribute to improve the implementation of suitable operating strategies of optimization and advanced control to achieve high operation performance.

Keywords: Bioreactor, modeling, ethanol, software sensors, artificial neural network.

1. Introduction

In continuous flash fermentation, the performance of the whole process is significantly influenced by the relationships between process variables in the reaction and separation units. Thus, the availability of an accurate mathematical model is important to improve process performance and to define the most suitable operating conditions to achieve a particular objective. Another important application of an accurate mathematical model is the monitoring of the process states, which is an information source for the decision making in the ethanol production process. In general, monitoring of ethanol profiles is carried out as off-line analysis, often with a significant time delay between sampling and availability of the analysis results.

In industrial fermentation plants analytical instruments are difficult to calibrate, mainly due to the characteristics of industrial culture media, such as turbidity of the culture, presence of dissolved CO_2, among others [1, 2]. Nowadays the software sensor is the preferred alternative for monitoring state variables in biotechnological processes [3].

The so-called software sensors are algorithms for on-line estimation of state variables and model parameters that are not measurable in real-time [4]. This approach is a promising research area with significant impact on industrial operation [5].

More recent works on software sensor for solving biotechnological complex problems are presented by Osorio et al. [2], Lee et al. [6] among others. The major purpose of using software sensors in bioprocesses is to assess the quality of the final product and to

validate on-line analyzers, providing redundant measurements. Artificial Neural Networks (ANN) has been dominant in literature in the field of software sensor design.
Artificial Intelligence techniques such as ANNs have been widely applied for bioprocess modeling, monitoring and control. This technique is sought to efficiently combine all available knowledge and to direct the development toward an improved process operation strategy. Besides, ANNs can be used to offer adaptive solutions, since the reestimation of their parameters is a straightforward procedure [7]. These characteristics are suitable for analyzing data from more complex processes such as the continuous flash fermentation, which have a large number of state variables.
In this work a relatively simple ANN-based software-sensors for monitoring a continuous flash fermentation process is proposed. On-line measurements, such as temperature, dilution rate, pH and turbidity in the fermentor, as well as temperature, pressure and feed flow rate in the flash vessel, are considered as inputs to the software sensor used to infer ethanol concentration in the fermentor and ethanol concentration in the condensed stream from the flash vessel. The ANN-SS is evaluated considering the accuracy with which it describes the experimental dataset.

2. Case study: Continuous flash fermentation

2.1. Process Description
Figure 1 show a schematic diagram of the continuous flash fermentation process investigated in this study. This process consists of three interconnected units: fermentor, cell recycle system (tangential microfiltration) and vacuum flash vessel (ethanol-fermented broth separation unit).

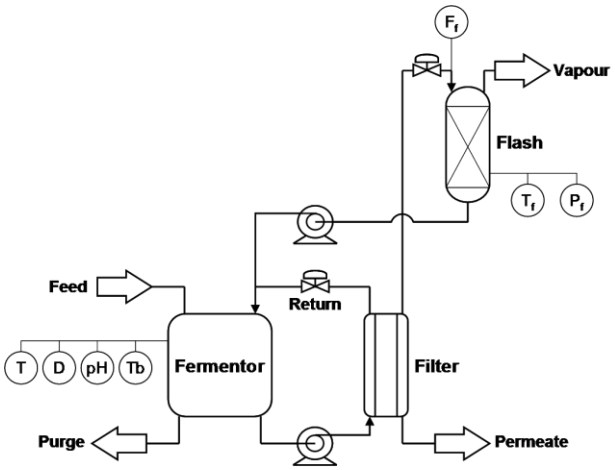


Figure 1. Schematic diagram of the continuous flash fermentation process

The total volume of the system was approximately 5 L. The following on-line variables are taken into consideration in this study: Pressure in the flash, P_f (mmHg); Flash feed flow rate, F_f (L/h); Temperature in the flash, T_f (^oC); Temperature in the fermentor, T (^oC); Dilution rate in the fermentor, D (h^{-1}); pH in the fermentor, pH; Turbidity in the fermentor, Tb (%).
The process is started-up with the addition of the inoculum in a solution of sugar molasses from an industrial plant diluted to 180 g/L of total reducing sugars. It is operated in semi-batch mode until all the substrate is consumed, with addition of diluted molasses only to complete the volume of ethanol evaporated in the flash vessel. The

objective is that biomass reaches a high concentration before the continuous fermentation begins. The end of the start-up stage is monitored by the stabilization of the turbidity as well as of the condensate volume readings. The continuous extractive fermentation is initiated by turning on a peristaltic pump that removes the permeate in the filtration system. The feed flow rate is adjusted to maintain a constant volume in the system. The dilution rate is adjusted to 0.1 h^{-1}, which correspond to a residence time of 10 hours.

After the steady state is attained, the biomass concentration is 30 g/L (cell viability of 95%) when the feed substrate concentration is of 180 g/L. This steady state is maintained during three times the residence time (in this case, about 30 hours). After the end of this stage, the input and output streams are adjusted to provide a dilution rate of 0.15 h^{-1}, which corresponds to a residence time of 6.67 hours, under the same feed substrate concentration and cell volume in the fermentor. This procedure is repeated for the dilution rates of 0.20 and 0.35 h^{-1}, which correspond to residence times of 5 hours and 2.85, respectively. The flash vessel is operated with a feed flow rate of 200 L/h, pressure ranging from 115 to 205 mmHg and temperature varying from 33.8 to 34.5^{o}C, keeping the ethanol concentration in the fermentor between 50 and 65 g/L. The flash vessel works at this temperature to eliminate the need of heat exchangers (which is the most expensive equipment in alcoholic fermentation plants), since the heat of fermentation is removed from the fermentation broth in the flash vessel during the vaporization of part of the ethanol.

The continuous extractive fermentation has shown several advantages, such as low vinasse generation due to the possibility of feeding molasses at higher concentrations, which reduces costs in waste treatment, and the potentiality of eliminating one distillation column from the process. Further detail of technical features can be found elsewhere [8].

3. Artificial neural network-based software sensor

In the present study, two ANNs, one for each output (concentration of ethanol in the fermentor, P_{ferm}, and condensed ethanol from the flash, P_{flash}). Tests in the current study and previous investigations [7] have shown that this approach lead to better results avoiding unnecessarily complex structures.

When ANNs are used to build inferential measurement systems, a study of their architecture is crucial in order to provide a reduction in the dimension of the input space, which can remarkably reduce the time needed for training [5]. Another important issue is to identify among large numbers of possible input variables the ones that are significant for the desired response.

In this work, a Plackett-Burman (PB) design [9] is used to test several combinations of the input variables and their effects on the ANN-SS prediction accuracy. The set of variables was selected using prior knowledge of the continuous flash fermentation process. An array of primary sensors provides seven on-line measurements (see Figure 1): P_f; F_f; T_f; T; D; pH and Tb. The PB design allows to test up to N-1 factors (design variables) in N trials, where N is a multiple of 4. When the number of examined factors is smaller than N-1, the design is completed with dummy variables, which enables the calculation of the standard error for the factors.

In the PB design used in this study, the input variables were varied at two levels; high (presence) and low (absence). A design with 12 trials was used, which leads to 4 degrees of freedom, considering that there are seven input variables. A total of 12 architectures were tested using the input variables corresponding to the PB matrix.

The appropriate number of neurons in the hidden layer of the 12 architectures was found by the cross-validation technique in order to avoid model over-fitting and to achieve good generalization from the training dataset. This technique splits the data sample into a training dataset and a validation dataset. Then MLP neural networks with different numbers of hidden nodes are trained with the training dataset, and their performance are evaluated on its ability to make correct predictions of the validation dataset in terms of mean square error (MSE) (equation 1). The lowest MSE of the validation dataset were used as response, where the corresponding number of hidden neurons is also reported. The effects were calculated using the software Statistica 7.0 (Statsoft).

$$MSE = \sum (d_k - g_k)^2 \qquad (1)$$

In Equation 3, g_k is the prediction of the neural networks and d_k is the desired output, which in this study are P_{ferm} and P_{flash}.

In this study, both input and output data to the ANN were normalized to the range [0.1, 0.9]. Small random values are used to initialization of weights and biases. Subsequently, the standard backpropagation learning algorithm [10], based on a gradient descendent method implemented in FORTRAN is employed to train each network describing ethanol concentrations. Training was stopped after 2000 epochs.

4. Results and Discussion

Two representative datasets containing 4000 input/output patterns were presented in a randomized sequence to the two neural networks, one for each output (P_{ferm} and P_{flash}). These patterns were randomly selected of a database containing 4420 patterns corresponding to approximately 200 hours of continuous flash fermentation at different dilution rates (0.1, 0.15, 0.20 and 0.35 h^{-1}). The validation, i.e., the predictive capability of the neural networks, was assessed on a different sequence of experimental observations, comprising the remaining 420 patterns of the database.

The development of the ANN-SS started with the evaluation of the influence of the considered input variables on the MSE of the desired responses using PB design. Various combinations of a total of seven input variables were tested to build the software-sensors. It was be seen that the ANN-SS that uses the seven inputs led to the lowest MSE values.

Figure 2 shows the effects of the input variables in the accuracy of prediction of the software sensor for the two responses. For the concentration of ethanol in the fermentor, the highest effect was that of the fermentor temperature, followed by the effects of flash temperature and flash pressure. For the concentration of condensed ethanol, the most important effects were that of flash temperature and dilution rate. These results show the strong influence of input variables from the flash vessel in the fermentor and vice-versa.

The performance of the ANN-SS with the lowest MSEs for the P_{ferm} and P_{flash} are shown in Figures 3a and b, respectively. The dilution rates corresponding to the data are shown in dashed lines. It can be seen that the ANN-SS has a good agreement with the experimental validation data. The measured data versus the prediction values of concentrations of ethanol in the fermentor and condensed ethanol concentration are shown in Figures 4a and b. These plots allow visualizing the prediction quality of the software sensors in a clear way. The correlation coefficients (r^2) values represent the fraction of the variance explained in the experimental observation by the ANN-SS.

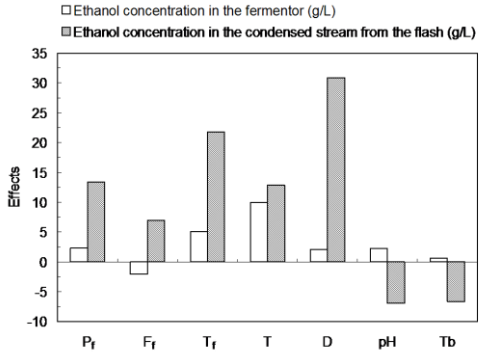


Figure 2. Effect of the input variables on P_{ferm} and P_{flash} from results of Plackett-Burman design

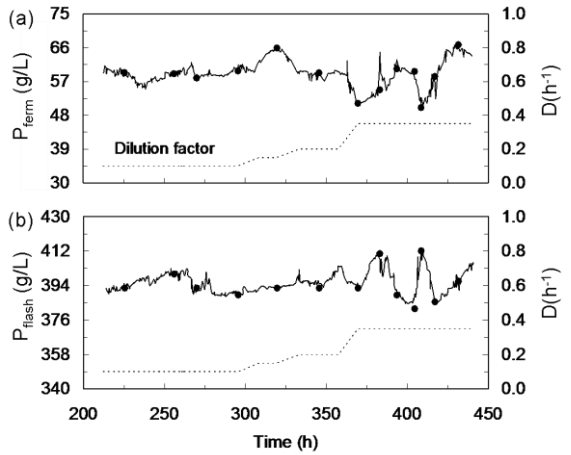


Figure 3. Experimental (solid symbols) and performance of the ANN-Software sensor (continuous lines) using the validation dataset: (a) Ethanol concentration in the fermentor and (b) Ethanol concentration in the condensed stream from the flash. The dilution rates corresponding to the data are shown in dashed lines

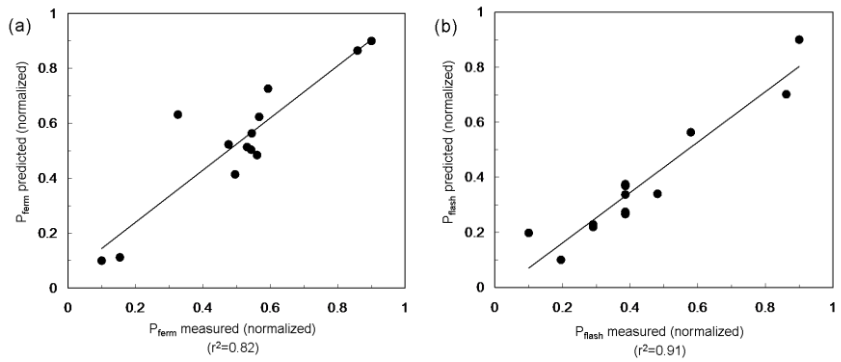


Figure 4. Regression plot from the best ANN-SS architecture during validation: (a) Ethanol concentration in the fermentor and (b) Ethanol concentration in the condensed stream from the flash

5. Concluding Remarks

The development of ANN-based soft-sensors applicable to a continuous flash fermentation process at laboratory scale is addressed in this work. Plackett-Burman design was used to evaluate the influence of various considered input variables in the concentrations of ethanol in the fermentor and of condensed ethanol from the flash vessel. The results have shown that it is possible to accurately infer these concentrations from easily measurable input variables. An important additional advantage of this approach consists in the use of a system of (relatively) cheap sensors such as thermocouples, pH meter and turbidimeter in order to infer the concentrations of ethanol.

The analysis of the effects of the input variables in the concentrations of ethanol in the fermentor and condensed from the flash vessel confirmed close relationships between process variables in the reaction (fermentor) and separation (flash vessel) units. The use of software sensor based on ANNs represents thus an accurate model-based approach which is expected to contribute to improve the implementation of suitable operating strategies of optimization and advanced control to achieve high operation performance.

6. Acknowledgements

The authors acknowledge FAPESP (process number 06/51646-4), and CNPq for financial support.

References

Macedo, 2003, Estado da arte e tendências das tecnologias para energia. Brasília: Ministério da Ciência e da Tecnologia. CTEnerg Secretaria Técnica do Fundo Setorial de Energia. CGEE Centro de gestão e Estudos estratégicos.

D. Osorio, J.R. Pérez-Correa, E. Agosin, M. Cabrera, 2008, Soft-sensor for on-line estimation of ethanol concentrations in wine stills, J. Food Eng., 87, 571-577.

S.I.T.A. Soons, M. Streefland, G. van Straten, A.J.B. van Boxtel, 2008, Assessment of near infrared and "software sensor" for biomass monitoring and control, Chemom. Intell. Lab. Sys., 94, 166-174.

G. Bastin, D. Dochain, 1990, On-line Estimation and Adaptive Control of Bioreactors, Elsevier, Amsterdam.

J.C.B. Gonzaga, L.A.C. Meleiro, C. Kiang, R. Maciel Filho, 2009, ANN-based soft-sensor for real-time process monitoring and control of an industrial polymerization process, Comput. Chem. Eng., 33, 43-49.

M.W. Lee, S.H. Hong, H. Choi, J.H. Kim, D.S. Lee, J.M. Park, 2008, Real-time remote monitoring of small-scaled biological wastewater treatment plants by a multivariate statistical process control and neural network-based software sensors, Process Biochem., 43, 1107-1113.

E.C. Rivera, A.C. Costa, R.R. Andrade, D.I.P. Atala, F. Maugeri Filho and R. Maciel Filho, 2007, Development of adaptive modeling techniques to describe the temperature-dependent kinetics of biotechnological processes, Biochem. Eng. J., 36, 157-166.

D.I.P Atala and F. Maugeri Filho, 2004, Processo fermentativo extrativo a vácuo para produção de etanol, C12P 7/14, C12R 1/865. Brazil, PI 0500321-0.

R.L. Plackett, J.P. Burman, 1946, The design of optimum multifactor experiments, Biometrika, 33:305-325.

C.M. Bishop, 1995, Neural networks for pattern recognition, Oxford University Press, Oxford, UK.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Using a Linear APC Controller on a Non-monotonic Process

Luciana M. Galvão,[a] Leo Lincoln,[b] Chad Segura,[c]

[a] *Luciana M. Galvão, Aspen Technology – Advanced Process Control NALA Professional Services Team, Av. Faria Lima 3729 / 5º, CEP 04538-905, São Paulo - SP, Brazil.*
[b] Leo *Lincoln, Braskem S.A. – UNIB Polo Petroquímico de Camaçari, Rua Eteno s.n., Camaçari – BA, Brazil.*
[c] *Chad Segura, Aspen Technology – Advanced Process Control Global Professional Services, 2500 CityWest Blvd.- CEP 77042- Houston –TX, USA.*

Abstract

This paper demonstrates an application of a linear MPC for the control of a non-monotonic process for which the steady state gain changes the sign depending on the operating conditions. This strategy of using a linear controller on a non-monotonic process is justified if, for instance, there is already a large installed base of linear controllers at the plant site without an existing base of nonlinear controllers and the process belongs to this simple subset of non-monotonic processes. Here, the practical application of this strategy on a cyclopentadiene to di-cyclopentadiene reactor in an isoprene production unit is discussed and the results for this implementation are also presented. The benefit is shown by the reduction in the variability of product and also the reduction of the isoprene losses. In addition, a more stable feed composition to the downstream production unit has been observed. This solution was reached using the Hysys® rigorous steady-state simulator from Aspen Technology, AspenIQ™ for dynamic modeling and the DMCplus® as the linear MPC. Using these tools, it was possible to detect the process non-monotonic behavior and control the dependent variables in a stable manner. This method of using the rigorous model is especially useful when the dead time is big compared to process time constant.

Keywords: non-monotonic process, advanced process control, MPC, cyclopentadiene, Isoprene.

1. Introduction

Nonlinearities are usually faced on Multi-variable Process Control projects. For lower degrees of non-linearity, linear controllers can be adapted to overcome the changes in process gain inherent in nonlinear processes. However, processes which are non-monotonic in nature present a special challenge for APC projects. In a few of these cases, if one can force the process to remain on one side of the zero gain point (point of gain sign inversion), one can usually safely use the linear controller as one has forced the process to remain in a monotonic region. However, if one must operate the process on both sides of the zero gain point such as in processes with uncontrolled disturbances or changes in the controlled variable target, a nonlinear controller is generally advised.

A complete APC (Advanced Process Control) solution has been developed for the Braskem Isoprene Unit. The execution group joins participants from Braskem and Aspentech. The DMCplus® was the primary choice to control all sections of this production site. It would also be required to remain the reaction section with the same control software as the rest of the unit, for keeping better communication between the sub-controllers. The challenge is the reaction section shows a very non-linear behavior in operating window. Due to the process and operating characteristics of the Isoprene reactor, a linear solution for control this system is no longer an option. This issue was solved through the use of simulation (Hysys®), prediction and online (AspenIQ™) tools combined with DMCplus® (MPC Aspen Technology, Inc.). The reaction sub-controller is able to communicate with other DMCplus® sub-controllers and it controls a process where the gains sign change under normal process behavior.

1.1. Process Description

The Braskem Isoprene Separation unit produces Isoprene (99.5%), Dicyclopentediene (DCPD) and Piperylenes. The unit recovers the isoprene (2-methyl-1,3-butadiene) from a feed stream by an extractive distillation process. The feed is C_5 mixed compounds that contents Cyclepentediene (CPD). The product is carried out in a number of process steps. The focus of this work is on the reactor section. The key economic driver for this unit is the isoprene losses reduction.

Since CPD volatility is very close to isoprene's, the separation of these compounds by distillation is practically impossible. By CPD dimerization, it becomes more feasible to purify the isoprene stream and free it of CPD contaminant. The reactor effluent composition has a large effect on the overall process downstream, and affects directly the final product specification. After passing through the reactor, these same reactions keep throughout the unit. Consequently, there is an optimal outlet value different from that of complete conversion of CPD. The optimal value will depend on the overall unit behavior. So, the reactor sub-controller should be able to interact with other unit sub-controllers.

1.2. Reaction

The reaction rates (showed in Figure 1) are affected by feed flow, by the inlet temperature and by CPD concentration in the inlet stream (inlet composition). As the reaction from CPD to DCPD is exothermic, the reactor (plug flow reactor) becomes hotter along the path if the CPD inlet concentration is high. The CPD inlet concentration also affects significantly the reaction rates. If reaction rates are too high, it can reduce isoprene throughput. On the other hand, if there is not enough CPD conversion, then CPD will contaminate the final product (isoprene). The main reaction kinetics model parameters used in the reactor simulation were extracted from literature, and were developed by Seidov (1964) and Muja (1975) at different temperatures ranges.

The reaction rate is function of the concentration and system temperature. It was considered that Arrhenius law can be a good approximation to relate the temperature with the reaction velocity. Because of confidentiality, it was not possible to present the values from F (Frequency Factor) and E (is the activation energy) used in simulation.

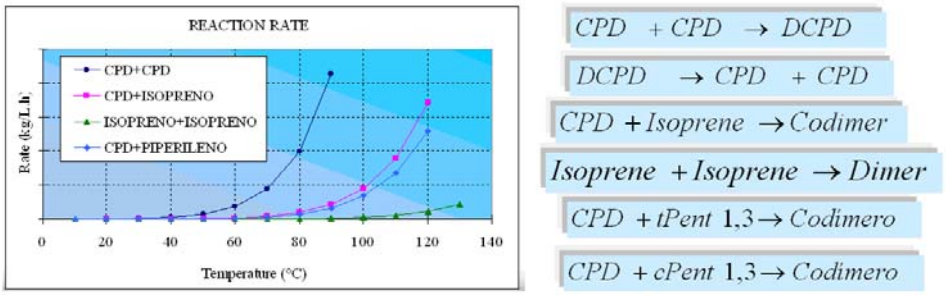


Figure 1: Reactions: homogeneous, in liquid phase, non-catalytic and exothermic. The scale was removed here because of confidential reasons (based on Seidov (1964) and Muja (1975)).

2. Controller Development

2.1. Process Control Challenge and Simulation Development

Figure 2 shows the reactor regulatory control, where the unit's feed stream is heated by a steam heat exchanger. The pressure control will keep the outlet stable and its effect is negligible here, because it is a liquid phase reaction. The only handle for reaction control is the inlet temperature. The inlet CPD concentration is the most relevant disturbance and it affects the temperature along the reactor path. The CPD in the feed changes frequently and abruptly (from 3% to 20%). This disturbance has an inflection point (around 6% CPD), when the steady state gain (disturbance against CPD outlet composition) changes from positive to negative.

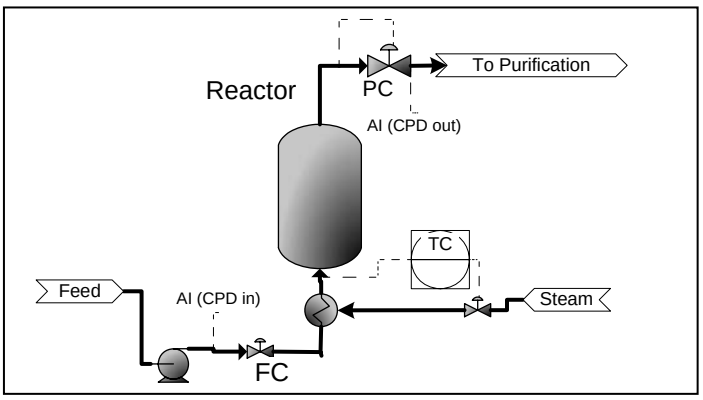


Figure 2 Reactor Regulatory Control,

Figure 3 came from the Hysys® simulation that was developed for this system. The steady-state simulation considered a Plug Flow Reactor, with no catalyst, without axial variation, adiabatic and with exothermic reaction. A number of case studies provided the data for regression. It became finally the online implementation for the control of the CPD outlet concentration. The simulations show that the steady state gain for

temperature (a controller manipulated variable) reverses sign, depending on feed composition (feedforward variable); and also the inflection point in the temperature (where the gain is zero) changes with the feed. It should also be noted that feed rate has an effect on this behavior.

The brown color zone, in Figure 3, should be avoided in process operation because of very high isoprene losses. The green and blue areas are the average conditions where operators try to keep the plant. The blue area is where the isoprene losses are at a minimum and the CPD contamination in the isoprene product gets most close to its limit. In the red area the isoprene will be out of specification. The blue area boundary (the optimum region) depends on downstream capacity and equipment efficiency, so the reactor controller should be able to act properly under various downstream limitations.

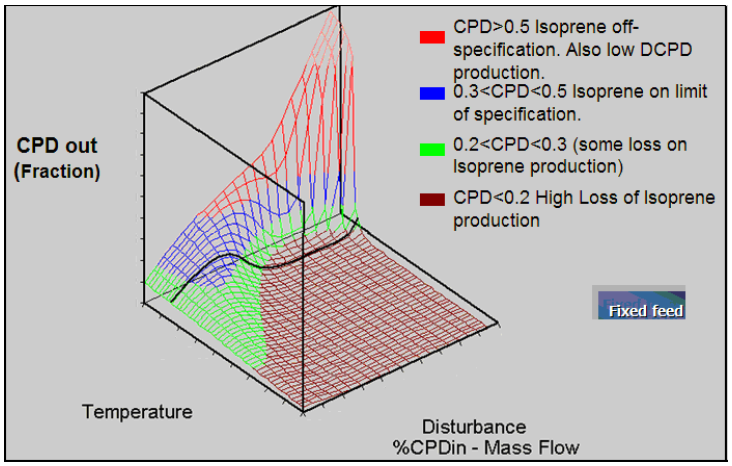


Figure 3. Effect of inlet CPD and inlet temperature on reaction, for various compositions, with fixed flow. Surface obtained by Hysys® Simulation steady state case study.

When CPD inlet concentration is low (below about 6%, for instance), the inlet temperature (control handle) should increase to lower the CPD outlet concentration (increase reaction rate) if CPD in feed rises. However, for high CPD inlet concentration (above 6%, for instance), the opposite action will be required. In other words, with high inlet CPD concentrations, there will be a very high conversion rate (and Isoprene losses as well, because of high reactor temperatures due to the exothermic reaction). In order to compensate for the high feed CPD concentration effect when CPD in the feed is high, temperature control should lower when CPD concentration in the feed rises.

The reactor has a long time to steady state (around 10-15 hours depending on feed rate). In addition, the dead time is approximately 6 hours. A traditional MPC controller with feedback would be virtually unable to control the outlet CPD at a target. The continuous varying controller gains, dead time and sign of the model gains would play havoc with the process.

2.2. Advanced Control Strategy

Figure 4 explains how the simulation was built and how the kinetic model was chosen. In the sequence, the results were checked against real data (at steady state condition)

and then several "Hysys® case studies" generated the data collection. Finally, a regression was done. The simulation data collection was used to build one temperature equation (by polynomial regression) as a function of: feed, CPD inlet concentration (%), and target CPD outlet (%). The calculated temperature is the required temperature to achieve the target CPD outlet, which is read at CPDout-ET spot inside the controller. The equation for CPD outlet prediction is the same polynomial equation used for temperature, but fed by the current temperature. It predicts the value that CPD outlet would achieve at steady state, if the current temperature remains the same.

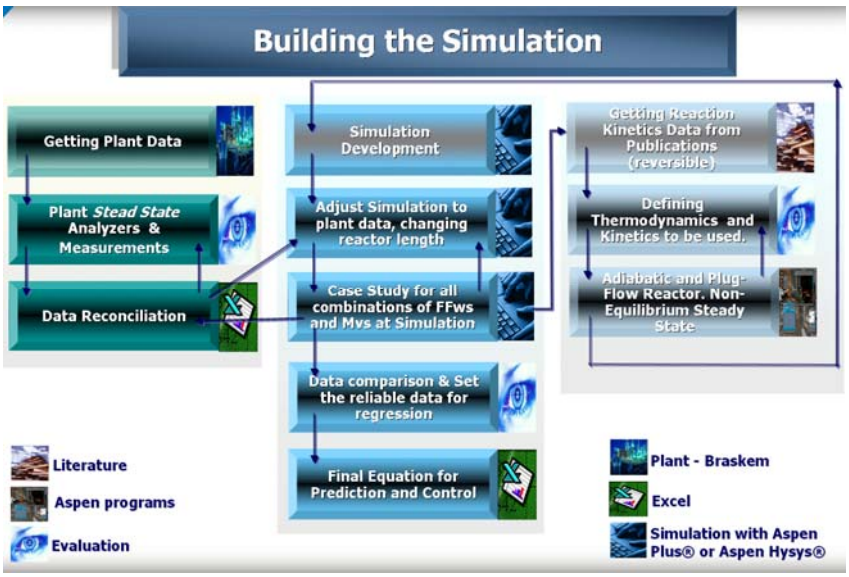


Figure 4: Scheme on how the simulation was developed.

DMCplus® model type is simple model. The controller reads the temperature result, (by IQ™ polynomial calculation) set at the External Target (ET). Then the controller will drive the unit smoothly to that temperature, respecting the other CV limits during dynamics and steady state scenarios. The controller will achieve the temperature-ET and then, at steady state, the CPD current value should be equal to CPDout-ET. The ET is a tool from DMCplus® that corresponds to a set point at the basic control; when driving the unit to the ET, the controller respects the limits and the weighs of other CVs. The CPD outlet ET should be set depending on how the downstream controllers can handle the isoprene losses. In any case, operators are free to specify the CPD outlet ET if required. Inside DMCplus® linear simple model, the CPD outlet variable was configured exclusively to receive the desired value at its ET (by operators or other tool). The CV of CDP outlet is normally OFF and its linear prediction is not used. The appropriate CPD prediction is the polynomial calculation result. The controller will drive to ET temperature, in order to achieve the CPD desired (at steady state).

Figure 5 provides a graphical view of stability rising. The outlet CPD shows a much more stable behavior, following the specified target. The isoprene savings depends on the target value that the operation forces the controller to be.

The model prediction presents a very accurate result when compared to plant data. Most of plant data have the outlet CPD at about 3% or lower. There is not much plant data for

higher values of outlet CPD, (about 5%). The final optimization of isoprene production sometimes requires higher values then 3% (the most usual target). The post controller commissioning stress-test was done. The error (difference between the regression of simulation data and pure simulation curves) average is around 0,14°C. With this new control strategy running, the outlet CPD was safely increased to 4% and some times 4,5%, reducing isoprene losses. Those outlet values were not previously achievable by the plant operators in a stable manner.

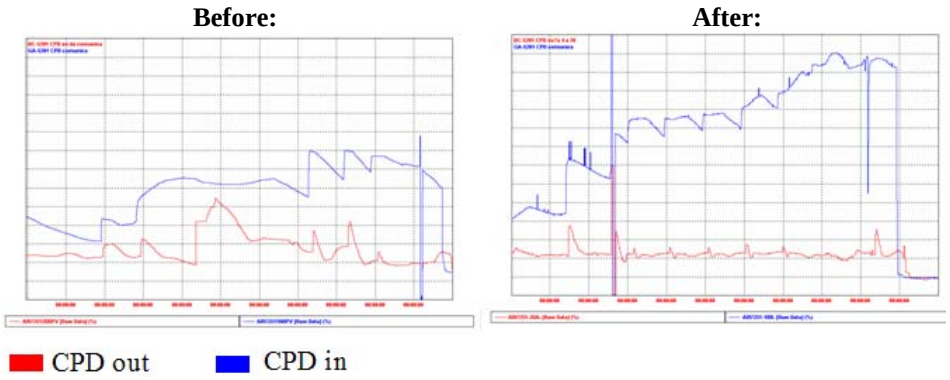


Figure 5: The Control Performance after and before the APC implementation. This is a 10 days comparison at same ranges and same scale. The actual numbers on scale were omitted due to confidentiality.

3. Conclusion

This controller is very stable and reaches or surpasses its original expectations. The controller approaches 100% uptime. It also:

- Reduces Isoprene Losses: it was possible to operate the reactor all time in the target temperature to maximize isoprene recovery.
- Demonstrates that rigorous modeling can help to understand the processes and avoid extra-testing in complex systems.
- Shows that if combined with AspenIQ™ and rigorous modeling, DMCplus® can handle non-linear, non-monotonic control for small subsets of variables.

The approach described in this paper can be applied to other processes when there is:

- A long time to steady state process, with a large dead-time;
- Evidence of gains with sign reversal (small controllers only);
- And there exists a simple regressed equation which can adequately predict the process behavior.

4. References

I. Muja, G. Andreescu, M. Corciovei, R. Fratilonium, 1975, Cinetica y Termodinamica unor Reatti Diels-Alder II. Studiul Cinetic al Reatiei de Dimerizare a ciclopentadienei si al reatiilor de codimerizare cyclopentadiena-izopren Revista de Chimie, 26, Nr 12,USA

Seidov N. M., Karirov R. A., Bakhshizade A. A. Azerb Khlm. Zh., 1964, No, 58, 81-86

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Bypass Design for Control and Optimization of Heat Exchanger Networks

M. Escobar and J.O. Trierweiler
Group of Intensification, Modelling, Simulation, Control and Optimization of Processes
Department of Chemical Engineering, Federal University of Rio Grande do Sul
Porto Alegre, Brazil (e-mail: escobar@enq.ufrgs.br/ jorge@enq.ufrgs.br)

Abstract

Disturbance propagation during a HEN operation may make the control difficult if the HEN is improperly design. Utility flow rates and bypasses are widely used for effective control of process stream target temperatures, but the number of utility units is usually much less than the number of process streams in the network. This paper addresses the optimal bypass design in heat exchanger networks. It consists in a model-based iterative procedure considering a worst-case disturbance rejection with minimum economic penalty. The methodology is demonstrated using a case study with 3 different structures, making possible a comparison between different options on a quantitative basis, taking into account the optimal operation attainable with minimum total annual cost.

Keywords: heat exchanger network control, controllability, bypass design.

1. Introduction

During the last decades, different approaches were proposed to design the control system in order to accommodate setpoint changes and to reject load disturbances in HENs. Mathisen et al. (1992) provided a heuristic method for bypass placement. Papalexandri and Pistikopoulos (1994 a, b) introduced a systematic framework for the synthesis or retrofit of a flexible and a structurally controllable HEN using a MINLP formulation. Aguilera and Marchetti (1998) developed a procedure for on-line optimization and control system design of a HEN also using a MINLP. Yan et al. (2001) proposed a model-based design for the development a retrofit HEN with optimal bypass placement using a simplified model for disturbance propagation and control.

The procedure proposed in this paper is an evolution of the method introduced by Yan et al. (2001). Our approach considers a more rigorous model for the heat exchangers, an algorithm for automatic pairing selection , and additional bypass are included to reduce the utility consume. Two side effects can occur by the bypass HEN control strategy, since usually increases the (i) interaction and (ii) initial investment. This work analyzes the HEN control, bypass design and minimization of utility consumption simultaneously.

2. Operation and Control of HENs

During HEN operation, degrees of freedom or manipulated inputs are needed for control and optimization. In a HEN with n_s streams and n_u utility units, at least $n_u - n_s$ extra

available manipulations must be used to make the operation structurally feasible, where all target temperatures can be controlled independently. We assume in this work that only a single bypass is used and a stream split is not used as a manipulated variable. In order to deal with positive and negative disturbances, the heat exchanger has to be designed with a steady-state flow rate for the bypass stream different than zero.

For a given HEN, a bypass with a specific nominal value u_{nom} can be added without changing the main HEN structure and operating point if the same heat load needs to be maintained. But besides the opportunity to reject disturbances ($i.e., \delta T^t = 0$), its installation must cause an increment of the heat transfer area ($i.e., \delta A/A_0$). A trade-off between disturbance rejection and costs must be considered during the bypass nominal design. Figure 1 illustrates qualitatively how outlet targets temperatures and the increment of area of the bypassed heat exchanger if the same heat load is maintained.

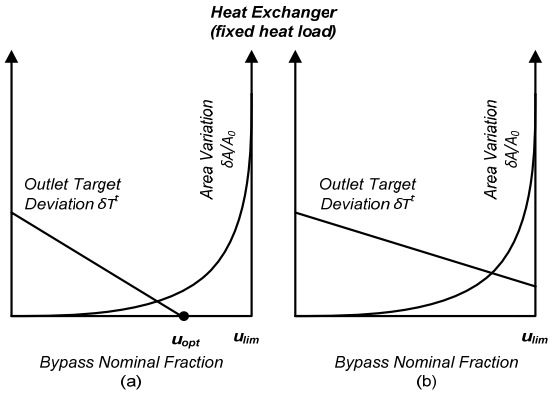


Fig. 1. Qualitative relations between a stream target temperature deviation and the area change as a bypass increases. (a) Complete disturbance rejection when $u_{nom} \geq u_{opt}$; (b) Incomplete disturbance rejection with any bypass fraction.

3. Bypass Design for Control and Optimization

The HEN here is separated into two parts, the inner HEN and the outer HEN according to Glemmestad (1997). While the inner HEN consists of all process exchangers, splitters and mixers, the outer consists of the utility exchangers and target values. The inner HEN contributes with n free variables. For energy optimization the temperatures upstream the utility exchangers, outlets of inner HEN, are to be as close to the stream target temperatures as possible. A model to disturbance propagation and control must characterize the system behaviour under control. It requires a model where the possible control actions are taken into account, represented in matrix form by the equation (1).

$$\delta T^t = G_u \delta u + G_d^t \delta T^s + G_d^w \delta w \tag{1}$$

The vector δT^t correspond to the deviation target temperatures related to the deviations of supply temperatures (δT^S), heat capacity flowrates (δw), and the bypass fraction by the matrices, G_d^t, G_d^w and G_u respectively. The system model for a given configuration is obtained using a heat exchanger model, structural information relating these units and the steady-state data. The unit model is based in the linearization of the model described by the set of equations (2), and (3) in agreement with the representation in Fig. 2.

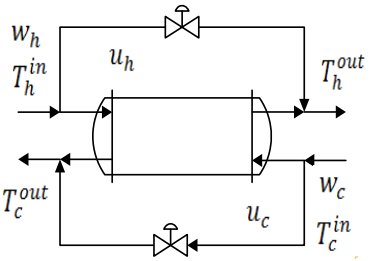


Fig. 2. General structure of a heat exchanger with bypasses.

$$\begin{bmatrix} T_h^{out} \\ T_c^{out} \end{bmatrix} = \begin{bmatrix} \frac{R_h - 1}{R_h - a}(1 - u_h) + u_h & \frac{1 - a}{R_h - a}(1 - u_h) \\ \frac{R_h(1 - a)}{R_h - a}(1 - u_c) & \frac{a(R_h - 1)}{R_h - a}(1 - u_c) + u_c \end{bmatrix} \begin{bmatrix} T_h^{in} \\ T_c^{in} \end{bmatrix} \quad (2)$$

$$R_h = \frac{T_c^o - T_c^{in}}{T_h^{in} - T_h^o} = \frac{w_h(1 - u_h)}{w_c(1 - u_c)} \;\; and \;\; a = exp\left(\frac{UA}{(1 - u_h)w_h}(1 - R_h)\right) \quad (3)$$

For a given HEN, the vectors $\delta T^{s(+)}, \delta T^{s(-)}, \delta w^{(+)}, \delta w^{(-)}, \delta T_{max}^{t(+)}, \delta T_{max}^{t(-)}$ are defined. It is made the assumption of the worst-case design, i.e. the maximum positive and negative deviations of the HEN target temperatures, which occurs at the extreme disturbance values of supply temperatures and heat capacity flow rates if we consider the linear model (eq.1). It is calculated the maximum positive $\delta T_d^{t(+)}$ and negative $\delta T_d^{t(-)}$ deviations of target temperatures considering no control actions (eq. 4 and 5), and it is determined necessary control correction vectors to each case $d^+ = \delta T_{max}^{t(+)} - \delta T_d^{t(+)}$ and $d^- = \delta T_{max}^{t(-)} - \delta T_d^{t(-)}$.

$$\delta T_d^{t(+)} = G_d^{t\,h} \delta T^{s(+)^h} + G_d^{t\,c} \delta T^{s(+)^c} + G_d^{w h} \delta w^{(+)^h} - G_d^{w c} \delta w^{(-)^c} \quad (4)$$

$$\delta T_d^{t(-)} = G_d^{t\,h} \delta T^{s(-)^h} + G_d^{t\,c} \delta T^{s(-)^c} + G_d^{w h} \delta w^{(-)^h} - G_d^{w c} \delta w^{(+)^c} \quad (5)$$

It must be selected the subset of at most n manipulated variables δu_s pairing and the subset of n target temperatures that must be controlled using bypass fraction. The pairing is based on the partial RGA ($RGA^{\dagger}$) defined similarly to the regular RGA substituting the inverse matrix by pseudoinverse of the non square gain matrix G_u (Yang *et al*., 2001). The basic rule is the same as that in regular RGA, where each controlled variable is paired to a manipulated variable such that the corresponding relative gain is positive and as close to 1 as possible. The best combinations are enumerated.

A pair is selected (δu_s) and the necessary control action $\delta {u_s}^{(+)}, \delta {u_s}^{(-)}$, is calculated by solving the optimization problem ($\min \; \|G_u M \delta u_s - d\|_2$) using d^+ and d^- respectively, and the nominal values is updated according to practical limitation $\left(u_{new} = -min\{\delta {u_s}^{(+)}, \delta {u_s}^{(-)}\}\right)$. The convergence criterion is checked ($|\delta u_{new} - \delta u_{old}| \leq \varepsilon$), where ε is the permissible computational error. The model is retrofitted updating the matrices G_u, G_d^t, and G_d^w with the selected nominal bypasses values and all the procedure is repeated until the convergence. At the end the new areas of heat exchangers bypassed are estimated and the model is used to calculate stream outputs deviations and the utility consumption is estimated to the worst case design. The new Total Annual Cost (TAC) is estimated the procedure must be repeated to other possible pair, the best solution is that with the lower TAC.

4. Case Study

In this section 3 different HENs structure depicted in Figure 3, are used to illustrate the design procedure and make a proper selection. The Table 1 lists the design data and the disturbance information need to compute the worst case design.

The degrees of freedom analysis for the $HEN01$, $HEN02$ and $HEN03$ pointed out the number of bypasses that can be used for control and optimization for each HEN ($n = 3$), i.e. the dimension of the space spanned by the bypasses in the target temperature set. A common strategy will use only one bypass to control the temperature target of stream if it is not possible using a utility flowrate. For the $HEN01$ and $HEN02$ only the stream $C1$ must be controlled using a bypass, and two bypasses will be used during operation to minimize the utility consumption. Considering the potential bypass sequence presented in vector $\delta u = \left[u_{1,1}^{h} \;\; u_{1,1}^{c} \;\; u_{1,2}^{h} \;\; u_{1,2}^{c} \;\; u_{2,1}^{h} \;\; u_{2,1}^{c} \;\; u_{2,1}^{h} \;\; u_{2,1}^{c}\right]^{T}$ the upper limit attainable considering the situation of no driving forces (infinity area) for the $HEN01$ is expressed in the vector $u_{lim} = [0.152 \;\; 0.237 \;\; 0.495 \;\; 0.818 \;\; 0.167 \;\; 0.083 \;\; 0.167 \;\; 0.633]^{T}$ if is assumed a $\Delta T_{min} = 10^{0}C$ the vector $u_{lim} = [0 \;\; 0.1 \;\; 0.43 \;\; 0.79 \;\; 0.09 \;\; 0 \;\; 0 \;\; 0.56]$. These limits must be considered during the bypass nominal design, associated with a controllability measure to select the appropriated pairing, i.e. taking into account the economic penalty.

With the matrices G_u and $RGA^{\dagger}$, the potential manipulated variables are sorted by all positive values as close to one as possible creating a priority order. One could think that if we select the prior pair were the best, but this procedure could result in the selection of the same heat exchanger to be bypassed for different controlled outputs, which would not desirable, since there is only one degree of freedom per exchanger. In addition, even if this pair is a possible one, no economical penalty is considered. But the priority order reduces drastically the number of possibilities to be enumerated, and define an appropriated sequence for this enumeration (probably the best pair will be one of the first). The strategy used here consider that keeping the control of the inner HEN This strategy implies that disturbances from the inner to the outer HEN when a utility unit is present are permissible, but to reject disturbances the utility consumption may increase if the disturbance combination results in a near worst case scenario.

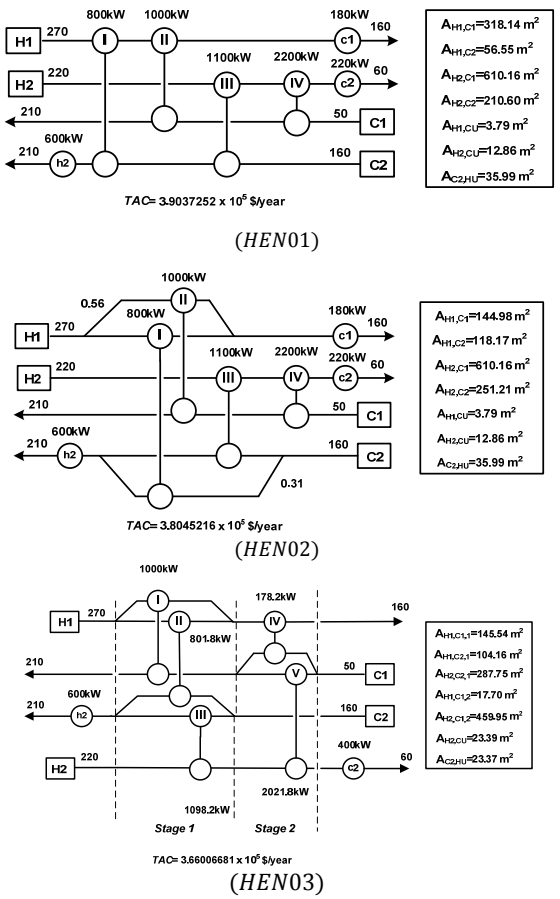


Fig 3. Synthesized HENs for the Case Study using different approaches.

Table 1. Design data for the Case Study.

Stream	T^t (ºC)	T^s (ºC)	F (kWºC^{-1})	h (kW m^2 ºC^{-1})	$\delta T^{(+)}$ (ºC)	$\delta T^{(-)}$ (ºC)	$\delta w^{(+)}$ (kWºC^{-1})	$\delta w^{(-)}$ (kWºC^{-1})
H1	270	160	18	1	2	-2	3.6	-3.6
H2	220	60	22	1	2	-2	4.4	-4.4
C1	50	210	20	1	2	-2	4	-4
C2	160	210	50	1	2	-2	10	-10
CU	15	20		1				
HU	250	250		1				

Cost of Heat Exchangers (y^{-1}) =4000+500[Area (m^2)]$^{0.83}$
Cost of Cooling/Heating Utility = 20 (kW^{-1}y$^{-1}) /200 ($kW^{-1}y^{-1})

In order to compare alternatives it is necessary to estimate the range of the utility consumption needed to reject disturbances in the two worst cases. It will provide the range of variation during operation. To avoid this variation, and make a fair comparison between the three alternatives the inner HEN is controlled using as most bypasses as possible according to the rank of the matrix energy balance. The best results obtained to the retrofit of each structure using the bypass design procedure proposed is summarized in the Table 2.

As only 3 bypasses could be manipulated independently, one target temperature is controlled using the utility flow rate, which has been selected based on the cheaper utility. The results show that different investment levels are needed to each case. Comparing the three HENs the last one needs more investment, but the costs associated with the critical consumption assumed by the utility (used to estimate the operating cost) controlled target result in the cheaper solution.

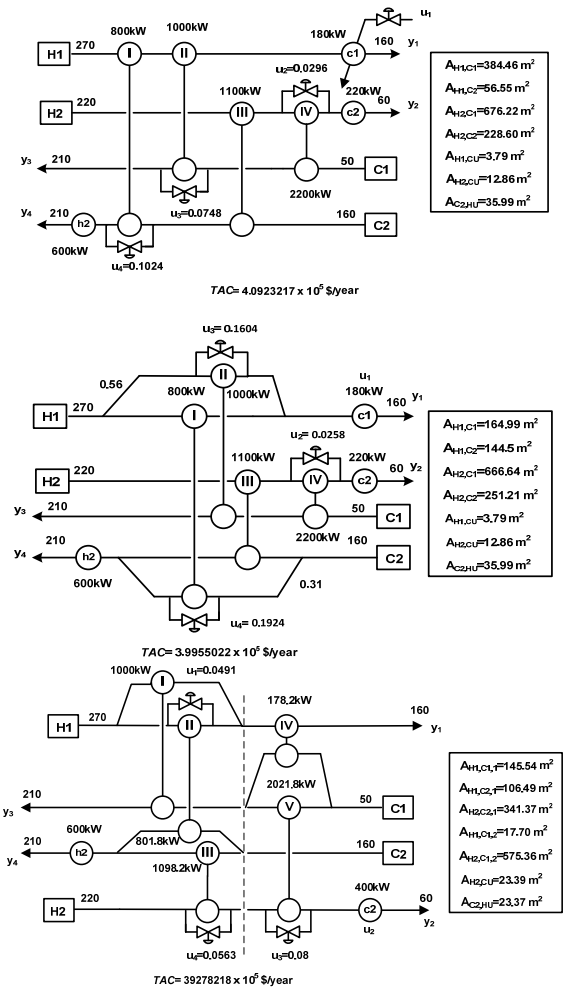


Fig 4. Retofitted HENS using bypass design.

Table 2. Bypass Nominal Design for the 3 structures.

Structure	***Controlled set***	***Manipulated set***	***Nominal Bypass Value***	***RGA Number***	$(\delta A/A_0)$ *%*	***I-I₀*** *($/year)*	***TAC*** *($/year)*
	T_{h2}^t	$u_{2,1}^h$	0.0296		20.93		
$HEN01$	T_{c1}^t	$u_{1,1}^c$	0.0748	0	10.85	22506	409294
	T_{c2}^t	$u_{2,2}^c$	0.1024		9.03		
	T_{h2}^t	$u_{2,1}^h$	0.0258		9.29		
$HEN02$	T_{c1}^t	$u_{1,1}^h$	0.1604	0	18.32	20572	399550
	T_{c2}^t	$u_{1,2}^c$	0.1926		19.23		
	T_{h1}^t	$u_{1,2,1}^c$	0.0491		3.92		
$HEN03$	T_{c1}^t	$u_{2,1,2}^h$	0.0800	1.44	25.27	25811	392782
	T_{c2}^t	$u_{2,2,1}^h$	0.0563		18.69		

This procedure must reject non controllable HENs, and provide a decision based on the trade-off between controllability and total costs, for this case the cheaper solution shows interaction a little bigger as can be seen by the *RGA* number. The final structures with control schemes are depicted in Fig. 4. For a final decision the dynamic model must be analyzed. The number n is an upper bound on the number of bypasses used to control. The investment cost increases, but it makes possible an operation close to the minimum utility consumption, if the maximum number of bypasses is designed. Therein, there is a trade-off that must be explored.

5. Conclusions

The design of a cost effective HEN capable of be controlled has both economical and operational significance. The controllability depends on the HEN structure, but to be evaluated it is not necessary to design the controller, but it must be selected a set of manipulated and controlled variables. It was presented a systematic model based framework for designing an appropriated control system, selecting the manipulated set, and design bypasses with minimum economic penalty. Through the prediction of the disturbances on controlled variables, it is possible to estimate the bypasses nominal fractions able to reject these disturbances solving an analytical problem per iteration resulting in a fast and robust procedure. The results demonstrate that the 3 structures analyzed are highly controllable with similar total annual cost. In order to make a final decision a comparative controllability analysis including the dynamic performance should be performed.

References

N. Aguilera, and Marchetti, J. L. , 1998, Optimizing and controlling the operation of heat exchanger networks. AIChE Journal, 44(5), 1090–1104.

B. Glemmestad, 1997, Optimal Operation of Integrated Processes, Studies on Heat Recovery Systems. Ph.D. thesis, Norwegian University of Science and Technology, Trondheim, Norwat.

K. W. Mathisen, S. Skogestad, and T. Gundersen, 1992, Optimal Bypass Placement in Heat Exchanger Networks,'' AIChE Meeting, New Orleans, LAŽ1992.

K. P. Papalexandri, and E. N. Pistikopoulos, 1994, A Multiperiod MINLP Model for the Synthesis of Flexible Heat and Mass Exchange Network, Comput. Chem. Eng., 18, 1125Ž.

Q. Z. Yan, Y. H. Yang, and Y. L. Huang, 2001, Cost-Effective Bypass Design of Highly Controllable Heat-Exchanger Networks, AIChE J., 47(10), 2253-2276.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Experimental study of a polycondensation reactor control by NMPC

Reinaldo A. Teixeira, Galo A. C. Le Roux

LSCP/CESQ - Department of Chemical Engineering, Polytechnic School - University of São Paulo. Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo, Brazil

Abstract

In this work, an implementation of the nonlinear predictive controller (NMPC) to a real 1,25 liter experimental polycondensation reactor is presented. The reaction studied is the reaction of dimethyl terephthalate (DMT) with Ethylene glycol (EG) producing esters and methanol (ME). The objective is to control the reactor temperature by means of the NMPC controller. A simplified model was derived from a complex phenomenological model. The simplified model was adjusted to experimental data. Two different implementations are presented. The performances of the NMPCs are compared with that of the traditional PID controller. NMPC with simple state feedback has the best performance because the state of the process is introduced into the model at each sampling time. Furthermore, NMPC with state feedback had a performance better than the PID in all the trajectories that were studied. The NMPC controller with simple state feedback is flexible because there is no need to adjust the model when there are disturbances in the process. This flexibility is important in the polymer industry, which generally works with multiproduct batch reactors, different batch sizes and variable product qualities.

Keywords: Experimental reactor, nonlinear predictive control, polycondensation, supervisory control.

1. Introduction

The process of batch polycondensation presented in this study is the reaction of dimethyl terephthalate (DMT) and ethylene glycol (EG), generating products such as esters and methanol (ME). The removal of methanol promotes the progress of the reaction to obtain a high degree of conversion. Reactors operating in this regime are known as semi-batch. The operation of reactors in semi-batch mode is a very usual practice in the polymer reaction engineering (Richards and Congalidis, 2006). The major concerns for process control are: safety, product quality, productivity and easy of scaling up (Bonvin, 1998). The supervision and control of the process in batch operation are of great importance in order to achieve these goals, (Ruppen et al., 1997).

Some difficulties encountered in batch processes are the absence of steady-states (natural characteristic of batch reactor) and the uncertainty of the mathematical model (Rho et al, 1998). Polycondensation processes involving the separation of the condensate (which are small molecules produced in the reaction of polycondensation of PET, e.g., water or methanol) have a high degree of nonlinearity when they reach the vapor-liquid equilibrium because of the interactions between the manipulated variables, generating control problems. For this case, the linear predictive controller (LMPC) and the traditional PI and PID may show a poor performance (Pottmann and Seborg, 1996).

Due to the control difficulties in the operation of polycondensation processes, decisions are often left to the operator, which is subject to operational risk (Bonvin, 1998). This justifies the experimental application of NMPC to polycondensation reactors (Henson, 1998, Nagy et al., 2007).

2. Experimental reactor and data acquisition system implementation

The batch reactor has a capacity of 1.25 liters. The system is supervised by a PLC (programmable logic controller). The reactor is jacketed laterally. Air is blown through this jacket in order to cool the reaction system.

The heating is supplied by a resistance at the reactor bottom. There is a thermal resistance (RTDs) for measuring the temperature of this resistance. A second RTD is inserted inside the reactor and measures the temperature of the reaction medium (Tr, reactor temperature). A packed–bed column is attached to the vapor outlet. The top of the column is connected to a partial condenser made up of a jacket through which cooling water flows and inside which vapor condenses in a bundle of vertical tubes. The flow of water used for cooling the partial condenser is handled through a peristaltic pump, also connected to the PLC. Another RTD is inserted into the partial condenser to measure the temperature at which the volatile products are condensed (Tcp, condensate temperature). The process is supervised by the SCADA software IFIX® 2.20 from GE Fanuc.

The architecture used for supervision and control action calculation consists of a Pentium 1.8 GHz computer with 1GB of RAM connected via RS-232 to the PLC. The IFIX 2.20 program has a VBA (Visual Basic for Applications) integrated, that allows to interact with external programs (eg, an executable NMPC controller). MATLAB 7.2 was used in order to develop the NMPC controller. The NMPC program is an executable file that allows receiving reactor temperature data from the IFIX program and produces an output value that is implemented in the experimental reactor.

The NMPC controller manipulates not only the heating power (Qr), but also the air-cooling flow to the reactor jacket (Qac) and in this case the NMPC has to solve a MINLP (mixed integer nonlinear programming) problem.

2.1. NMPC Algorithm

The NMPC algorithm is based on the minimization of an objective function, J(u), which a quadratic function of the errors of the controlled and manipulated variables that quantifies the performance of the controller within the prediction horizon (P). The control actions vary within a horizon M < P, known as the control horizon (M). Only the first control action computed is implemented in the process. Q, R and S are positive definite weighting matrices that penalize the controlled variables, the manipulated variables and the incremental manipulated variables respectively.

The flow rate of cooling is a binary variable, and the NMPC problem leads to a MINLP. The implementation of NMPC with MINLP problem is solved by exhaustive enumeration. In this case, there are four possible combinations, as the control horizon is 2. The combinations are:

1-heat in all control horizon (two heating actions);

2-heat as the first control action and then cooling;

3-cooling as the first control action and then heat;

4-cooling in the entire control horizon, in this case there is no optimization variable, and then the algorithm only evaluates the value of the function.

The solutions for each of these problems (except the fourth) are calculated using the successive quadratic programming in MATLAB. The lowest value of the four situations corresponds to the control action that is implemented in the process (reactor).

3. Simplified model of the experimental polycondensation reactor

A simplified model was derived from a complex phenomenological model. The complex model in details can be found in Carrillo Le Roux and Teixeira (2004). In the complex model, balances and equilibrium equations are considered for the reactor and the column with the partial condenser at the top.

The simplified model consists of a reactor and one external tank, that simplified represents the column by allowing the reflux of some liquid with a given composition to the reactor. The model assumptions adopted are:

- The column behaves as a tank where Ethylene glycol and methanol are accumulated.
- The hold-up of vapor phase is negligible;
- The heat loss through the walls of the external accumulator is negligible;
- The relation between the molar hold-up and the flow liquid is based on Billet (1995) correlation;
- The composition of the reflux to the reactor is given.

The simplified tank equations are represented by only mass balances.

The energy balance in the reactor, takes into account the thermal inertia of various physical components of the reactor and the resistance, and the heat transfer between these physical components is considered by phenomenological constants. The simplified model was adjusted to experimental data obtained in an open loop test (see figure 1). The test signal was applied in the heating power (Qr) and cooling flow (Qac). "Nothing" is applied to regions where no heat or cooling takes place.

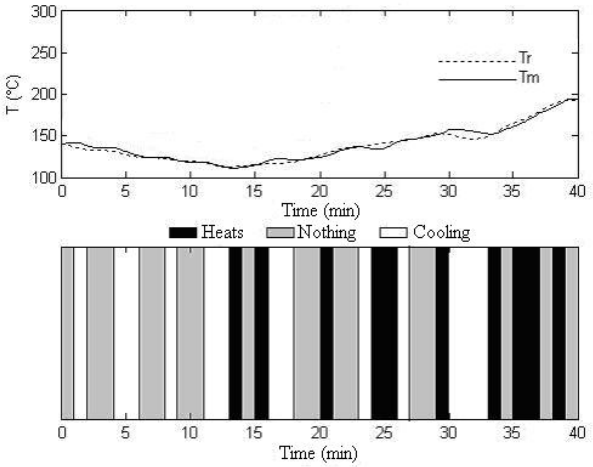


Figure 1 - Simplified Model adjusted with the experimental data obtained in an open loop test.

4. Application of the NMPC

The NMPC controller was used to close the temperature loop of the reactor. For the temperature on the top of the column, a PI controller was used. The controlled variable is the temperature of the experimental reactor (Tr) by manipulating of the heating power

(Qr) and the flow of the cold air flow in the reactor jacket (Qac). To control the partial condenser temperature (Tcp) the cooling flow (Qw) of cold water is manipulated.
The tuning parameters used in NMPC controller are M = 2 and P = 4. These were obtained from simulations where the performance of the controller was compared taking into consideration the necessity of a reduced computational task (Carrillo Le Roux and Teixeira, 2004). The batch duration is 120 min and at the end of the batch the controller is allowed to extrapolate from this point in the calculations. The weighting matrices Q = 1, R = 0, S = 1e-3, the sampling period Δt = 1 minute and the constraints used in inputs are: umin = 0 and umax = 100. For the PI controller, the tuning is: Kp = 0.5%/°C and Ti = 0.01 min/repetitions and sampling period is 1 second.
In all the experiments, the reactor is loaded with 1.5 moles of DMT (dimethyl terephthalate) and 3 moles of ethylene glycol (EG). Zinc acetate at a concentration of 3.75e-4 (mol cat/mol DMT) is used as a catalyst. In order to favor the progress of the etherification reaction, the set point for the column condenser temperature is 67 °C, allowing preferably the removal of ME. The trajectory used as the set point for the reactor starts at 140 °C and goes to 180 °C at a rate of 4 °C/min, thereafter the rate is 0.25 °C/min until 200 °C, keeping this temperature until 120 min;
To compare the performance of the NMPC controller a PID was used, with tuning parameters Kp = 1.45%/°C, Ti = 0.28 min/repetition, Td=1.4 min with sampling period 6 s. The comparison is presented in Figure (2) and (3) and the values of the ISE and IAE are presented in Table 1. Comparing the results, the PID controller had a performance equivalent to that of the NMPC controller. Notice that the model (see the model temperature, Tm) in figure (3) and the reactor temperature agree during the steps where NMPC control has a good performance. However, between the times 100 and 120 minutes the model temperature gets far to the reactor temperature (Tr) and the performance of the NMPC controller decreases. To increase the performance of the NMPC controller, a simple strategy of state feedback was implemented: the measured reactor temperature is fed as the initial condition of the reactor temperature in the reference model. The results for the NMPC with simple feedback are presented in figure (4) and table 2. The previous NMPC is thus called simple NMPC. The performance of NMPC with simple state feedback increases significantly and is better than the previous NMPC and PID controllers.

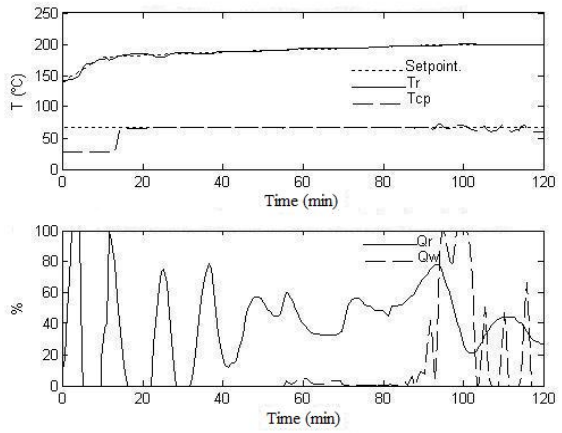


Figure 2 - Results of the PID controller applied to the experimental reactor.

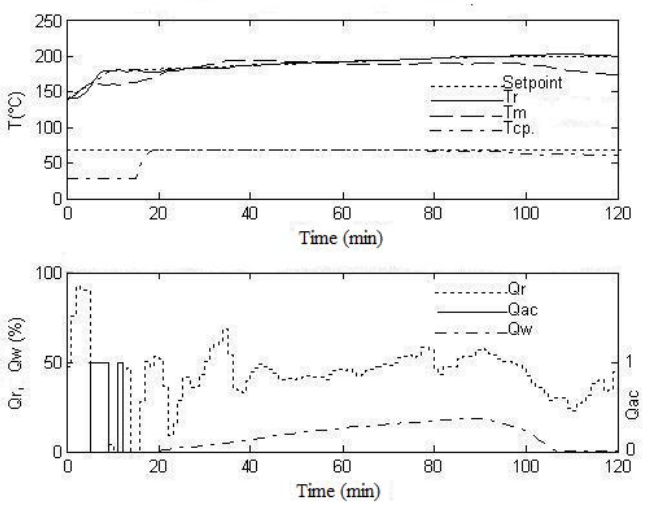


Figure 3 - Results of the simple NMPC controller applied to the experimental reactor.

Table 1 – Comparison of NMPC and PID controllers by ISE and IAE values.

	NMPC	PID controller
ISE	696	621
IAE	198	241

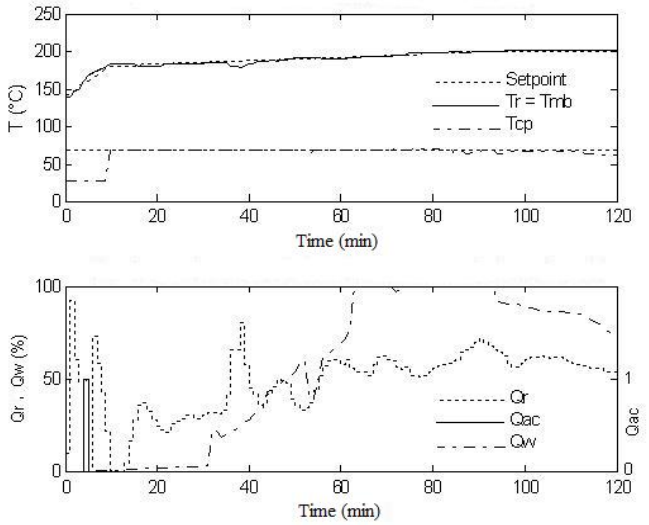


Figure 4 - Results of the NMPC with state feedback in the experimental reactor.

Table 2 – Comparison of the NMPCs and PID controllers by ISE and IAE values.

	NMPC controller with simple state feedback	Simple NMPC	PID controller
ISE	585	696	621
IAE	186	198	241

5. Conclusions

Two NMPC controllers were implemented in an experimental reactor for polycondensation and their performances were compared with that of a PID controller using ISE and IAE metrics. The PID controller showed a performance similar to that of the Simple NMPC controller. The performance of the Simple NMPC controller was degraded because the model does not represent well the process in some ranges. Therefore, an NMPC with simple state feedback was proposed it had a performance better to that of the Simple NMPC and the PID controllers. An advantage of the NMPC with simple state feedback is that it isn't necessary to adjust the model when disturbances in the process take place. This flexibility is important in the polymer industry, which generally works with multiproduct batch reactors, different batch sizes and variable product qualities.

Acknowledgements

The authors acknowledge the Coordenação de Aperfeiçoamento de Nível Superior, CAPES and FAPESP for their financial support.

References

R Billet,1995, Packed Towers in Processing and Environmental Technology, VCH Publishers, Inc., New York,.

D. Bonvin,1998, Optimal operation of batch reactors - a personal view, J. Proc. Cont., 5-6, p.355-368.

G. A. Carrilo Le Roux; R. A.Teixeira, Application of nonlinear predictive control to a semi-batch polycondensation reactor, Industrial Engineering & Chemistry Research,43(23) p. 7303- 7311, 2004.

M. A. Henson,1998, Nonlinear model predictive control: current status and future directions, Comput. Chem. Engng, 23 p.187-202.

N. Kaistha; M. S.Johnson, C. F.Moore, M. G. Leitnaker. Online batch recipe adjustments for product quality control using empirical models: Application to a nylon-6.6 process, ISA Transactions, Vol. 42 p.305-315, 2003.

Z. K. Nagy, B. Mahn, R. Franke, F. Allgöwer. Evaluation study of an efficient output feedback nonlinear model predictive control for temperature tracking in an industrial batch reactor, Control Engng Pratice,Vol. 15 p.839-850, 2007.

M. Pottmann, D. E. Seborg. A nonlinear predictive control strategy based on radial basis function models, Comput. Chem. Engng., Vol. 21, No. 9, p. 965-980, 1997.

J.R. Richards J. P. Congalidis, Measurement and control of polymerization reactors, Comput. Chem. Engng., Vol. 30 p. 1447-1463, 2006.

D. Ruppen, D. Bonvin, D.W.T. Rippin, Implementation of adaptative optimal operation for a semi-batch reaction system, Comput. Chem. Engng., 22, p.185-199, 1997.

H.-J. Rho, Y.-J. Huh, H.-K. Rhee, Application of adaptative model-predictive control to a batch polymerization reactor, Chem. Eng. Sci., 53, p.3729-3739, 1998.

R. A. Teixeira, Aplicação de controlador preditivo não linear a um reator de policondensação, M.Sc. Thesis (in portuguese), Escola Politécnica da USP, 2003.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Control loop Performance Assessment and Improvement of an Industrial Hydrogen Generation Unit

Luís G. S. Longhi,[a] Alex S. Reginato,[b] Leandro P. Lusa,[c] Santiago S. Gonzalez,[d] Thiago D. Fleck,[e] Cristhian A. C. Cortez,[f] Herbert C. G. Teixeira[g]

[a,b]*Alberto Pasquilini – REFAP S A., Av .Getúlio Vargas, 11001, Canoas/RS, 92420-221, Brazil, {longhi,alex.reginato}@petrobras.com.br*
[c,d,e]*Trisolutions Soluções em Engenharia, Rua General Bento Martins, 24/111, Porto Alegre/RS, 90010-080, Brazil, {leandro,santiago,thiago}@trisolutions.com.br*
[f,g]*Petróleo Brasileiro S.A. – PETROBRAS, CENPES, Rio de Janeiro/RJ, Brazil, {cristhiancortez,herbertteixeira}@petrobras.com.br*

Abstract

This paper describes the CLPA (Control Loop Performance Assessment) and improvement of an industrial hydrogen generation unit (HGU) located at Alberto Pasqualini – REFAP S.A., a petroleum refinery (whose capacity is 189,000 barrels per day) located at the south of Brazil, subsidiary of Petróleo Brasileiro S.A. (PETROBRAS), and also belonged to REPSOL YPF. At first, we describe the HGU unit and its process purpose. After, we introduce the motivation for the CLPA for this unit and describe the main tool used to carry out this task. This tool is entitled BRPerfX (petroBRas PERFormance indeX) and it was developed by the Federal University of Rio Grande do Sul, Trisolutions enterprise and CENPES (the PETROBRAS research center). The necessary capital to develop this tool was supported by PETROBRAS. A base case was developed for comparison. This base case is composed by a time interval of typical operation before the project startup. This allows showing the improvements made after the control loop assessment is finished. The main steps of the work methodology are also presented, with focus on the planned and implanted actions to improve the dynamical performance of the unit. Finishing the paper, we present the main results achieved. These include not only the better dynamic behavior of the loops but also the economic results related to the index online monitored in the CLPA tool.

Keywords: control loop performance assessment, process control, industrial application.

1. Introduction

Since the publication of the seminal work of Harris [1], the interest in control loop performance assessment (CLPA) has reflourished in the control community. This occurred because, for the first time, this community clearly saw that was possible to deal with the important problem of online evaluating the performance of control loops with a realistic framework. In subsequent years several works concerning this subject arose in the open literature and some tools started to be developed by some control and automation enterprises. Instead of purchasing some of those tools, PETROBRAS, at dawn of current millennium, decided to support Federal University of Rio Grande do Sul and the Trisolutions enterprise aiming the development of a new tool for CLPA.

This new tool has evolved since then and resulted in the BRPerfX software. This tool was applied to the REFAP's HGU last year and the presentation of this project and its results is the main concern of this paper.
In the sequence of this work we describe the HGU and the BRPerfX software in section 2. In section 3, the main results are presented. These results encompass the project main steps, the development of the base case, the actions made to improve the dynamic performance, the evidences of that improvement, and the estimation of the economic gains after the actions made by our technical team. Finally, we present some brief concerns about the overall work in the conclusions section.

2. Materials and Methods

2.1. HGU Description
The REFAP HGU started up in December 2005 and produces hydrogen up to 99.9 % purity level from natural gas. The unit objective is to supply hydrogen for the subsequent diesel hydro treating unit according to its demand. As the diesel is the main REFAP's product and it can only be sold after being hydro treated, the continuous and stable operation of HGU is fundamental for attending one of the most important strategic enterprise objectives. The HGU is composed by three main stages (Fig. 1):

2.1.1. Desulfurization and Pre-heating
In this stage, the natural gas is fed in the unit and receives a hydrogen injection, being immediately sent to heating in the convection section of the reformer furnace. There, it acquires the necessary temperature to start up the reactions.

2.1.2. Conversion
Here, the pre-heated load receives steam and is converted to CO, CO_2 and H_2. After, the final conversion from CO to CO_2 and H_2 is carried out in the shift reactor.

2.1.3. Purification (PSA – Pressure Swing Adsorption)
The PSA purpose is to purify the hydrogen produced in the HGU until it reaches a concentration of 99.9 % (volumetric). It utilizes the molecular sieve process. The gas is admitted at the bottom of the vessel where the impurities filtering and retention occurs.

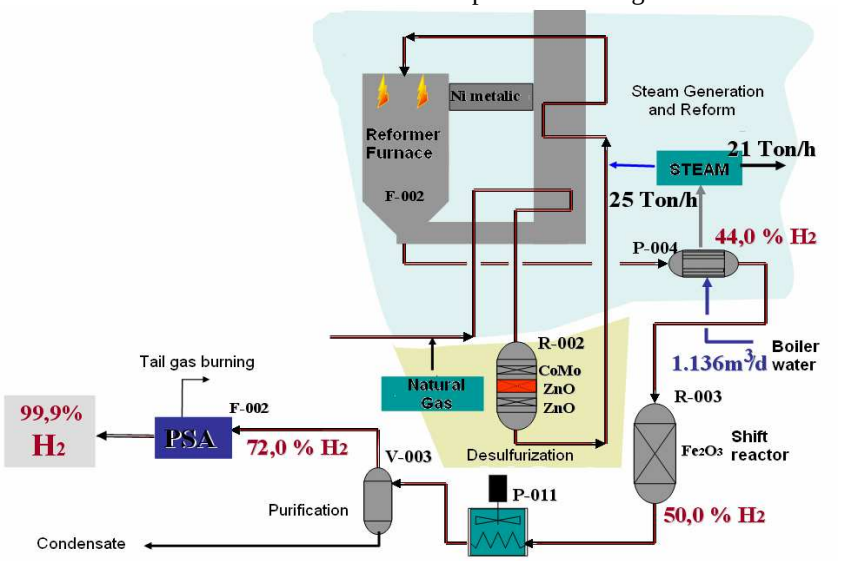


Fig. 1: HGU scheme. The HGU produces 550,000 Nm3/d of H2.

2.2. The BRPerfX tool

The BRPerfX is a real-time tool for performance monitoring and evaluation of control loops developed by Federal University of Rio Grande do Sul and Trisolutions enterprise and supported by PETROBRAS. It allows the identification of the main responsible for the bad control performance of a plant with large number of control loops. This is possible due to the state of art level of indexes and metrics available in the software. These includes: (1) oscillation detection indexes, (2) performance metrics based on minimum variance index, (3) traditional statistics, (4) Valve analysis, (5) controllers service factors, (6) PID tuning changing monitoring, (7) several graphical analysis and, (8) customized weighted grades for equipments, units and plants.

The application works online without any interference in the plant operation. The essential information required for its correct work are the PV (Process Variable), SP (Set-Point), CO (Controller Output) and, MO (Operation Mode). To get these data, it uses an OPC (OLE for Process Control) DA (Data Access) interface, a standard to connectivity among systems largely adopted in industry. The software architecture is composed by three main parts: a Windows service, a database and, a Web application. A typical view of the web user interface is shown if Fig. 2

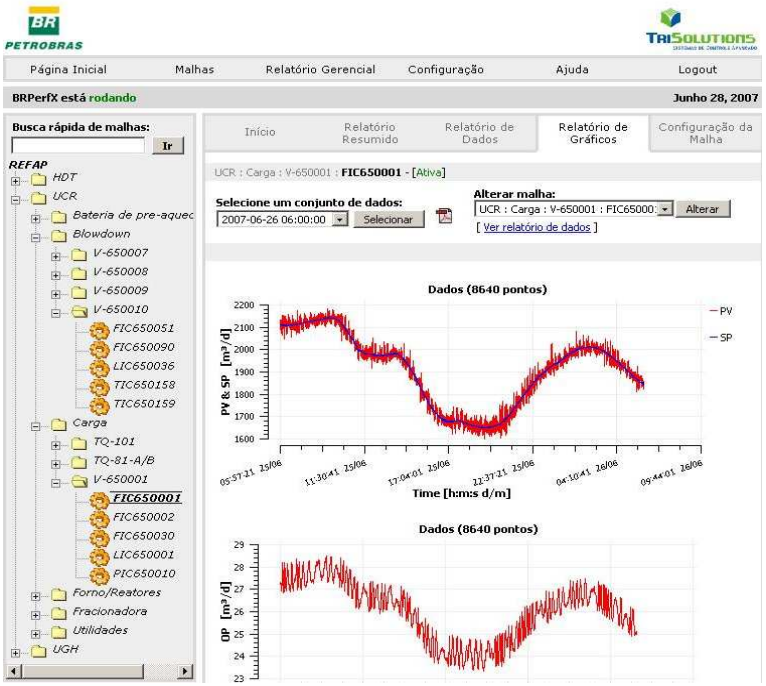


Fig. 2: BRPerfX user interface in web environment.

3. Results

3.1. Project Stages

To be successful, an industrial control management project should be composed by four basic steps plus cycles for continuous improvement. The time ranges, estimated in the following lines, are the ones for a typical unit with about 50 to 100 control loops. These steps are: (1) software installation and configuration – this usually takes one week. For the HGU case it was very fast, since the software was already installed. (2) control loop

evaluation and elaboration of a base case – at least one month, depending on plant operating conditions. This stage is important because it furnishes the reference for future performance comparisons, (3) Corrective actions – the range time depends on problem's nature. The time to solve tuning problems can be estimated about 10 loops for week. Maintenance and project problems have much slower solving velocity. (4) Process revaluation and economic gains estimate – this will show the relevance of the work for the high level enterprise staff. It can be estimated about, at least, one month to collect data and one week to compute the gains. This will be explained in section 3.5.
All these steps must be followed by new cycles for continuous improvement and preventing performance degradation.

3.2. The Base Case

The base case for this work was done with data collected from 2008 August 8 up to 2008 September 8. In this time range the unit was under stable operation and no actions were made by the control engineers. After a carefully analysis of data, the base case was defined as the one briefly presented in Table 1.

Table 1 – The base case

Number of loops:	42	
Problem classification	#loops	Percent
DCS configuration	2	5%
Other	3	7%
Maintenance	1	2%
Tuning	21	50%
Operation	4	10%
Project	1	2%
No problems detected	10	24%
Saturation (>= 25% of time)		
Loops	13	31%
Operation mode (>=25% of time in manual)		
Loops	9	21%

3.3. Actions Made

The actions made by the control engineers were focused mainly in tuning and DCS configuration problems. A new cascade controller was also included after a review in the HGU control philosophy. This work took about two discontinuous months. The results, in comparison with the base case, are presented in Fig. 3.

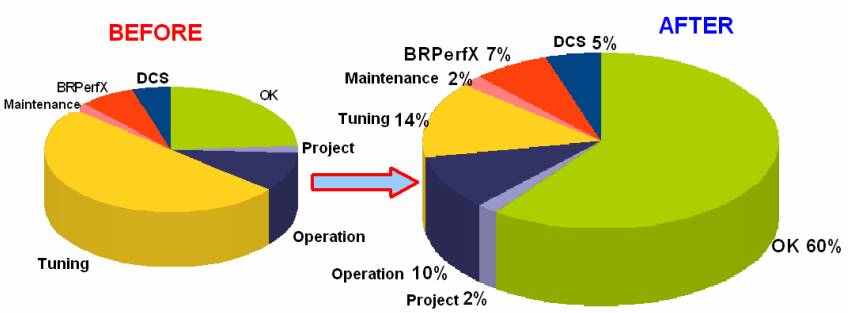


Fig. 3: Evaluation of control loops before and after the actions made by the control engineer team.

3.4. Dynamic Improvement

The process improvement can be clearly seen by looking some pictures available in the BRPerfX software. Due to space limitations a small fraction of them are presented in this section. Some parts of the graphics are written in Portuguese (software language).

The first example concerns the O_2 excess in the reformer furnace, controlled by the controller AIC702002. This controller is the master of a cascade that operates in the air fan blower. This controller had a very conservative performance, being very slow and oscillating. As a consequence, the operators had orders to operate at a higher than recommended oxygen excess. After the control engineers work, the standard deviation of the loop goes down from 0.41 to 0.06 (Fig. 4). This also created a potential to operate at lower levels of air excess. Despite being a small furnace, this allows a fuel gas reduction and a monthly economic gain about R$ 10.000,00, or USD 4,400.00 / month (using the 2009 February, 6 exchange rates).

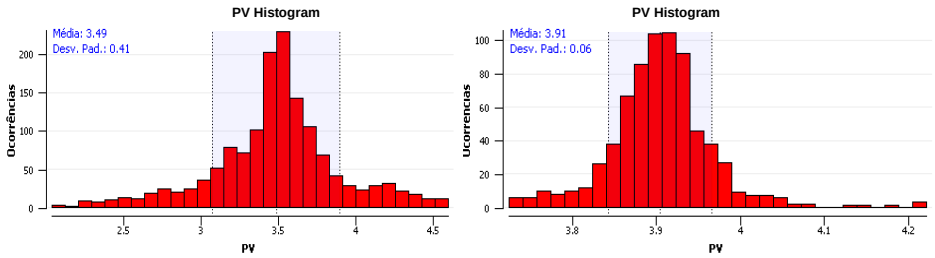


Fig. 4: BRPerfX histogram of AIC702002 before and after the control engineers actions.

The second example is the pressure control of the reformer furnace pilot burner, PIC702051. After our action, its very high oscillations have vanished, as Fig 5 shows.

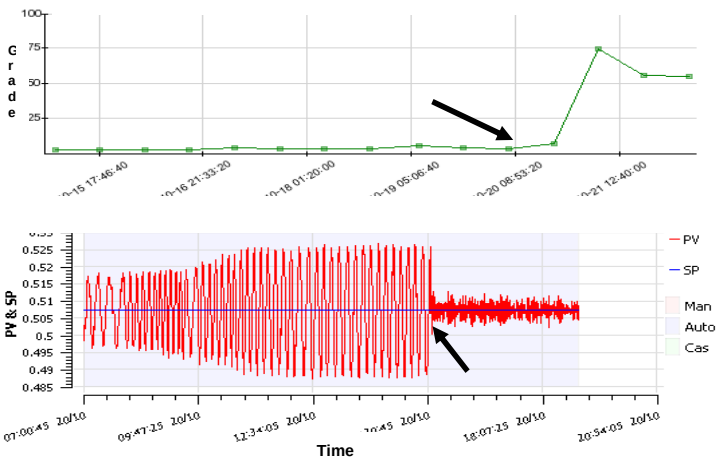


Fig. 5: Evaluation of PIC702051. In the top, the one week control loop grade evolution. Below, it is shown the PV-SP graphics, for a 12 hour period. The arrows point the time we have acted in.

Another illustrative example of the insights obtained by using the software occurred in the natural gas recycle pressure controller, at the feed HGU section – PIC702003. This controller presented very high oscillations with nearly 2 minutes period, as can be seen in the power (Fourier) spectra computed by the software (Fig. 6). After retuning, the

performance improvement - also shown in Fig. 6 – is obvious. The spectra became more distributed and less dense in the higher frequencies.

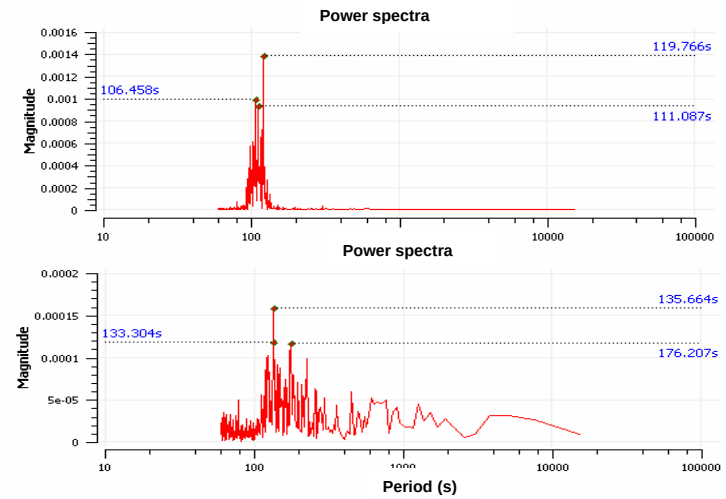


Fig. 6: Power spectra after (top) and before (bottom) the control loop performance improvement.

More interesting than showing the better performance, the new index (TuniPID) [2] included in the software indicates to the user that the control can be improved, but not by a PID because the control loop is very close to the best achievable PID performance.

3.5. Economic gain

Despite showing the dynamic behavior is much better now than ever in the HGU short history; this has no significant meaning for the enterprise managers and maybe for other people but control engineers. So, it is fundamental to compute the economic gains of the whole work. This computation varies from one unit to another and there is no applicable methodology avoiding a process analysis. Facing this fact, REFAP's and Trisolutions engineers internally developed a framework to compute gains from regulatory control improvement for chemical plants. This is based on comparisons with base case performance, accounting the price evaluation (for reactants, intermediate products and fuels) and, the formalization of the basic tangible mechanisms for plant profit: selectivity, quality, productivity, fuel efficiency, loss reduction, among others. After one week of careful calculations, we estimate the official number for the economic gain of this work about USD 85,000.00 per month, plus the intangible gains related to plant stability and the consequent reduction of scheduled and/or unplanned plant shutdown.

4. Conclusions

This work presents a project that applies new academic developments to solve practical industrial problems. Due to space limitations, we conclude only stating that the critical factors to its successful implementation were: IT infrastructure, reliable tool choice and, staff and management commitment. We thank to all refinery staff for their gentle help.

References

[1] Harris, T. J., 1989, Assessment of control loop performance, The Canadian Journal of Chemical Engineering, 67, p. 856-861.

[2] Farenzena, M., 2008, Novel Methodologies for assessment and diagnostics in control loop management, Ph.D. Thesis, Federal University of Rio Grande do Sul.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Advanced Control and Optimization of a Natural Gas Plant - Benefits of the new regulatory control strategy

Mario C.M. Campos,[a] Marcos V. C. Gomes,[b] Vicente D. Moreira,[c] Marcos F. Lima,[d] José Renato Silva,[d]

[a]*PETROBRAS / CENPES / Engenharia Básica/ Automação, Equipamentos Dinâmicos e Confiabilidade, Av. Horácio Macedo, 950 - Ilha do Fundão, Rio de Janeiro, 21949-915, Brazil (e-mail: mariocampos@petrobras.com.br).*
[b]*PETROBRAS / CENPES / P&D de Gás, Energia e Desenvolvimento Sustentável / Gás Natural / Célula de Otimização*
[c]*PETROBRAS / UN-RNCE / ENGP, Natal, Brazil*
[d]*PETROBRAS / UN-RNCE / UTPF, Natal, Brazil*

Abstract

This paper presents the design of a new regulatory control strategy, as part of an advanced control and optmization project, for the natural gas processing plant of PETROBRAS (UPGN-III in Portuguese symbols) in Rio Grande do Norte. This phase of the implementation of the advanced control and optimization system (multivariable model based constraint control) allowed an increase of about 4% in the Liquefied Petroleum Gas (LPG) per charge and composition, this is the principal product of this unit, representing an income of US$ 1.300.000 a year. The new control system also reduced in 30% the number of unscheduled unit's downtime due to an improvement of its stability.
Keywords: advanced control system, regulatory control, industrial application, loop performance assessment.

1. Introduction

The advanced control system in petrochemical plants is an industrial reality (Qin and Badgwell,, 2003). There are many advantages for the process units in using these systems as more stability, respect to the restrictions and an increase in the profitability and safety. PETROBRAS has been investing in the development of these systems for several years. It is a consolidated technology in its refineries with many controllers for FCC and distillation units. However, the application in Natural Gas Plants (UPGNs in Portuguese symbols) is recent, although the optimization of these plants can bring great economical earnings, besides to increase the energy efficiency and to minimize the emissions.

Installation and maintenance of advanced controllers with a good performance is a big challenge. Their performance is influenced by instrumentation problems, bad tuning of the regulatory control, bad identification of the dynamic models of the process (Ender, 1993)(Kern, 2007), no measured disturbances, etc. In this article, we will discuss some tools that were used to the diagnosis and tuning of the regulatory control and the instrumentation, which is an important phase of the implementation of an advanced control in a Natural Gas Plant. We also show some results obtained in a real case during this phase.

2. Advanced Control System

Process control aims to maintain certain variables in their desirable operational limits and could be visualized as a pyramid. In the base of this pyramid, the first level is the regulatory control, that uses PID controllers (Campos and Teixeira, 2006)(Ogata, 1982) and is configured in the digital systems (DCS - Distributed control system or PLC - Programmable logical controllers). In a second level, we have the advanced control systems that use for instance Model Predictive Control (MPC). This algorithm considers the interaction between control loops, and includes an optimization layer of the industrial plant. These algorithms are usually implemented in a process computer that communicates with the DCS using for example OPC protocol (OPC, 2008). The outputs of this advanced control are usually the setpoints of the PID controllers of the regulatory control. So if there is a problem in the advanced control, the plant operation continues with the last PID setpoints in the DCS.

Model predictive control tries to control several process variables simultaneously acting in many manipulated variables. The advanced control algorithm used in this natural gas plant is called CPM (SSP-Laplace) and was developed by PETROBRAS/CENPES.

CPM algorithm incorporates a dynamic model of the process, used to obtain the control actions that minimize the effects caused by plant disturbances. The multivariable and dynamic model of the process allows to increase the performance of the control system mainly when we have some of the following problems:

- Multivariable and coupled control loops;
- Process with long dead times or long time constants.

Best operating setpoint is obtained by an optimization module of the CPM. This module determines this point based on economical criteria, considering the process constraints.

An advanced control system won't reach the expected economical benefit if turned off constantly by the operators. Therefore, the instruments, valves and the regulatory control (PIDs) should operate appropriately. Besides, training is also a key factor to the technology absorption for the plant personnel (operators and process engineers), in order to use the system adequately. The phases to implement an advanced control system are: Functional project (select the economical objectives, the controlled and manipulated variables and disturbances), Diagnosis, evaluation and tuning of the regulatory control and instrumentation; Identification of the process models; Installation and tuning of the model predictive control and Training, maintenance and documentation;

In this article we will focus in the results of the phase associated with the diagnosis, evaluation and tuning of the regulatory control and instrumentation. The project of advanced control for the natural gas plant (UPGN-III) of Guamaré - Rio Grande do Norte (Northeast of Brazil) began in March of 2007.

3. Description of the Natural Gas Plant (UPGN-III)

The natural gas plant of PETROBRAS/UN-RNCE (UPGN-III) recoves the heavy fractions (LNG) contained in the natural gas. First, this gas is chilled and after the expander its temperature is enough to condense the heavy hydrocarbon fractions (C_{3+}). The flow is sent to a distillation column that separates the gas in the top, known as residual gas, from the liquid (LNG). The residual gas exchanges heat with the feed in a cold box. Afterwards, it goes to the suction of the turbo-expander compressor that is in series with the sale gas compressor. Economically the objective is obtain minimum lost of C_{3+} in this gas (Campos et al., 2008).

4. Diagnosis, evaluation and tuning of the regulatory control

The good performance of the regulatory control is fundamental to the success of the advanced control system. This is due to the fact that the advanced system acts usually changing the regulatory setpoints. Therefore, if a problem exists in the regulatory control then the performance of the advanced system is affected.

Common problems associated with the regulatory control are: tuning of the PID controllers (oscillation, stability, etc.), control strategy (interaction, degrees of freedom, etc.), and instrumentation problems (out of scale, measurement resolution, sizing, etc.).

So, it is clear the importance of the diagnosis and evaluation of the regulatory control. To help us in this task, reducing the subjectivity, PETROBRAS developed with Federal University of Rio Grande do Sul (UFRGS in Portuguese symbols) the software “BR-PerfX”. It calculates some key performance indicators and generates graphs that are used to quantify the quality of the regulatory control (Kempf, 2003)(Farenzena and Trierweiler, 2008). From the analysis of these results it is possible to identify the problem affecting the regulatory control.

If the problem is the tuning of PID controller then the software "BR-Tuning" is used. This software was developed from a partnership between PETROBRAS and the Federal University of Campina Grande (UFCG in Portuguese symbols).

BR-Tuning (Schmidt et al., 2008)(Arruda and Barros, 2003) implements a group of techniques for open and close loops identifications and proposes new tuning parameters. It communicates directly with the process automation system (DCS or PLC) using the OPC protocol (OPC, 2008).

The main problem found in this plant (UPGN-III) was the turbo-expander flow control (FIC-01) operating continuously in manual mode. When the control was putted in automatic mode occurred a strong interaction with the pressure control (PIC-01), resulting in oscillations. It was the cause of many trips of the unity, and was a great operators' concern. Figure 1 displays the old control strategy of the unit.

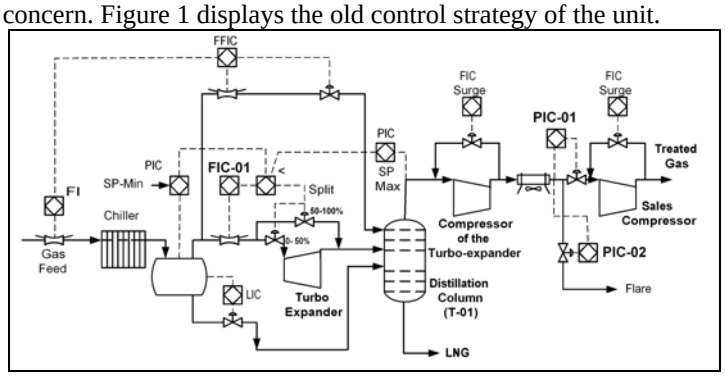


Figure 1 – Old regulatory control strategy.

The problems with the old control strategy were the following ones:

- The setpoints of the pressure controls in the suction of the compressor associated with the turbo-expander (PIC-01 and PIC-02) were difficult to be adjusted, because they should be changed if the turbo-expander was on or off. For example, if there were a trip in the turbo-expander, these setpoints should be changed quickly to lower values in order to avoid a trip in the whole unit due to high pressure in the distillation column (T-01). This trip was observed many times in this plant.

- The degree of freedom was not well used, because the control strategy was fixing many variables of the compressor associated with the turbo-expander. Therefore the control system was not stable in automatic mode.
- When the compressor associated with the turbo-expander opened its anti-surge recycle valve (Campos and Teixeira, 2006), the discharge pressure fell and the controller (PIC-01) closed the suction valve of the sales compressor, taking in a not correct manner this machine also to surge. This event took some times the unit to oscillate and to trip, generating losses.
- The control pressure of the distillation column (T-01) didn't operate in automatic mode in spite of being the most important column of the plant. Abrupt variations in this pressure increase the losses of LNG.

A new control strategy was proposed, according to figure 2. Basically, we changed the location of the process variables for the pressure control (PIC-01 and PIC-02). These controllers are now monitoring the pressure of the distillation column (T-01).

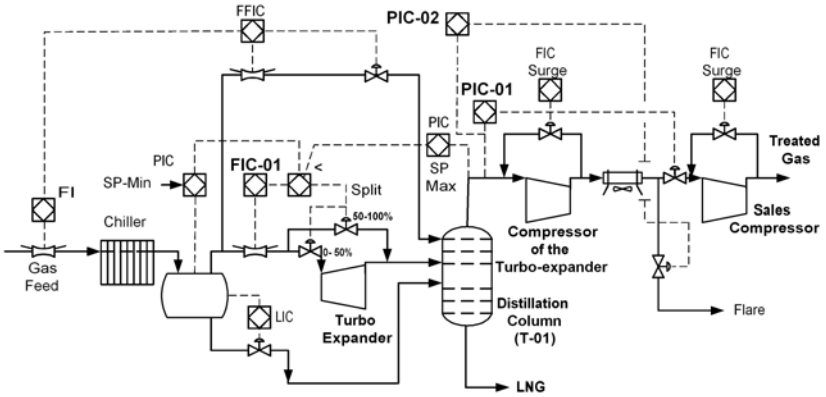


Figure 2 – New regulatory control strategy to UPGN-III.

The turbo-expander flow control (FIC-01) in this new strategy will define the compressor speed but the discharge pressure is free to accommodate changes in compressors efficiency. This new regulatory control strategy does not the old problems associated with the unit trip. Another advantage is that pressure control is now in automatic with a good performance (figure 3), manipulating the suction valve of the sales compressor (PIC-01).

The plant flow control (FIC-01), that manipulates the gas to the turbo-expander, is now operating in automatic mode continually (figure 4: FIC-01 performance - setpoint: green curve, flow: blue curve, and the valve opening: yellow curve). The next phase of the advanced control project will be the model identification and the tuning of the model predictive controller. This new phase will be discussed in a next paper.

5. Economical benefits and safety result in this project phase

With this new regulatory control strategy, the unit is easier to be operated and it is easier to find a better operating point, allowing to increase the profitability of the plant. Besides, more than fifty control loops were analyzed and tuned. Many levels controls were tuned in order to minimize the movement of the associated flow, improving the stability.

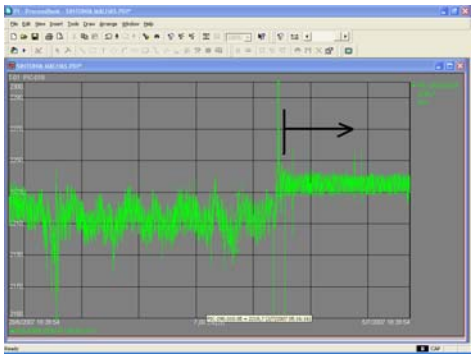


Figure 3 – Greater stability of the distillation column pressure.

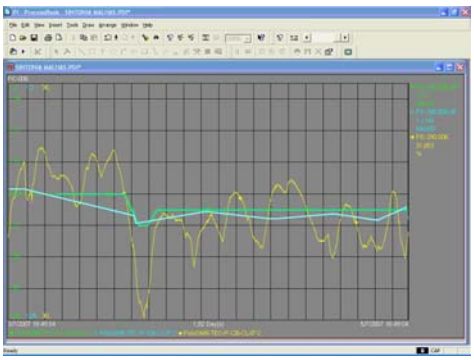


Figure 4 – Flow control in automatic, manipulating the gas to expander.

LNG of the T-01 (figure 1) is sent for another distillation column (not shown in figure 1) where Liquefied Petroleum Gas (LPG) is obtained, which is the principal product of this unit (UPGN-III). In order to calculate the benefits of this project, we defined a performance indicator based on the LPG production. This indicator considers the production of LPG in relation to the C_{3+} in the feed gas:

$$Indicator = \frac{LPG\ \ Flow}{Feed\ \ Flow} * \frac{1}{\%\, C_3^+ \text{ in the Feed}}$$

After the tuning of this new control strategy, this indicator passed from 3,17±0,17 to 3,30±0,08. So it was obtained an increase of about 4,1% in the production of LPG. This result represents an increase in the profit of the unit of approximately US$ 1.300.000/year. We can also observe a reduction in the variability of the plant (smaller standard deviation). Figure 5 shows the performance indicator improvement after the implementation and tuning of the new control strategy. Another benefit is associated with the reduction of the number of trips (non-programmed stops) of the unit. The medium time of the unit stopped passed from 729 minutes a month (May of 2006 to March of 2007) to 485 minutes a month (July of 2007 to May of 2008). This means a reduction of about 33%. It also should be noticed that less trips implicates in less emissions. So, the system also improved the health, safety and environment aspects (HSE) of the plant.

6. Conclusions

This paper presented the results obtained with the implementation of a new regulatory control strategy, which is an important phase to the implantation of the advanced control

system. It showed that the old control strategy was unstable and responsible for many unit trips. A new regulatory control strategy was able to solve these old problems.
The next phases of this project will be the identification of the models and the tuning of the predictive controller.

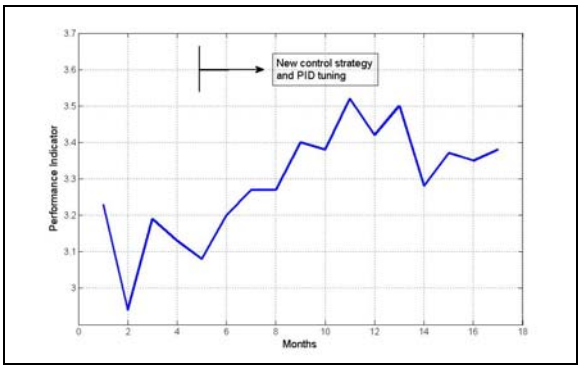


Figure 5 – Better performance indicator after the implementation of the new control strategy.

However, it is worthwhile to emphasize the economical results obtained with the implantation of this new regulatory control strategy. Only this phase allowed an increase of approximately 4% in the LPG production (an increase in the profitability of US$ 1.300.000/year). Other observed benefit was an expressive reduction of 33% in shutdown time of the unit.

Acknowledgements

We are grateful to many colleagues from PETROBRAS and Federal University of Rio Grande do Norte (UFRN) that contributed and participated in this work.

References

Arruda, G. and Barros, P., 2003, “Relay based gain and phase margins PI controller design”, IEEE Transactions on Inst. And Meas. Tech., 52(5), 1548-1553.

Campos, M. and Teixeira, H., 2006, “Controles típicos de equipamentos e processos industriais”, Ed. Edgard Blücher, São Paulo.

Campos et al., 2008, “Sistema automático de identificação de modelos dinâmicos necessários ao Controle Avançado – Aplicação na UN-RNCE”, Primeiro CICAP – Congresso de Instrumentação, Controle e Automação da PETROBRAS, May, Rio de Janeiro.

Ender, D., 1993, “Process Control Performance: Not as good as you think”, Control Eng., 1993, pp.180.

Farenzena, M. and Trierweiler, J., 2008, “Fronteiras e desafios em gerenciamento de malhas de controle”, In: COBEQ 2008 - Congresso Brasileiro de Engenharia Química, Recife.

Kempf, A., 2003, “Avaliação de Desempenho de Malhas de Controle”, Dissertação de Mestrado, Departamento Eng. Química, Universidade Federal do Rio Grande do Sul, UFRGS.

Kern, G., 2007, “Summiting with multivariable predictive control”, Hydrocarbon Processing,

Ogata, K., 1982, “Engenharia de Controle Moderno”, Ed. Prentice/Hall do Brasil.

OPC Foundation, 2008. Site: <http://www.opcfoundation.org/>, Accessed in: 03/17/08.

Qin, S. and Badgwell, T., 2003, “A survey of industrial model predictive control technology”, Control Engineering Practice 11, 733-764.

Schmidt et al., 2008, “BR-Tuning Ferramenta para sintonia de controladores PID”, Primeiro CICAP – Congresso de Instrumentação, Controle e Automação da PETROBRAS, May, Rio de Janeiro.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Versatile Biodiesel Production by Catalytic Separative Reactors

Anton A. Kiss

AkzoNobel Research, Development & Innovation, Process & Product Technology, Arnhem, The Netherlands, tony.kiss@akzonobel.com

Abstract

This study proposes an integrated biodiesel production via a two-step process that combines the advantages of using solid acid and base catalysts with the integration of reaction and separation. Such an integrated separative reactor is flexible to treat any range of free fatty acids present in the fatty raw material. Computer aided process engineering tools such as AspenONE are used for process design and simulation of a plant producing 10 ktpy biodiesel from animal fat and bio-ethanol.

Keywords: reactive distillation, green catalysts, solid acid / base, biofuels

1. Introduction

The recent steep increase in fossil fuel prices associated with governmental restrictions on discharge of green-house gasses shifted the worldwide trend to focus on renewable energy sources. Biodiesel is a very popular renewable fuel, currently produced from vegetable oils, animal fat or even recycled waste cooking-oil from the food industry.[1,2] Due to its properties similar to petrodiesel, biodiesel can be used in pure form, or may be blended with petroleum diesel at any concentration, in most modern diesel engines. As a *green fuel*, biodiesel that has many advantages over conventional petrodiesel: it is safe, renewable, non-toxic and biodegradable, it contains insignificant amounts of sulfur and its increased lubricity extends the life of diesel engines. In addition, it has a high cetane number (above 60 compared to 40 for petrodiesel), a high flash point (>130°C) and it emits ~70% fewer hydrocarbons, ~80% less CO_2, and ~50% less particles.[1-3]

Biodiesel is a mixture of fatty esters, currently produced by (trans-)esterification of tri-glycerides and free fatty acids, followed by several neutralization and purification steps. However, all the traditional methods suffer from drawbacks related to the use of liquid acid/base catalysts, heading to major economical and environmental penalties, especially considering the recent boost of the international biodiesel production rate. Worldwide, the production of biodiesel increased tremendously during the past 10 years, mostly in Asia, US, and Western Europe with Germany, France, Austria, Spain and UK among top consumers (Figure 1).

This work proposes a novel two-step biodiesel production process bases on reactive-separation using solid acid/base catalysts, thus simplifying the overall process and bringing significant benefits: high conversion and selectivity, elimination of conventional catalyst-related operations, no waste streams, as well as reduced capital investment and operating costs. The process design of the 10 ktpy fatty acid ethyl esters plant described here is based on experimental results, integrated in rigorous simulations performed using AspenTech AspenONE as computer aided process engineering tool.

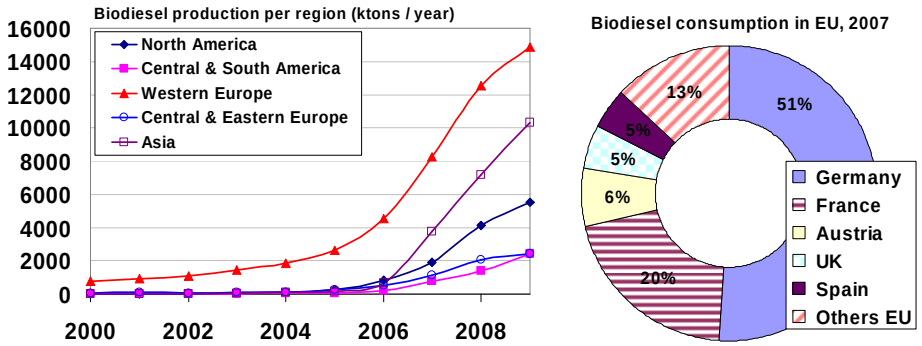


Figure 1. Biodiesel production per region (left), and biodiesel consumption in EU (right).

2. Problem statement

All conventional biodiesel production methods have associated optimal operating parameters and downstream processing steps, although much of the available literature emphasizes the base catalyzed route.[1-3] Traditional biodiesel processes employ liquid catalysts, such as H_2SO_4 or KOH.[3] The problem is that such catalysts demand neutralization, separation, washing, recovery, and salt waste disposal operations with serious economical and environmental consequences.

Nowadays, the surplus of waste oil available at industrial scale would allow production of very cheap biodiesel – a key benefit in the energy market. For example, in Brazil alone, more than 350 millions litres of biofuel are produced annually from animal fat. The problem with the animal fat or waste-oil, is that it becomes useless within 24 hours since it turns so acidic due to the increased free fatty acids (FFA) content, that it is more appropriate for making soap than for biodiesel. To solve these problems, we propose a sustainable two-step process based on the esterification of FFA's in a separative reactor using solid acids,[4,5] followed by trans-esterification of the remaining tri-glycerides (TG) using conventional or solid base catalysts.

1: R-COOH + EtOH ↔ R-COO-Et + H_2O (esterification)

2: TG + 3 EtOH ↔ 3 RCOO-Et + Gly (trans-esterification)

The integrated reactive distillation equipment proposed in this work is able to shift the chemical equilibrium and drive the esterification reaction to completion by continuously removing the fatty esters products and water by-product.[6,7] The raw materials consist of waste-oil or animal fat – mainly a mixture of free fatty acids – and a light alcohol, such as methanol or (bio-)ethanol. A key feature of this work is the replacement of anhydrous ethanol by its hydrous azeotrope, thus leading to further reduction of production costs.

Table 1 presents an overview of the available solid acid and base catalysts for biodiesel production by (trans-)esterification.[4-6] In this work we selected the metal oxides as solid acid catalysts for FFA esterification (first step) and calcium ethoxide as solid base catalyst for the trans-esterification of the remaining tri-glycerides (second step).

Table 1. Advantages and disadvantages of the acid/base catalysts tested for (trans-)esterification.

Catalyst type	**Benefits**	**Drawbacks**
Ion-exchange resins (Nafion, Amberlyst)	Very high activity Easy regeneration	Low thermal stability Possible leeching
TPA ($H_3PW_{12}O_{40}$)	Very high activity	Soluble in water
TPA-Cs ($Cs_{2.5}H_{0.5}PW_{12}O_{40}$)	Super acid sites	Low activity per weight
Zeolites (H-ZSM-5, Y and Beta)	Controlable acidity and hydrophobicity	Small pore size Low activity
Sulfated metal oxides (zirconia, titania, tin oxide)	High activity Thermally stable	Deactivates in water, but not in organic phase
Niobic oxide (Nb_2O_5)	Water tolerant	Average activity
Calcium oxide / CaO	Low temperatures	Long reaction times
Calcium methoxide / $Ca(OMe)_2$	High yield, reusable	High reactants ratio
Calcium ethoxide / $Ca(OEt)_2$	High yield, short times	High reactants ratio
Li-dopped zinc oxide / ZnO	Low temperatures	Long reaction times
KF loaded on Eu_2O_3	Short reaction times	Incomplete yields

3. Simulation methods

The simulation methods available are given in Table 2. Each method has important benefits but also certain drawbacks and the requirements can differ significantly. The amount of data required by the rigorous method is practically not feasible in practice while the shortcut method leads to low-fidelity models only, with limited applications. For practical reasons, the hybrid approach gives the best results. In this work the experimentally determined kinetic parameters were used [5,6] but the fatty components were lumped into one fatty acid/ester compound, according to the reaction:

$$\text{R-COOH} + \text{EtOH} \leftrightarrow \text{R-COO-Et} + H_2O$$

Table 2. Simulation methods for biodiesel production: requirements, benefits and drawbacks.

	Rigorous method	**Shortcut method**	**Hybrid method**
Requirements	Properties for all species. VLL data and BIP's for all pairs of components. Kinetic parameters for all reactions possible.	Properties for single fatty acid/ester/tri-glyceride. VLL data for the system ester/glycerol/alcohol. Asumed conversion (no kinetic parameters).	Single or reduced list of fatty acid/ester/TG. Short list of VLL data and BIP's for components. Reduced list of kinetic parameters, few reactions.
Benefits	Easy optimization of reaction and separation. High fidelity model. Usable for many plants. Easy comparison for various feedstocks.	Simple model. Fast simulations. Easy-to-build mass and energy balance. No data needed for all species present.	Optimization possible for reaction and separation. Certain ability to compare various feedstocks. Better model fidelity. Fast simulations for RTO.
Drawbacks	Slow simulations and convergence problems. Expensive measurements. Limited RTO and model based control usage.	No comparison possible for various feedstocks. Low-fidelity model. Less ability to use RTO.	More effort to build component list and get kinetic parameters. More work to find VLL data and regress BIP's.

4. Results and discussion

The properties of the fatty components were determined experimentally, or estimated using state-of-the-art contribution methods such as UNIFAC – Dortmund modified.[5-9]

Figure 2 (left) shows the residue curve map (RCM) for the ethanol-water-glycerol ternary mixture. The presence of the ethanol-water azeotrope does not hinder the biodiesel production process, since ethanol is a reactant and not a high purity product. Moreover, as water by-product from the esterification reaction further dilutes the ethanol, its hydrous azeotrope can be used directly as a lower-cost feedstock.

Vapor pressure is perhaps one of the most important properties with a critical effect in modeling reactive separations. Figure 2 (right) shows the vapor pressure of most common fatty acids and esters. At ambient pressure the boiling points are relatively high, exceeding 300 °C. Although high purity products are possible by reactive distillation, the high temperature in the reboiler – caused by the high boiling points, is in conflict with the thermo-stability of the biodiesel product. However, this problem can be avoided by working at lower pressure or by allowing ethanol in the bottom product.

Figure 3 presents the flowsheet of a two-step biodiesel production process based on a reactive distillation column (RDC) as the key unit for esterification or pre-treatment of free fatty acids (FFA). The process proposed here was rigorously simulated and optimized using AspenTech AspenONE. The production rate considered for the plant designed in this work is 10 ktpy fatty acid ethyl esters (FAEE). Note that the kinetic parameters used in the simulation were previously reported in the open literature.[5,6]

The RDC is operated in the temperature range of 100–250 °C, at ambient pressure. Out of the 15 stages of the integrated unit, the reactive zone is located in the middle of the column (10 stages). The fatty acid is pre-heated then fed as hot liquid on top of the reactive zone while a stoichiometric amount of alcohol introduced in the bottom of the reactive zone, thus creating a counter-current V-L flow regime over the middle reactive section. The reflux ratio is very low (RR=0.1) as returning water to the column is detrimental to the chemical equilibrium. Water by-product is removed in top, then separated in a decanter from which only the fatty acids are recycled to the column while water is recovered at high purity and hence reusable as industrial water on the same site.

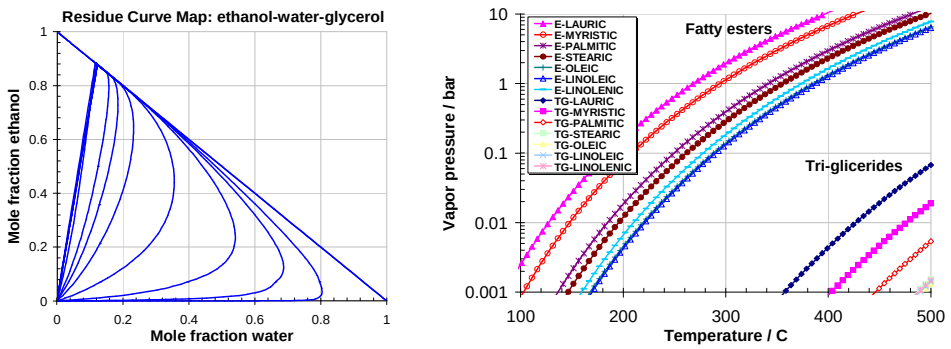


Figure 2. RCM ethanol-water-glycerol (left), Vapor pressure of fatty esters vs temperature (right).

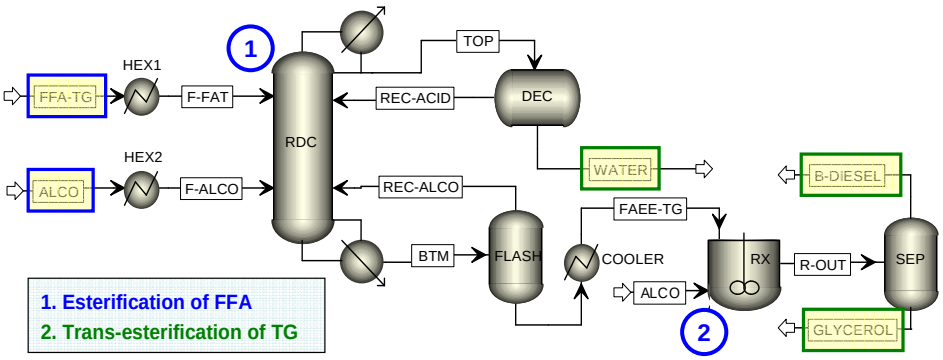


Figure 3. Flowsheet of biodiesel production by catalytic reactive distillation.

The fatty esters are delivered as high-purity bottom product of the RDC. The hot product is flashed first to remove the traces of ethanol, then it is send to the trans-esterification reactor to further convert the remaining tri-glycerides to fatty esters. In this work we considered the worst case scenario, with 100% FFA in the feedstock.

The mass and energy balance is given in Table 3. High purity products are possible, the purity specifications exceeding 99.9%wt for the final biodiesel product (FAEE stream). Note that the total amount of the recycle streams (REC-ACID and REC-ALCO) is not significant, representing only ~0.5% of the total biodiesel production rate.

Figure 4 (left) shows the liquid composition profiles in the reactive distillation column. The concentration of fatty acids decreases while the concentration of fatty esters increases from the top to bottom. Similarly, the ethanol concentration decreases while water concentration increases from bottom to top. The temperature and reaction rate profiles in the RDC are presented in Figure 4 (right). As expected, the reaction rate exhibits a maximum in the middle of the column, in the reactive zone. Moreover, the concentration of water is low in the reactive zone, hence the catalyst activity is not affected. Nevertheless, the concentration of reactants is relatively high and the temperature is sufficiently high to allow high reaction rates and complete conversion.

Table 3. Mass and energy balance of a 10 ktpa biodiesel production process, based on RD.

	F-ACID	F-ALCO	BTM	TOP	REC-ACID	REC-ALCO	FAEE	WATER
Temperature K	418.1	352.2	507.9	372.7	303.1	372.6	303.1	303.1
Pressure atm	1.036	1.036	1.017	0.987	1	1	1	1
Vapor Frac	0	0	0	0	0	0	0	0
Mass Flow kg/hr	1094.918	264.202	1257.312	109.246	0.152	7.312	1250	109.094
Volume Flow l/min	22.986	6.156	28.953	1.983	0.003	0.154	23.774	1.839
Enthalpy Gcal/hr	-0.892	-0.394	-0.886	-0.405	0	-0.006	-1.035	-0.413
Mass Flow kg/hr								
ETHANOL	0	253.568	2.622	0	0	0.857	1.765	0
ACID	1094.918	0	0	0.115	0.107	0	0	0.008
WATER	0	10.635	0.029	109.089	0.003	0.013	0.016	109.086
ESTER-E	0	0	1254.661	0.043	0.042	6.442	1248.219	0
Mass Frac								
ETHANOL	0	0.96	0.002	0	0	0.117	0.001	0
ACID	1	0	0	0.001	0.702	0	0	0
WATER	0	0.04	0	0.999	0.02	0.002	0	1
ESTER-E	0	0	0.998	0	0.278	0.881	0.999	0

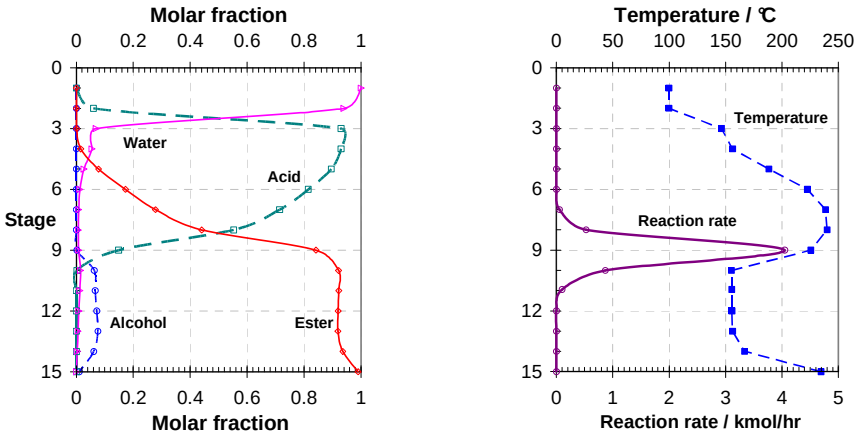


Figure 4. Profiles in RDC: liquid composition (left), temperature and reaction rate (right).

5. Conclusions

An innovative two-step biodiesel process based on separative reactors using solid catalysts was developed in this study using computer aided engineering tools such as AspenTech AspenONE. The novel two-step process proposed here improves considerably the biodiesel production and reduces drastically the number of downstream processing steps. The major benefits of this sustainable process are:

- Flexible integrated reactor suitable for a large range of fatty raw material with up to 100% FFA content, such as: frying oils, animal tallow, tall oil, waste vegetable oil.
- Straightforward and robust process with no soap formation, no catalyst-related waste streams, and sulfur-free biodiesel as solid acids do not leach into the product.
- Effective use of the integrated reactor volume leading to high unit productivity.
- Efficient use of the raw materials: complete conversion and high selectivity, stoichiometric reactants ratio, FFA conversion to esters and not to soap waste.
- Reduced equipment costs, with up to ~40% savings on the total investment costs.
- Competitive operating costs due to the integrated design and the elimination of conventional steps: handling of homogeneous catalyst and corrosive solutions, separation and disposal of salts, waste water treatment, excess alcohol recovery.

References

1. M. Balat, H. Balata, 2008, Energy Conversion and Management, 49, 2727.
2. M.G. Kulkarni, A. K. Dalai, 2006, Industrial & Engineering Chemistry Research, 45, 2901.
3. K. Narasimharao, A. Lee, K. Wilson, 2007, J. Biobased Materials & Bioenergy, 1, 19.
4. T. Okuhara, 2002, Chemical Reviews, 102, 3641.
5. A. Kiss, A. C. Dimian, G. Rothenberg, 2006, Advanced Synthesis & Catalysis, 348, 75.
6. A. Kiss, G. Rothenberg, A. C. Dimian, F. Omota, 2006, Topics in Catalysis, 40, 141.
7. A. Kiss, A. C. Dimian, G. Rothenberg, 2008, Energy & Fuels, 22, 598.
8. S. Steinigeweg, J. Gmehling, 2003, Ind. Eng. Chem. Res., 42, 3612.
9. C. Dimian, F. Omota, A. Bliek, 2004, Chem. Eng. & Proc., 43, 411.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal Scheduling under Variable Electricity Pricing and Availability

Pedro M. Castro,[a,b] Ignacio E. Grossmann,[b] Iiro Harjunkoski[c]

[a]*DMS, Inst. Nacional Engenharia Tecnologia e Inovação, 1649-038 Lisboa, Portugal*
[b]*Dep. Chemical Engineering, Carnegie Mellon University, Pittsburgh PA 15213, USA*
[c]*ABB Corporate Research Center, Wallstadter Str. 59, 68526 Ladenburg, Germany*

Abstract

This work addresses the scheduling of continuous plants subject to energy constraints related to time-dependent electricity pricing and availability. A Resource-Task Network continuous-time formulation is presented that can address these issues together with multiple intermediate due dates. To the best of our knowledge, it is the first model of its type that is able to effectively incorporate time-variable utility profiles. The validity of the approach is illustrated through the solution of a few example problems. When compared to a simple manual scheduling procedure, significant electricity savings can be achieved by switching production from periods of high to low electricity cost.

Keywords: energy, continuous-time formulation, decision-making tool.

1. Introduction

Enterprises are currently under pressure to produce at the lowest possible cost within continuously changing economic constraints. To achieve this goal, they must actively look at the best operating practices and optimize these both globally and locally. Within this overall goal, scheduling plays an important part.

This work is motivated by a real industrial problem. It involves the final stage of a multiproduct plant with electricity intensive parallel equipment units, where scheduling involves deciding when each unit has to produce a certain product. The products are then sent to storage units, where they remain until dispatching takes place. Meeting customer demands on time is vital, and for this reason, in some plants, no other factors are taken into account besides trying to keep the storage units full in order to be able to fulfill the orders. Plant scheduling is difficult due to the following factors: large combinatorial size arising from the number of equipment units, products and storage units; various operating and contractual constraints; liberalized electricity market with nontransparent billing practices. Due to the inherent complexity, the operator scheduling choices may be far from the optimal ones.

The most challenging aspect of plant scheduling is undoubtedly the incorporation of energy constraints related to electricity pricing and availability. We consider the case where in the planning stage, contracts are agreed between the electricity supplier and the plant, which often specifies maximum levels of power usage. If electricity consumption exceeds this threshold, the plant incurs in stiff penalties, whereas underproduction costs the same as planned production. Electricity cost varies significantly throughout the day, and this must be taken into consideration within the modeling framework. Another important aspect concerns meeting the sales forecasts, which typically occur at the end of each day.

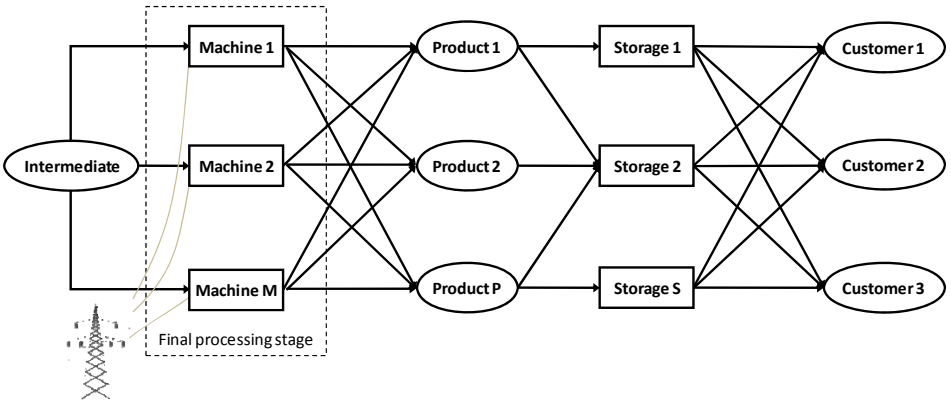


Figure 1. Final processing stage of industrial case study

Accounting for events that occur at predetermined points of time, being either a change in electricity cost level or the occurrence of a demand point, may be relatively easy or extremely complex, depending on the type of time representation employed. This is straightforward with a discrete-time approach. In contrast, with continuous-time, the timing of all event points is determined by the solver and thus it is much harder to relate the events with the points of change.

This paper presents a new continuous-time formulation that effectively handles time dependent cost parameters and discrete demand points. Incorporation of the former aspect within a continuous-time formulation has not been reported before to the best of our knowledge, whereas the constraints used to model the latter aspect are conceptually similar to those used by Maravelias and Grossmann (2003). The proposed approach builds on the general multipurpose formulation of Castro et al. (2004), which can address problems involving batch and continuous tasks, efficiently. Indeed, for batch plants, a study (2006) has found it to be the best single time grid formulation. Nevertheless, the constraints that are given are only suitable for continuous plants.

The new formulation is built on a unified framework for process representation, the Resource-Task Network (RTN) of Pantelides (1994). This means that the model variables and constraints are written is terms of abstract entities like resources, tasks and event points, and so it has a much wider scope than the single-stage industrial case study used for illustrative purposes.

2. Problem definition

During the final stage of the process, an intermediate material is transformed into one of different final products, characterized by chemical composition and particle size distribution, through the use of electricity. These are then sent to storage units, where they wait until customer dispatch takes place. This process is illustrated in Figure 1.

Typical plant schedules are established over one week, so this will be the time horizon (H=168 h) assumed. Let M represent the set of machines, P the set of products and S the storage units. Machines are characterized by: (i) power requirements $pw_{p,m}$ [MW]; (ii) processing rates, $\rho_{p,m}$ [ton/h]. Let DY and HR be the days of the week and the hours of the day, respectively. Each product may have multiple demands over the week, at any hour of the day, $d_{p,dy,hr}$ [ton]. Storage units have known maximum capacities, cap_s [ton].

The energy contract signed by the plant and electricity provider establishes a certain pricing policy. Electricity cost is typically lower during the night and higher during the day and we will be using the energy policy given in Duarte et al. (2009). It consists of

three energy levels, E, with prices c_e of 0.0481, 0.0945 and 0.2162 [€/kWh]. The maximum level of total power consumption is specified by parameter $pwx_{hr,dy}$ [MW].

3. Mathematical formulation

The proposed model uses the RTN framework to make it as general as possible. Before employing a RTN model it is required to describe the process and/or process recipe as a sequence of tasks involving certain resources. When this is done, one needs to determine the structural parameters, which pass the information about the structure to the mathematical model. Due to limited space we will be skipping this important modeling phase and moving directly to the mathematical formulation.

Tasks ($i \in I$) are characterized by two sets of extent variables, one binary, $N_{i,t}$, and one continuous, $\xi_{i,t}$. The former identify the start of task i at time point t, while the latter give the amount handled by the task. We are implicitly assuming that the tasks are either instantaneous (start and end at t) or continuous, which can be made to last a single time interval without loss of generality (Castro et al., 2004).

The continuous-time formulation relies on a single time grid to keep track of events taking place. It uses $|T|$ event points that can be placed anywhere between the origin and end (H) of the time horizon. Traditionally, tasks starting at event point t have been assumed to start at the absolute time determined for that event point, T_t. However, the tasks have total freedom to start anywhere provided that they end before T_{t+1}. The advantage of not requiring tasks to start (end) exactly at a time point is the reduction on the number of event points required to represent a solution, which seriously affects computational effort.

Consider a simple example involving the execution of a single processing task at slot/interval t, see Figure 2. Based on the problem data, one can build an algorithm to generate: (i) demand points ($td \in TD$) and respective timing tfx_{td}; (ii) time periods of constant cost for electricity level e ($tp \in TP_e$), and respective starting and ending times, $lb_{e,tp}$ and $ub_{e,tp}$ [h]. In cases where the duration of the task is lower than the duration of the lowest cost period located inside the interval, the interval boundaries shown, T_t and T_{t+1}, can result from two consecutive demand points. Variable Ts_t will then give the starting time of the task executed during interval t (or the earliest starting time amongst all tasks executed). Note that no further event points are required.

The fundamental timing constraint states that the difference in time between two consecutive event points must be greater than the duration of the task taking place, eq 1. It is written for all equipment resources involved in the timing constraints (R^{TC}) and implicitly assumes that there can only be one task executed in such equipment unit at a certain time. That is ensured by the initial resource availability and the excess resource balances, which are not shown here. The novel constraints are given next. Eq 2 is used to guarantee that the tasks are fully executed within time interval t, while eq 3 ensures that tasks start no sooner than the time of the interval's lower boundary.

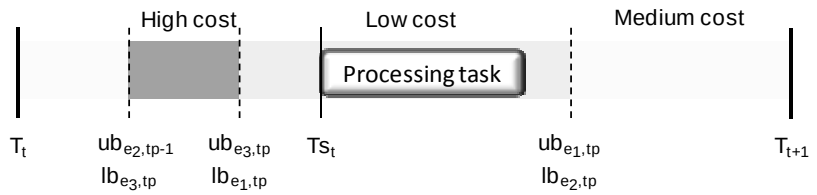


Figure 2. Continuous tasks executed at interval t do not necessarily start at event point t

$$T_{t+1} - T_t \geq \sum_{i \in I^c} \frac{\overline{\mu}_{r,i} \xi_{i,t}}{\rho_i^{\max}} \quad \forall r \in R^{TC}, t \in T, t \neq |T| \tag{1}$$

$$T_{t+1} - Ts_t \geq \sum_{i \in I^c} \frac{\overline{\mu}_{r,i} \xi_{i,t}}{\rho_i^{\max}} \quad \forall r \in R^{TC}, t \in T, t \neq |T| \tag{2}$$

$$Ts_t \geq T_t \quad \forall t \in T, t \neq |T| \tag{3}$$

We now define a new binary variable, $Y_{t,tp,e}$, which takes the value of one if, during interval *t*, tasks are processed within time period *tp* of energy level *e*. Similarly, binary variable $Y_{t,td}^{out}$, identifies whether or not event point *t* corresponds to demand point *td*. It can be assumed without loss of generality that there is: (i) one active time period of an energy level during *t*; (ii) one event point associated to due date *td*, eqs 4-5. If time interval *t* is located within time period *tp* of energy level *e*, then the starting time of tasks must be greater than the time period lower bound, eq 6. Likewise, they must end before its upper boundary (see Figure 2). Eq 7 is a big-M constraint that is only active if there is a continuous task being executed that belongs to energy level *e*. Finally, if event point *t* corresponds to demand point *td*, then the time values must match (eqs 8-9).

$$\sum_{e \in E} \sum_{tp \in TP_e} Y_{t,tp,e} = 1 \quad \forall t \in T, t \neq |T| \tag{4}$$

$$\sum_{t \in T \wedge t \neq 1} Y_{t,td}^{out} = 1 \quad \forall td \in TD \tag{5}$$

$$Ts_t \geq \sum_{e \in E} \sum_{tp \in TP_e} lb_{e,tp} Y_{t,tp,e} \quad \forall t \in T, t \neq |T| \tag{6}$$

$$Ts_t + \sum_{i \in I_e^c} \frac{\overline{\mu}_{r,i} \xi_{i,t}}{\rho_i^{\max}} \leq \sum_{tp \in TP_e} ub_{e,tp} Y_{t,tp,e} + H \cdot (1 - \sum_{i \in I_e^c} \overline{\mu}_{r,i} N_{i,t}) \quad \forall r \in R^{TC}, t \neq |T|, e \in E \tag{7}$$

$$T_t \geq \sum_{td \in TD} tfx_{td} Y_{t,td}^{out} \quad \forall t \in T, t \neq 1 \tag{8}$$

$$T_t \leq \sum_{td \in TD} tfx_{td} Y_{t,td}^{out} + H \cdot (1 - \sum_{td \in TD} Y_{t,td}^{out}) \quad \forall t \in T, t \neq 1 \tag{9}$$

The objective function is given in eq 10, which maximizes the total electricity cost [k€]. This is given by the sum over all electricity levels *e*, time intervals *t* and tasks *i*, of the product of electricity cost c_e [€/kWh], power consumption [MW], and task duration [h].

$$\min \sum_{e \in E} \sum_{i \in I_e^c} \sum_{r \in R^{UT}} \sum_{t \in T \wedge t \neq |T|} c_e \cdot (-\mu_{r,i}) \cdot \frac{\xi_{i,t}}{\rho_i^{\max}} \tag{10}$$

4. Computational Results

The performance of the model is illustrated through the solution of 5 test cases. Directly related to the complexity is the number of products, machines and storage units (|P|, |M|, |S|). Due to the differences in electricity cost among the energy levels, assigning production to the lower levels will have the biggest impact on the total cost. The resulting mixed-integer linear programming (MILP) problems were solved by CPLEX 11.1 up to a relative optimality tolerance=10^{-6} on an Intel Core2 Duo T9300 (2.5 GHz) laptop running Windows Vista Enterprise. The results are listed in Table 1.

The optimal solution for EX1 corresponds to a total electricity cost of €18,625. Eight out of 10 task instances are executed in low-cost periods with the remaining being in medium-cost levels. Enough power was provided to operate the machine throughout the time horizon. However, power shortages may force tasks to switch to a higher cost level. In order to show that the model can cope with constraints of this type, EX2 was solved. It is uses the same data of EX1 but now the maximum power consumption during the first seven hours of Tuesday and Thursday is set to zero. The cost increases by 15.8%, to €21,575, which is very significant considering such small changes in data. It serves to illustrate the impact that advanced scheduling tools can have on plant profitability.

Nevertheless, we do have to say that the continuous-time formulation is limited to small problems. Even for EX5, which features three products, two machines and two storage units, we had to specify low product demands (leading to a 33% capacity) to ensure that most tasks could fit into a single period of constant electricity cost. In this way, a good solution could still be found with a relatively small number of event points. However, it took already more than 1 hour to prove optimality for $|T|=11$ (€26,911), a value that is still 0.5% above the true optimum (found by a discrete-time formulation). The schedule and corresponding storage profiles are given in Figure 3. For larger problems, CT is intractable due to the following facts: (i) the computational effort is strongly dependent on |T|, with experience telling us that we get typically a one order of magnitude increase for a single increase in |T|; (ii) it is not straightforward to find a number that ensures feasibility; (iii) it may be even difficult for the solver to find out that the value of $|T|$ is insufficient, i.e. finding that the problem is infeasible.

Table 1. Computational results

problem	(P,M,S)	|T|	DV	SV	EQ	RMIP (€)	MIP (€)	CPU s	nodes
EX1	(2,1,1)	11	490	797	347	16620	**18625**	8.53	4252
		12	539	875	380	16620	18625	9.8	5442
EX2	(2,1,1)	11	490	797	347	16620	**21575**	26	20971
		12	539	875	380	16620	21575	21.6	15474
EX3	(2,1,2)	11	510	934	420	16620	**18153**	237	69046
		12	561	1025	460	16620	18153	1372	467962
EX4	(2,1,2)	11	510	934	420	18896	21349	178	46512
		12	561	1025	460	18896	**21175**	1470	237314
EX5	(3,2,2)	9	528	1089	562	25625	27222	7.18	3989
		10	594	1221	629	25625	27008	369	138426
		11	660	1353	696	25625	**26911**	4131	1295540

5. Conclusions

This paper has focused on the modeling of discrete events that occur at predetermined points in time with a continuous-time scheduling formulation. These included multiple intermediate due dates, utility availability and variable electricity costs. New sets of

constraints have been proposed that are part of a general model that relies on the Resource-Task Network process representation. The validity of the approach has been demonstrated on a few test cases adapted from a real industrial problem. Despite the major modeling breakthrough, the results have shown that only problems of small size can be handled effectively. More importantly, the paper has highlighted the importance of taking variable electricity costs into consideration when deriving the schedule. State-of-the-art scheduling formulations have the potential to achieve major savings when compared to procedures that are mostly focused on feasibility.

References

P. Castro, A. Barbosa-Póvoa, H. Matos, A. Novais, 2004, Ind. Eng. Chem. Res., 43, 105.

B. Duarte, L. Santos, J. Mariano, 2009, Computers Operations Research, 36, 1825-1834.

C. Maravelias, I. Grossmann, 2003, Proceedings ESCAPE-13 (Eds. A. Kraslawski and I. Turunen), pp. 215.

C. Pantelides, 1994, In Proceedings of the Second Conference on Foundations of Computer Aided Operations; Cache Publications: New York, 1994; pp 253.

M. Shaik, S. Janak, C. Floudas, 2006, Ind. Eng. Chem. Res, 45, 6190.

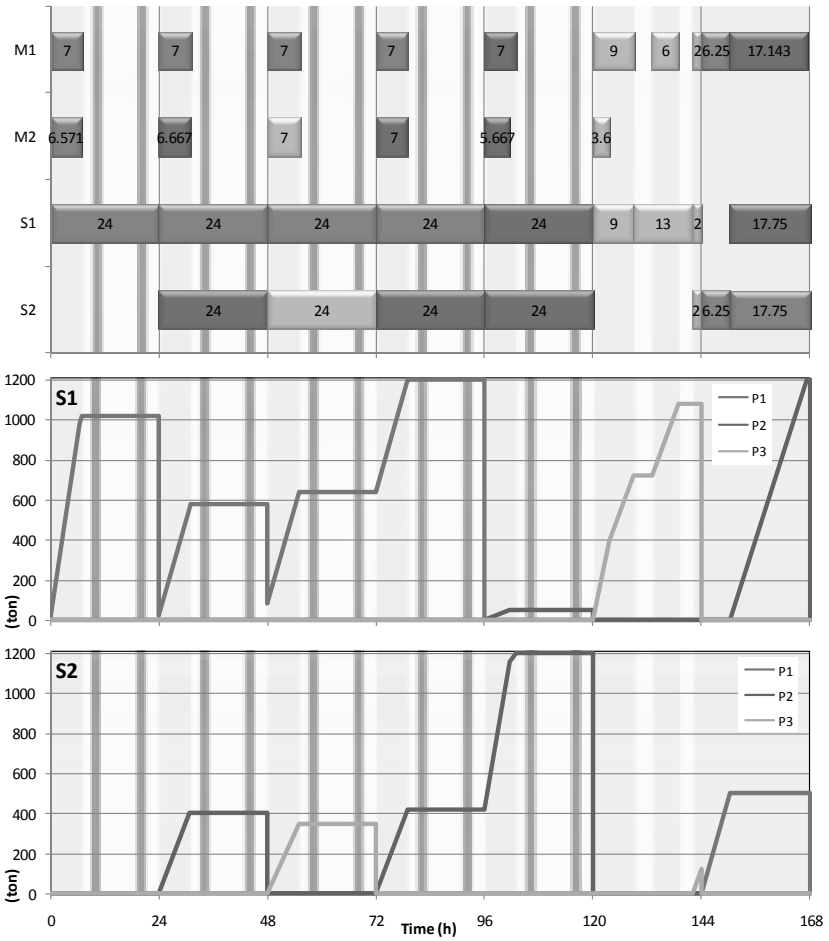


Figure 3. Best solution found for EX5.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Effect of substrate specific area on lignocellulose enzymatic hydrolysis: an experimental and modeling investigation

Chiara Piccolo[a], Gunnar Lidén[b], Fabrizio Bezzo[a]

[a] *DIPIC-Dipartimento di Principi e Impianti di Ingegneria Chimica, Università di Padova, via Marzolo 9, I-35131, Padova, Italy (chiara.piccolo@unipd.it)*
[b]*Department of Chemical Engineering II, Lund Institute of Technology, P.O. Box 124, 221 00 Lund, Sweden*

Abstract

Lignocellulose hydrolysis is a critical step in the enzymatic process for the production of bioethanol. Its high variability and dependence on materials and process parameters makes it very difficult to be described and optimized through a reliable modeling approach. In this work an experimental investigation has been carried out to assess the effect of the substrate specific area on the kinetics of lignocellulose hydrolysis. Data were used to estimate the parameters of a modified Langmuir adsorption model, embedding the accessible surface area as a critical parameter.

Keywords: surface area, adsorption experiments, Langmuir adsorption.

1. Introduction

Kinetics and further modeling studies of hydrolysis are useful in a different stages of processing of biomass to fermentable sugars. Mechanistic models are developed from the reaction mechanism, mass transfer consideration and other physical parameters that affect the extent of hydrolysis, and they vary in their complexity based on the intended use of the models [1].

Extensive research has been made to identify parameters that affect the rate of lignocellulose hydrolysis. Two main categories can be identified depending on whether parameters are mainly related to the substrate characteristics (lignin distribution, accessible surface area, particles size, crystallinity, degree of polymerization) or to the enzyme features (adsorption of enzyme prior to reaction, end-product inhibition, thermal and shear force inactivation, synergism, mass transfer limitation affecting the transport of the enzyme to the substrate) [2].

In the past 50 years, there has been a constant influx of research publications addressing the enzymatic kinetics of cellulose degradation. However, the kinetics of cellulose degradation is still not fully understood because of different competing effects that can hardly be distinguished from each other and that introduce large bias and variability in the estimation of kinetic parameters.

In the current work, an emphasis is given to the effect of the substrate specific surface area on the enzyme adsorption onto the substrate, which represents the essential prerequisite step of the hydrolysis process. In previous works it has been shown that the rate of adsorption is rapid compared to the actual hydrolytic activity of the enzymes,

thus making the amount of adsorbed cellulase an important factor in the effectiveness of the reaction [3]. The most common description of cellulase adsorption is the Langmuir isotherm (Eq. 1 in the following), derived assuming that adsorption may be described in terms of a single adsorption equilibrium constant and a specified adsorption capacity [4]. Embedding the effect of the specific surface in the model parameters accounting for the accessibility of the substrates represents a first attempt to deal with the analysis of adsorption in spatial terms so as to understand cellulose hydrolysis at a mechanistic level.

A three-step approach was adopted to investigate the phenomenon:

- surface area measurement and pores size distributions were determined for three different substrates through BET-technique;
- adsorption isotherms and hydrolysis experiments were performed;
- a modified Langmuir model was developed and adsorption equilibrium data and hydrolysis data were used to estimate the parameter of the model.

2. Materials and Methods

2.1. Enzymes

A commercial cellulase mixture (Celluclast 1.5L provided by Novozymes A/S, Bagsvaerd, Denmark) was used in the experiments. The cellulase enzyme had an activity of 46.8 FPU/ml.

2.2. Substrates

A comparative study was carried out on three different substrates: Avicel, and spruce pre-treated at two different conditions.

- Avicel (i.e. microcrystalline cellulose) was purchased from Sigma-Aldrich.
- Softwood hydrolyzates were prepared by steam explosion with SO_2 impregnation from spruce chips at Lund University. Two different treatment conditions were used: a) temperature 210°C, SO_2 content 2.5% (w/w) and residence time 5 min. The composition as determined by NREL standard procedure was 46.7% glucan, 1.9% mannan, 1.6% xylan, 1.2% galactan and 1.2% arabinan. Acid insoluble lignin was 42.0%; b) temperature 190°C, residence $time_e$ 10 min with the same SO_2 content as before. NREL composition analysis provided these results: 48.6% glucan, 2.9% mannan, 2.1% xylan, 1.2% galactan and 1.2% arabinan. Acid insoluble lignin was 35.7%.

The first pretreatment condition was known to produce hydrolyzate which more easily can be enzymatically degraded. Both pretreated spruce substrates were washed with distilled water and stored at 4°C.

2.3. BET measurements

This procedure, often used for the determination of the specific surface area of inorganic catalysts, is based on Nitrogen adsorption onto the material surface at different pressures.

Surface area, total pore volumes, at a relative pressure of about 0.994, and pore size distributions were determined on a Micrometrics ASAP 2400 after degassing for 10 days at 40°C.

Before degassing pretreated spruce materials were dried. Different drying conditions were chosen in order to investigate the impact on biomass inner structure of the dewatering process: samples were dried for 48 h in the oven at 105°C, and for 5 days at room temperature. Avicel, stored at room temperature, was straight subject to degassing.

2.4. Adsorption isotherms studies

Adsorption isotherms studies were conducted at constant reaction time and constant temperature with different enzyme-substrate ratios, for the three substrates whose BET surface had been assessed (i.e. Avicel and spruce pretreated at optimal and milder conditions). Different enzyme solutions were added with a constant amount of substrate in 50 mL plastic Falcon test tubes, to obtain a dry matter loading at 10g/L. The pH was adjusted at 4.8. After 90 minutes of incubation (60-90 minutes is the time needed to achieve maximal adsorption at the given conditions) at the selected temperature, the supernatant was collected, centrifuged and filtered. Cellulase residual activity in solution was assessed using the Bradford method for protein content determination. The experiments were performed first at 4°C to avoid changes in substrate properties due to extensive hydrolysis. The amount of bound cellulase is calculated from the difference between the initial cellulase or protein concentration and final free cellulase or protein concentration

2.5. Hydrolysis studies

Hydrolysis experiments were performed in 300ml baffled Erlenmeyer flasks, at pH 4.8 and 30°C, in distilled deionized water, using Avicel, the optimal pretreated spruce, and the non optimal pretreated spruce as substrate. The WIS (Water Insoluble Solids) content was 10g/L (i.e. about 1%w/v).

Two different enzyme loads were tested: 2ml and 0.067ml, corresponding roughly to 46.84 FPU/g WIS and 1.57 FPU/g WIS. Samples were withdrawn during the first 30 hours, centrifuged and the concentration of hydrolysis products was measured through HPLC.

3. Modeling approach

The most common description of cellulase adsorption is the Langmuir isotherm, derived assuming that adsorption can be described by a single adsorption equilibrium constant and a specified adsorption capacity [4]. The Langmuir isotherm may be represented as:

$$E_a = \frac{W_{max} K_P E_f}{1 + K_p E_f} \tag{1}$$

where E_a is the adsorbed cellulase (mg cellulase/L); W_{max} is the maximum cellulase adsorption = A_{max}*C (mg cellulase/L) with A_{max} the maximum cellulase adsorption per g cellulose (mg cellulase / g cellulose) and C the cellulose concentration (g cellulose/L); E_f is the free cellulase (mg cellulase/L); and K_P is the dissociation constant ($K_p = \frac{E_a}{E_f C}$) in terms of L/g cellulose.

The focus of the current study is not to propose a new phenomenological adsorption model but rather to present a models which is as simple as possible with physically meaningful parameters Although very complex and detailed models could be envisaged, the real issue is to produce reliable experimental data that can be used to properly identify them.

Accordingly, the structure of the Langmuir equation, which is widely used because it provides a good (or often very good) fit to experimental data and represents a simple mechanistic model that can be used to compare kinetic properties of various cellulase–cellulose systems, was maintained. Specific surface areas and a measure of pore accessibility were embedded in the model. Variables and parameters on Eq. 1 are thus

changed as follows: E_a is the adsorbed cellulase (mg cellulase/g WIS); W_{max} is the maximum cellulase adsorption (mg cellulase/g WIS) defined as:

$$W_{max} = A_{max} S_a \alpha \tag{2}$$

with A_{max} the maximum cellulase adsorption/m^2 on the substrate (mg cellulase / m^2 substrate), S_a the substrate specific area (m^2/g WIS) and α is the fraction of pores accessible to enzyme as calculated from pores size distribution; E_f is the free cellulase (mg cellulase/L); and K_P is the dissociation constant ($K_P = \frac{E_a}{E_f}$).

3.1. Estimation of the Model Parameters

A least squares regression was used to estimate the model parameters from the adsorption and hydrolysis experimental. In particular, in order to fit the glucose profiles, the kinetics proposed by Philippidis et al. [5, 6] was considered and adjusted to the new model variables definition. According to the Philippidis' model, based on the principles of heterogeneous catalysis, the reaction rate depends on the concentration of adsorbed cellulase (E_a) through:

$$r = k^1 E_a \tag{3}$$

where k^1 is a lumped specific rate constant [h^{-1}], which exhibits a Michaelis-Menten dependence on the cellulase concentration according to the following expression:

$$k^1 = \frac{kE}{K_{eq} + E} \tag{4}$$

where k is the maximum specific cellulose hydrolysis rate [h^{-1}], E is the total (free and bound) concentration of the enzyme and K_{eq} is the cellulase enzyme saturation constant [FPU/L]. The values of k and K_{eq} used for the estimation of the adsorption isotherm parameters were taken from Philippidis et al. [6].

4. Results and Discussions

4.1. BET measurements

Due to variations in the experimental conditions such as adsorption time, vacuum time and vacuum pressure, sample preparation and sample origin and features, a wide range of gross area values have been reported in the literature even for the same substrate. Typical superficial surface areas for Avicel exhibit a wide variability (in the range 1.8-22 m^2/g) [2]. For the steam pretreated spruce material, however, no previous measurements were found in literature. The surface area values, determined through BET procedure, are reported in Table 1.

Table 1. BET specific surface area values and fraction of the accessible surface determined for the substrates of interest.

Substrate	Drying at room T [sq. m/g]	Drying in the oven [sq. m/g]	Average value used in model identification	Average fraction accessible surface area
Avicel	1.1±0.0		1.1	0.453
Opt. pret. spruce	2.0±0.1	2.4±0.1	2.2	0.680
Non optimal pret. spruce	1.6±0.1	1.2±0.0	1.4	0.660

*Error of measurement ± 0.9%

For the "optimally" pretreated spruce, the BET surface areas were determined both for material dried at room temperature and material dried in the oven (105 °C). The values were only slightly different, i.e. the drying temperature does not have big impact on substrate structure, at least with concern to the surface area. The analysis of the pore size distributions, measured through the same procedure, is important in determining the fraction of pores accessible to the enzymes. Only pores with a diameter larger than 54Å are accessible to cellulose enzymes [7]. The fraction of accessible area, calculated from PSD data are also reported in Table 1.

4.2. Adsorption and hydrolysis experiments results

After incubation of various amounts of cellulase enzymes with a fixed amount of substrate, the free protein concentration was determined. From these data the adsorption isotherms (ratio amount of enzyme adsorbed/amount of enzyme added) were constructed **for** the different substrates (Fig. 1a). The error bars show the standard deviation of the data calculated from repeated experiments and analysis. By comparing the isotherms it can be observed that at the same WIS content the extent of adsorption is higher for the substrate characterized by the bigger BET surface area, i.e. the optimal pretreated spruce. The lowest amount of enzyme adsorbed was observed for Avicel, i.e. the substrate with the smallest surface area.

The profiles of glucose released during hydrolysis experiments at different enzyme loads (E1=46.84 FPU/g WIS and E2=1.57 FPU/g WIS) are shown in Fig. 1b.

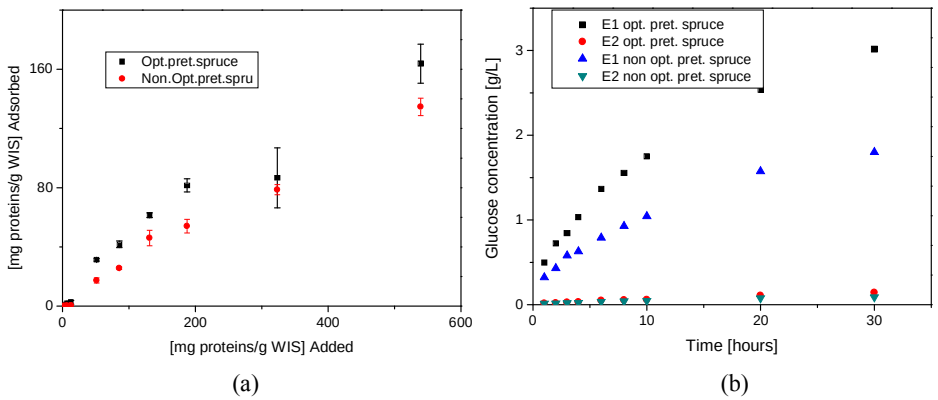


Figure 1 Adsorption isotherms (a) and glucose rate of release (b) for hydrolysis of pretreated spruce substrates.

4.3. Model identification

Data on free and bound enzyme concentration and glucose concentration were used to estimate the equilibrium adsorption constants, K_p and A_{max}. The values for the parameters are given in Table 2

Table 2. Parameter values for adsorption on pretreated spruce.

Parameter	Estimate	*t*-value	t_{ref}	Confidence intervals 95%	χ^2-value	χ^2_{ref}
A_{max}	92.208	7.511	1.653	12.204	229.66	216.53
K_p	0.0065	6.890		0.0009		

5. Final remarks

The investigation has highlighted the importance of the substrate specific surface area in the hydrolysis process for lignocellulosic materials. The experimental data were used to estimate the parameters of a modified Langmuir-type adsorption model taking into account the effect of the specific surface. Future work will aim at incorporating within a modeling framework the extent of productive and unproductive binding of enzyme, which can be assessed through a combined analysis of adsorption data and dynamic hydrolysis experiment data, performed on substrates characterized by sensibly different lignin content.

6. Acknowledgement

C.P. and F.B. gratefully acknowledge Fondazione Cariparo for Progetto Dottorati di Ricerca 2006 under whose framework this research has been carried out. Partial support from the University of Padova under Progetto di Ateneo 2007 (cod. CPDA071843): "Bioethanol from lignocellulosic biomass: process and equipment development" is gratefully acknowledged.

References

[1] S. Peri, S. Karra, Y.Y. Lee, M.N. Karim, 2007, Modeling intrinsic kinetic of enzymatic cellulose hydrolysis, Biotechnol. Progr., 23, 626-637.

[2] Y.H.P. Zhang, L.R. Lynd, 2004, Toward an aggregated understanding of enzymatic hydrolysis of cellulose: Noncomplexed cellulase systems, Biotechnol. Bioeng., 88, 797-824.

[3] W. Steiner, W. Sattler, H. Esterbauer, 1988, Adsorption of *Trichoderma reesei* cellulase on cellulose-experimental-data and their analysis by different equations, Biotechnol. Bioeng., 32, 853–865.

[4] B. Nidetzky, W. Steiner, M. Claeyssens, 1994, Cellulose hydrolysis by the cellulases from *Trichoderma reesei*: adsorptions of two cellobiohydrolases, two endocellulases and their core proteins on filter paper and their relation to hydrolysis, Biochem. J., 303, 817-823.

[5] G.P. Philippidis, D.D. Spindler, C.E. Wyman, 1992, Mathematical modeling of cellulose conversion to ethanol by simultaneous saccharification and fermentation process, Appl. Biochem. Biotechnol., 34-35, 543-556.

[6] G.P. Philippidis, C. Hatzis, 1997, Biochemical engineering analysis of critical process factors in the biomass-to-ethanol technology, Biotechnol. Progr., 13, 222-231.

[7] H.E. Grethlein, 1985, The effect of pore size distribution on the rate of enzymatic hydrolysis of cellulose substrates, Bio/ Technol., 3, 155–160.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

An outer approximation algorithm for the global optimization of regulated metabolic systems

Gonzalo Guillén-Gosálbez[a*], Carlos Pozo[a], Laureano Jiménez[a], Albert Sorribas[b]

[a] *Department of Chemical Engineering, University Rovira i Virgili, Tarragona, Spain*
[b] *Departament de Ciències Mèdiques Bàsiques, Universitat de Lleida, Lleida, Spain*

Abstract

Understanding the evolution of cellular metabolism requires a number of techniques able to deal with its complexity. Adaptive responses observed in evolutive studies are expected to consist of an optimal set of changes in enzymes activities fulfilling important physiological constraints. Within this context, we present a novel approach to identify enzyme activity regions that contain feasible biological responses in evolution. The framework presented also allows to optimize the enzyme activity changes required to maximize certain fluxes in biotechnological applications. The method relies on solving nonlinear programming models via global optimization techniques.

Keywords: Optimization, Power-law, Evolution.

1. Introduction

In natural cells, emergence of new designs results from evolution. The adaptive response of the cellular metabolism to different situations is attained by tuning gene expression and enzyme activity. Understanding the evolution of adaptive strategies is an important goal in Systems Biology.

The evolution of adaptive stress responses can be seen as a multi objective optimization problem. In that sense, the observed response represents an optimal (in some sense) combination of changes that ensure appropriate survival in the considered conditions. Evolution results in adaptations that are admissible solutions fulfilling important physiological constraints.

Within this general context, we introduce a novel approach that aims to identify enzyme activity regions containing feasible responses observed in evolution. The method introduced can also be employed to optimize biological systems in biotechnological applications. Our approach focuses on the properties of a particular class of non-linear models, the GMA (Generalized Mass Action) models that are based on the power-law formalism. The proposed algorithm is very efficient for realistic problems. The solutions found would represent the landscape in which evolutive solutions are expected. Comparison of our results and actual data shows the practical usefulness of the proposed method.

2. GMA representation

We shall consider a metabolic network that has p fluxes that can contribute to the change in the concentration of the pool of any of the n internal metabolites:

$$\frac{dX_i}{dt} = \sum_{r=1}^{p} \mu_{ir} v_r \qquad i = 1,...n$$

Here, μ_{ir} is a stoichiometric factor that indicates how many molecules of X_i are produced or used by the process v_r; it is a positive integer if the flux r produces X_i and it is a negative integer if the flux r depletes the pool of X_i. Each velocity can be represented by different functional forms, but, the so-called power-law formalism is one of the most convenient:

$$v_r = \gamma_r \prod_{j=1}^{n+m} X_j^{f_{rj}}$$

In this representation, X_j accounts for the concentration of metabolite j, γ_r is an apparent rate constant for flux r, and f_{rj} is the kinetic order of variable X_j in reaction r. Each kinetic order quantifies the effect of the metabolite X_j on flux r and corresponds to the local sensitivity of the rate v_r to X_j evaluated at the corresponding operating point. Using this representation, a Generalized Mass Action (GMA) model is defined as [1]:

$$\frac{dX_i}{dt} = \sum_{r=1}^{p} \mu_{ir} \left(\gamma_r \prod_{j=1}^{n+m} X_j^{f_{rj}} \right) \qquad i = 1,...n$$

In this expression, m indicates independent (external) metabolites.

3. Optimization model and solution strategy

Here, we present an optimization framework for GMA systems that will be later on taken as a basis for deriving the feasibility approach, which is the main contribution of this work. Non-linear optimization models based on the power-law formalism were first proposed by Voit [2]. In S-system representations, a transformation to logarithmic coordinates can be applied thus leading to linear optimization models. However, when the problem is represented by a GMA model, this technique cannot be applied.

In general, the problem of identifying the optimal values of v_r, γ_r and X_j that maximize a given criterion and satisfy at the same time the equations involved in the GMA representation can be posed as a nonlinear programming (NLP) as follows:

$$ONLP = \min \; U(v_r, \gamma_r, X_j)$$

$$s.t. \quad \sum_{r=1}^{p} \mu_{ir} v_r = 0 \qquad i = 1,...n$$

$$v_r = \gamma_r \prod_{j=1}^{n+m} X_j^{f_{rj}} \qquad r = 1,..., p$$

$$v_r, \gamma_r, X_j \in \Re_+$$

Model **ONLP** corresponds to a non-convex problem. Because of this, standard NLP techniques may get trapped in local solutions that are likely to be far away from the global optimum. This may lead to wrong conclusions when performing biological studies. To circumvent this limitation, we introduce a deterministic algorithm to globally optimize **ONLP** that is based on the works of Bergamini and co-workers [3] and Polisetty et al. [4]. The proposed method relies on hierarchically decomposing the problem into two levels, an upper level master problem **CMILP** and a lower level slave problem **RNLP**, between which the algorithm iterates until a termination criterion is satisfied (see Figure 1).

The master level of the algorithm entails the solution of a mixed-integer linear (MILP) problem, which is a relaxation of model **ONLP** (i.e., it rigorously overestimates the feasible region of **ONLP**), and therefore predicts a valid lower bound on its global optimum. In the lower level, the original problem is locally optimized in a reduced search space (i.e., model **RNLP**), thus yielding an upper bound on its global solution. The upper and lower level problems are solved iteratively until the bounds converge. Due to space limitations, technical details of the main features of the proposed algorithm are omitted. As mentioned before, this method can be employed in biotechnological applications in order to optimize a given bioprocess. In this work, as discussed in section 4, such method is employed to derive a tool to perform feasibility analysis in evolutive studies.

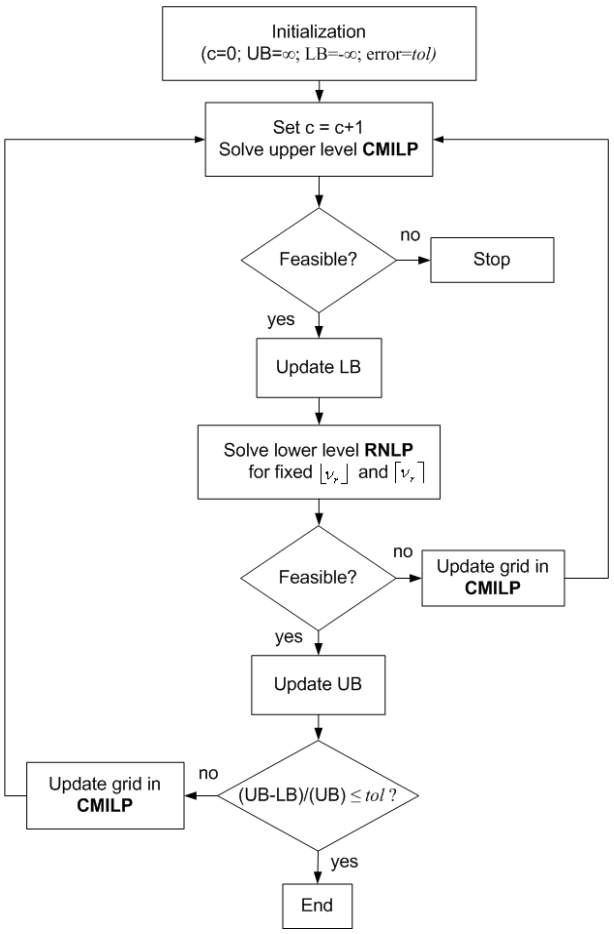


Figure 1. Proposed algorithm.

4. Feasibility approach

The algorithm previously presented can be used, after minor modifications, to identify regions that contain feasible solutions to the original problem **ONLP**, and discard others

in which no single feasible solution exists. Given the metabolic model, the goal is then to find the admissible changes at the level of enzyme activities that are compatible with a set of physiological and functional effective criteria.

From the mathematical point of view, this analysis requires the definition of a set of disjoint sets P_S^q ($P_S^q \cap P_S^{q'} = 0$ for all $q \neq q'$) such that their union contains the feasible space S of **ONLP** ($S \subseteq \bigcup_{q=1,...,Q} P_S^q$). In this work, for the sake of simplicity, we assume that each of these regions P_S^q is a hyper-rectangle described by a set of linear inequalities that impose lower and upper limits ($\lfloor \gamma^q \rfloor$ and $\lceil \gamma^q \rceil$, respectively) on the values of the apparent rate constants γ^q. Thus, we have:

$$P_S^q = \left\{ (v, \gamma, X) \in \Re_+^p \times \Re_+^p \times \Re_+^{n+m} : \lfloor \gamma^q \rfloor \leq \gamma^q \leq \lceil \gamma^q \rceil \right\} \qquad q = 1,...,Q$$

Hence, the feasibility analysis must determine whether these hyper-rectangles contain feasible solutions to **ONLP** or not.

The method devised to accomplish this task is based on the same ideas presented before and comprises two different levels. At the upper level, a master problem is solved to identify a region (i.e., hyper-rectangle) that may contain a feasible solution of **ONLP**. At the lower level, the prediction made by the master problem is checked by solving the original problem in a reduced search space. If a feasible solution is found, then integer cuts are added to the master problem in order to exclude the region containing such a feasible point. Otherwise, the master model is updated by refining its grid, until either a feasible solution is obtained in the lower level or the higher level problem turns out to be unfeasible.

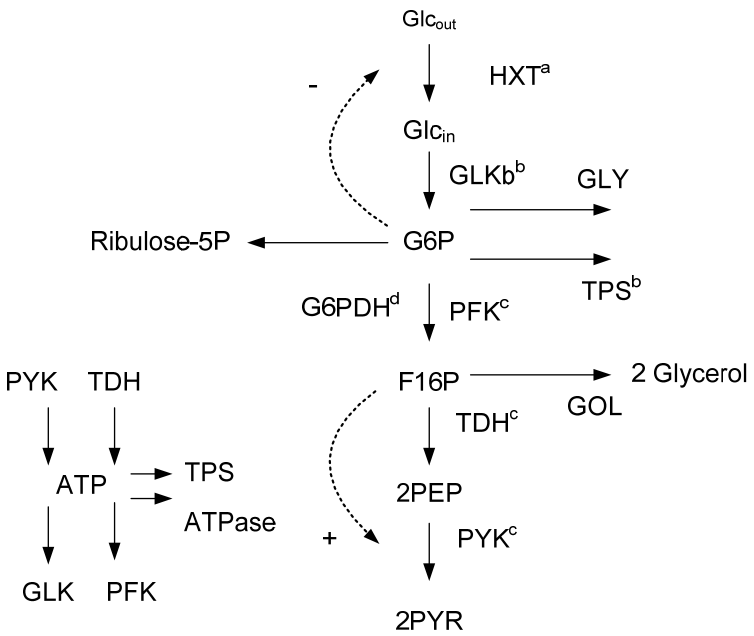


Figure 2. Scheme of the modeled pathways and ranges used for generation of the in silico gen expression profiles (GEPs).

5. Feasible adaptive response of yeast to heat shock

The capabilities of our method were illustrated through its application to the optimal adaptive response of yeast to heat shock (for a detailed description see [5]). Our model includes the core of the glycolytic pathway and the first step of the pentose phosphate pathway. It also accounts for the synthesis of glycogen, trehalose and glycerol, as shown in Figure 2. The notation used in this figure is as follows. Glcout: Extracellular Glucose; Glcin: Intracellular Glucose; G6P: Glucose-6-phosphate; F16P: Fructose-1,6-biphosphate; PEP: Phosphoenolpyruvate; PYR: Pyruvate; HXT: Hexose transporters (HXT1–4, HXT6–8, HXT12); GLK: Glucokinase/Hexokinase (GLK1, HXK1, HXK2); PFK: Phosphofructokinase (PFK1, PFK2); TDH: Glyceraldehyde-3-phosphate dehydrogenase (TDH1, TDH2, TDH3); PYK: Pyruvate kynase (PYK1, PYK2); GLY: Production glycogen; TPS: Trehalose 6-phosphate syntase complex (TPS1, TPS2, TPS3); G6PDH: Glucose 6-phosphate dehydrogenase (ZWF1).
The metabolic network was found to be specially sensitive to changes in two specific enzymes (i.e., PFK and TDH). For this reason, the feasibility analysis was performed on their domain, defining ten different sub-intervals for each of them. Hence, in this particular example, the feasibility analysis focuses on identifying, from the initial set of 100 hyper-rectangles, those containing feasible solutions to **ONLP** and those in which no feasible point exists.

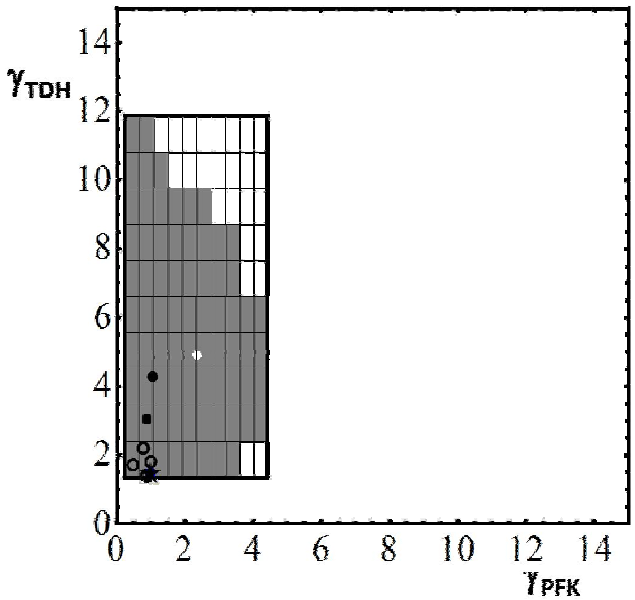


Figure 3. Feasibility analysis. White circle: Maximum rate of ATP synthesis; Black circle: Maximum rate of NADPH synthesis; Black square: Maximum rate of Trehalose synthesis; Black star: Minimum cost; Black empty circles: Experimental observations by Vilaprinyo et al. [5]

The algorithm was implemented in GAMS interfacing with CPLEX and CONOPT as main optimization packages. The total CPU time was less than one minute on an Intel 1.2 GHz machine. Results of this analysis are depicted in Figure 3.
In the figure, shady boxes represent hyper-rectangles that contain at least one feasible solution to the problem, whereas those in white have been proved to be unfeasible. For

comparison purposes, we have also depicted other solutions that are optimal in terms of some criteria: maximum rate of ATP synthesis, maximum rate of NADPH synthesis, maximum rate of Trehalose synthesis and minimum cost. This last metric (i.e., the cost) measures the overexpression of the enzymes.
As can be seen, experimentally observed responses [5] fall within the feasible region predicted by the algorithm. Interestingly, they allocate especially close to the minimum cost solution, that is, the one that would minimize the overexpression of the enzymes. Additionally, the maximum Trehalose rate solution is also near. This probably indicates some importance in the adaptation process.

6. Conclusions

This work introduced a systematic method for identifying the enzyme activity changes that allow a system to meet a set of physiological constraints while optimizing a parameter in the network. The approach presented relies on formulating nonconvex nonlinear problems that are solved via global optimization techniques.
The approach presented was applied to study the optimal adaptive response of yeast to heat shock. Experimental data fall well within the feasible region predicted. The closeness of those points to the minimum cost solution suggests that a conservative strategy where minimum changes are done is the preferred adaptive response. On the computational side, our method proved to be very efficient for medium size problems. The solutions found are intended to shed light on both, biotechnological and evolution studies.

7. Acknowledgments

Financial support received from the Spanish “Ministerio de Educación y Ciencia” (projects DPI2008-04099, PHB2008-0090-PC and BFU2008-00196), the Spanish “Ministerio de Asuntos Exteriores” (projects A/8502/07, HS2007-0006 and A/020104/08) and from “Generalitat de Catalunya” (FI programs) is fully appreciated.

References

Voit EO: Computational Analysis of Biochemical Systems. A Practical Guide for Biochemists and Molecular Biologists. Cambridge, U.K.: Cambridge University Press 2000.

Voit EO: Optimization in integrated biochemical systems. Biotechnol Bioeng 1992, 40(5):572-582.

Bergamini ML, Aguirre P, Grossmann IE: Logic-based outer approximation for globally optimal synthesis of process networks. Computers and Chemical Engineering 2005, 29:1914-1933.

Polisetty PK, Gatzke EP, Voit EO: Yield optimization of regulated metabolic systems using deterministic branch-and-reduce methods. Biotech Bioeng 2008, 99(5):1154-1169.

Vilaprinyo E, Alves R, Sorribas A: Use of physiological constraints to identify quantitative design principles for gene expression in yeast adaptation to heat shock. BMC Bioinformatics 2006, 7:184.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Controlled Formation of Self-assembled Nanostructures with Desired Geometries: Robust Dynamic Paths to Robust Desired Structures

Earl O. P. Solis, Paul I. Barton, George Stephanopoulos

Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA

Abstract

This paper discusses the design principles underlying the controlled formation of nanostructures with desired geometries through a hybrid top-down and bottom-up approach: top-down formation of the physical domains with externally-imposed controls and bottom-up self-assembly of the nanoscale particles to form the desired structure. We propose a two-phase approach for this design problem. The first phase guarantees a robust desired structure, and the second allows the desired structure to be reachable from any initial particle distribution in the physical domain. Both phases require the solution of combinatorially-constrained quadratic optimization problems. The dynamics of the self-assembly process is described through a multiresolution view of the system. Crucial to the achievement of the design goals is the need to break the ergodicity of the system.

Keywords: Robust nanostructures, self-assembly dynamics, ergodicity breaking.

1. Introduction

The fabrication of structures with geometric features at the nanometer length scale is essential to the manufacturing of future electronic, magnetic and optical devices composed of nanoparticles. One major challenge for the formation of complex nanoscale structures is the precise positioning of particles into the desired geometries. The work described below focuses on the design principles and methodologies underlying the controlled formation of such structures through a hybrid top-down and bottom-up approach: top-down formation of physical domains with externally-imposed controls (degrees of freedom), and bottom-up generation of the desired structure through the self-assembly of the nanoscale particles, driven by interparticle interactions and interactions with the external controls. This type of approach is seen experimentally in templated self-assembly techniques (reviewed in Koh, 2007), e.g., crystallization on template surfaces (Aizenberg et al., 1999), crystallization of colloids in optical fields (Burns et al., 1990), and DNA-programmed placement using DNA crystals as scaffolds (Le et al., 2004).

Many of the current self-assembly techniques are only applicable to periodic, close-packed structures, which simply require judicious nanoparticle design. The hybrid design strategy we propose in this paper can be used to systematically engineer more complex non-periodic, non-close-packed structures, which require external controls to guide the self-assembly process. To implement the design strategy we need to solve the following key problems: (1) *Static Problem*: the placement of externally-imposed controls and determination of their intensities in order to ensure a stable self-assembled

desired structure; (2) *Dynamic Problem*: the use of time-varying external controls to ensure that the desired final structure can be reached with the highest probability from any initial particle distribution. This paper is organized as follows. Section 2 introduces the two case studies discussed throughout this paper. Section 3 and 4 will focus on the static and dynamic problems, respectively.

2. Self-assembly Model Systems

The model system we use to demonstrate our system design strategies is an isomorph of the Ising model. Figure 1a shows 1- and 2-D desired structures in finite lattice domains. The 1D example has a volume, V=8, and particle number, N=5; the 2D example has a volume, V=64, and particle number, N=19. The former system only has 56 total configurations in its phase space and is used to compare dynamic Monte Carlo (MC) simulation results to Boltzmann probability distribution calculations in the Canonical prescription. All dynamic MC simulations were performed using the "virtual-move" MC algorithm (Whitelam and Geissler, 2007).

Ergodicity-breaking is a key component of our design strategy. An ergodic system is one that exhibits eventual access to all other phase space states from any particular system state. For the controlled self-assembly processes described in this paper, the externally-imposed controls offer degrees of freedom, which are used to decrease the volume of phase space accessible to the system, i.e., decrease the number of states accessible from a given state. Such systems exhibit *non-ergodicity*, and their phase space can be decomposed into ergodic subsets called *components.* Transitioning between any two components separated by large energetic barriers is not very probable, as it would require either an unreasonably high system temperature or long period of time.

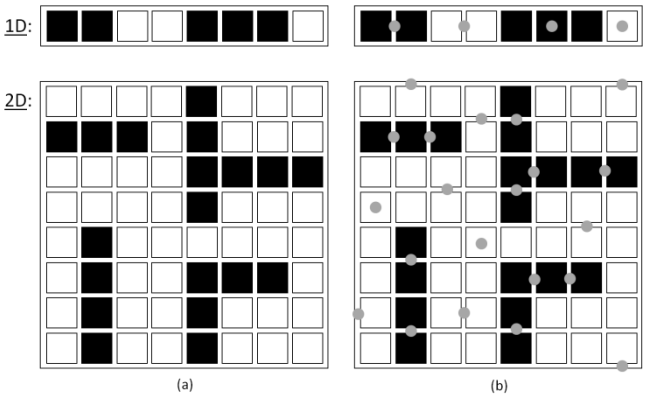


Figure 1: (a) Example desired configurations in finite 1- and 2-D lattice volumes. (b) The location of point conditions as external controls in the 1- and 2-D example systems.

For both example systems, we assume that the particles do not rotate but simply translate throughout the discrete system volume. We also assume that the particles are indistinguishable and negatively charged with a charge equal to -1. The binary interaction potential energy between particles is given by

$$\beta E = z_i z_j J_{ij} = z_i z_j \left[\frac{1}{r_{ij}} - \frac{1}{r_{ij}^{\varepsilon}}\right], \tag{1}$$

where r_{ij} is the positional distance between lattice sites i and j. Variables z_i and z_j are binary with value 0 for an empty lattice site or 1 for an occupied site. This is a simple

model that is intended to simulate a long-range Coulombic repulsion and short-range attractive potential. For real systems, more complex interaction energy models will be required, but the above model serves as a phenomenological one.

The static design strategy involves two steps that create an ergodic component with the desired configuration as its only member: (1) define the locations of the system degrees of freedom needed to guarantee a robust desired configuration, and (2) optimize the intensities of each degree of freedom. The dynamic strategy is similar in that it restricts the system phase space, but it does so progressively over time until it reaches the component defined by the static problem.

3. Static Problem Strategy

Assuming that the system has achieved the desired structure, we must guarantee that this structure is maintained, i.e., we have to ensure that the desired structure is robust. To do this, we specify the necessary system degrees of freedom: the location of point conditions (attractive or repulsive point charges) and their intensities (charge values). The degrees of freedom we use are isotropic and well- or barrier-forming, depending on the charge. Because the presence of two wells forms a barrier and vice versa, only one type of point condition is needed to define the general features of the energy landscape.

In defining the point condition locations, we essentially separate the system into subvolumes of occupied lattice sites that require attractive point conditions and unoccupied sites that require repulsive point conditions. To find the locations of the minimum number of well- and barrier- forming point conditions needed, we developed an algorithm which covers the system volume with *tiles* of specific shapes. Specifically, for each point condition we assign a tile which encompasses the local area that the point condition is intended to influence. Thus, the minimum number of tiles needed to cover the barrier-forming regions determines the minimum number of negative point charges, and the minimum number of tiles needed to cover the well-forming regions determines the minimum number of positive point charges. We select the type of point condition with the smallest number necessary to cover the appropriate regions. The details of the set-covering method can be found in Solis et al. (2009).

For the 1- and 2-D example systems, Figure 1b shows the solution to the set-covering method. The 1D example has 2 well-forming and 2 barrier-forming point conditions. We choose to use the well-forming type. The 2D example has 11 well-forming point conditions and 10 barrier-forming point conditions. We choose the 10 barrier-forming point conditions since it is smaller in number.

Given the minimum number of point conditions and their locations, we want to generate an energy landscape in the physical domain that ensures a robust desired structure. The solution to the following problem determines the robustness of the desired structure against statistical fluctuations:

$$\begin{aligned} &\max_{\boldsymbol{s} \in \boldsymbol{S}, \delta} \quad \delta \\ &s.t.\ E(\boldsymbol{s}, \boldsymbol{z}) - E(\boldsymbol{s}, \boldsymbol{z}_d) \geq \delta, \forall \boldsymbol{z} \in \boldsymbol{\zeta}^{\alpha} \backslash \boldsymbol{z}_d, \end{aligned} \tag{2}$$

where $\boldsymbol{z}_d$ represents the desired structure, $\boldsymbol{s}$ represents the point condition strengths, and $\boldsymbol{\zeta}^{\alpha}$ represents the subset of competing configurations that make up component α. The solution to this optimization problem ensures that the difference in potential energy, δ, between any configuration in the ergodic component, to which the desired structure belongs, and the desired structure is as large as possible. The δ value is therefore a

measure of the desired structure's robustness. If δ is sufficiently large, then the static problem is solved. If not, additional point conditions are needed. Since the minimum number of point conditions is of the same type, i.e., well- or barrier-forming, the additional point conditions will be selected from the opposite type.

For the 1- and 2-D examples, the solution to the optimization problem (2) gives the output $\delta = 19.8$ and $\delta = 0.1$, respectively. If, for instance, we would like the value of δ for the 2D example to be larger to satisfy a system robustness criteria, we can find the competing configuration(s) whose energy differs the least from the desired structure, i.e., the energy difference is equal to δ. The configurational differences between the desired and competing configuration(s) inform us of the *constraining feature(s)* of the desired structure. We can select the point condition(s) of the opposite type that energetically influences the constraining feature(s) and solve the optimization problem (2) again with the added point condition(s). In the 2D example, the constraining feature is the particle located in the lattice row 6 and column 7. Hence, we activate the point condition near it (row 6 and column 6.5) and solve the optimization problem again to find an increased value for the minimum energy difference, $\delta = 0.6$.

4. Dynamic Problem Strategy

Simply guaranteeing that the desired configuration is the minimum energy configuration within its component does not mean that the system, starting from any initial state, has a high probability of reaching the desired state at a particular temperature or within a reasonable timeframe. In fact, the formation of large barriers between the desired configuration and neighboring states in the solution to the static problem above creates a rough potential energy landscape, and given that the system has an unknown initial configuration, the problem of becoming kinetically trapped in a metastable configuration needs to be addressed.

We propose the use of time-varying degrees of freedom to restrict the volume of accessible system phase space states. To do this, we have adopted a multiresolution view of the desired structure. Figure 2 shows a systematic course-graining of the system particle number for the 1D example system. The proper use of time-varying degrees of freedom helps force the system to go through progressively smaller subsets of phase space, guaranteeing a higher likelihood of achieving the desired state when compared to simply using the static problem solution alone.

Similar to the static problem formulation above, the following optimization problem is solved for each stage in the dynamic process:

$$\max_{\boldsymbol{s} \in \boldsymbol{S},\, \delta^{(i)}} \delta^{(i)} \quad s.t.\ E\left(\boldsymbol{s}, \boldsymbol{z}^{(i)}\right) - E\left(\boldsymbol{s}, \boldsymbol{z}_{\boldsymbol{d}}^{(i)}\right) \geq \delta, \forall \boldsymbol{z}^{(i)} \in \boldsymbol{\zeta} \backslash \boldsymbol{z}_{\boldsymbol{d}}^{(i)}, \tag{3}$$

where $\boldsymbol{z}_{\boldsymbol{d}}^{(i)}$ represents all configurations that belong to the desired coarse-grained density specification of stage i and $\boldsymbol{\zeta}$ represents all configurations in phase space. The main principle behind this optimization formulation is to maximize the energy gap, $\delta^{(i)}$, between the configuration in $\boldsymbol{z}_{\boldsymbol{d}}^{(i)}$ with the highest energy and the configuration in the remainder of phase space (excluding the configurations that satisfy the desired density specification) with the lowest energy. In turn, this will also try to push the overall minimum energy state to be within $\boldsymbol{z}_{\boldsymbol{d}}^{(i)}$.

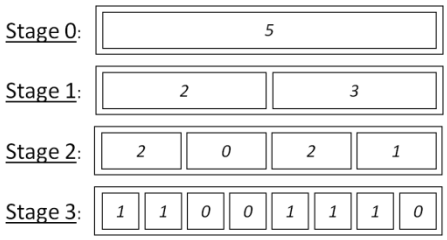


Figure 2: The systematic coarse-graining of the desired configuration, where Stage 3 is the desired detailed structure. The coarse-grained particle number specifications defined in the lower stages contain the desired structure but also contain progressively larger subsets of the phase space. Stage 0 contains all 56 states.

Given that the point condition locations in solving the static problem were sufficient to guarantee a robust desired configuration, we utilize the same set of point conditions for solving the dynamic problem. The solutions to this problem for Stages 1-3 of the 1D example system are shown in Figure 3.

The dynamic process for the 1D example system proceeds as follows: The system starts in Stage 0 with no point conditions and $k_BT = 1$; dynamic MC simulation results show that the system is ergodic at this temperature. Upon reaching equilibrium, we then impose the Stage 1 point conditions on the system at the temperature that maximizes the probability of being in the Stage 1 desired state, found through dynamic MC simulations to be $k_BT = 0.1$. Upon reaching equilibrium with the Stage 1 point condition intensities, we then impose the Stage 2 point condition intensities at a temperature ($k_BT = 0.25$) that maximizes the probability of the Stage 2 desired state starting from any configuration that satisfies the previous level's density specification. The same step is performed for Stage 3, where dynamic MC simulation results allow us to maintain $k_BT = 0.25$. After the system equilibrates, the solution to the Static Problem is used to maintain the desired configuration.

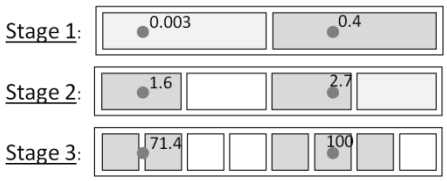


Figure 3: The dynamic point condition intensity solution to the optimization problem defined in Equation 3 for the 1D example system.

Figure 4 compares the results at the different stages of the simulation of the dynamic process to simulation results using the same point condition intensity values at each stage but starting from a random initial configuration. The time-varying process using the solutions to the optimization problems defined in Equation 3 shows a significant improvement in producing the desired configuration. However, there is room for further improvement, which comes from two further refinements: (1) A smoother transition between the point condition intensities and temperatures at each stage. Simulation results show that a simple linear transition is not favorable. This is because as the values of the external controls change in time, they induce higher energy barriers. Thus, the rate of change in the intensity values must decrease in order to allow equilibration in the current subset of phase space. (2) Though the system exhibits non-ergodic behavior at the temperatures used in the dynamic process, some of these temperatures are near the glass transition temperature, where there is a transition between ergodic and non-ergodic behavior. Utilizing a smoother evolution between the

stages of the dynamic process allows us to end each stage at lower temperatures where the system exhibits purely non-ergodic behavior, therefore increasing the probability of remaining in the desired subset of phase space. This gradual transition also allows systems "trapped" in metastable undesirable states to find the desirable state within a reasonable amount of time, thus increasing the probability of achieving the desired state.

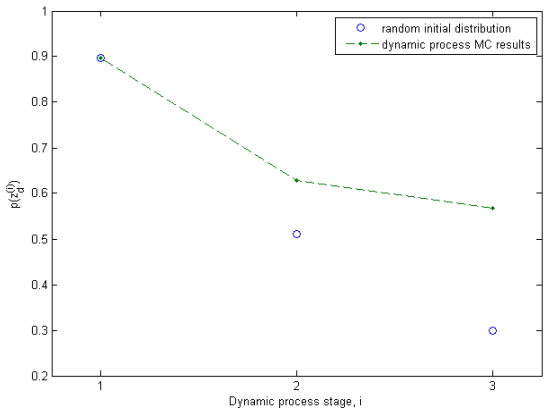


Figure 4: Dynamic MC results for the dynamic process compared to simulations with random initial states at the conditions of each stage in the dynamic process.

Conclusions

In this paper we proposed a two-phase approach for the controlled self-assembly of nanoparticles to form structures with desired geometries. The first phase provides us with the necessary system degrees of freedom to guarantee a robust desired structure. The second phase utilizes a dynamic process with time-varying degrees of freedom to increase the probability of reaching the desired structure from any initial particle distribution. Both phases require solving combinatorially-constrained quadratic optimization problems in order to find the external controls needed to dynamically reach the desired nanostructure and maintain it within a specific level of robustness.

Acknowledgements

Earl Solis and George Stephanopoulos acknowledge with gratitude the support for this work provided by Mitsubishi Chemical Holding Company.

References

J. Aizenberg, A.J. Black, G.M. Whitesides, 1999, Control of Crystal Nucleation by Patterned Self-Assembled Monolayers, Nature, 398, 495-498.

M.M. Burns, J.M. Fournier, J.A. Golovchenko, 1990, Optical Matter: Crystallization and Binding in Intense Optical Fields, Science, 249, 749-754.

S.J. Koh, 2007, Strategies for Controlled Placement of Nanoscale Building Blocks, Nanoscale Research Letters, 2, 519-545.

J.D. Le, Y. Pinto, N.C. Seeman, K. Musier-Forsyth, T.A. Taton, R.A. Kiehl, 2004, DNA-templated Self-assembly of Metallic Nanocomponent Arrays on a Surface, Nano Letters, 4, 2343-2347.

S. Whitelam, P.L. Geissler, 2007, Avoiding unphysical kinetic traps in Monte Carlo simulations of strongly attractive particles, Journal of Chemical Physics, 127, 154101.

E.O.P. Solis, P.I. Barton, G. Stephanopoulos, 2009, Controlled Formation of Nanostructures with Desired Geometries, FOCAPD '09 Proceedings.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Conversion of Glycerol to Liquid Fuels

Carlos A. Henao, Dante Simonetti, James A. Dumesic, and Christos T. Maravelias

Department of Chemical and Biological Engineering University of Wisconsin
1415 Engineering Drive, Madison - Wisconsin, USA.
E-mail address: christos@engr.wisc.edu

Abstract

A significant fraction of the total petroleum supply is used for transportation in the form of liquid hydrocarbons. This fact, along with the increasing demand for oil in developing countries has led to substantial research efforts in the area of renewable liquid fuels. One alternative is the conversion of vegetal oil and animal fat into bio-diesel. However, this comes with the production of significant amounts of glycerol, a byproduct that will become abundant if large scale bio-diesel production is implemented. In order to increase the total biomass to fuel efficiency during the production of bio-diesel and address the overproduction of glycerol, a novel integrated Glycerol Reforming (GR) + Fischer-Tropsch (FT) process is presented. The novelty of this process lies in the use of a Rhenium-based catalyst for the conversion of aqueous glycerol to synthesis gas (syngas). This step reduces significantly the cost of syngas production in traditional green FT processes. This work presents a preliminary process synthesis and an economic evaluation for a medium capacity GR-FT plant. The results show that the integrated process is economically attractive, and that there is room for further improvements trough the use of systematic process design and optimization methodologies.

Keywords: Liquid Fuels, Renewable Energy, Glycerol Reforming, Fischer-Tropsch.

1. Introduction

Petroleum feedstocks are currently used to provide liquid hydrocarbons for transportation fuels, and concerns regarding diminishing petroleum reserves have fostered the development of renewable transportation fuels to supplement or replace those derived from petroleum (Corma et al. 2007, Huber et al. 2007, Bozell 2006, Chheda et al. 2007, Ragauskas et al. 2006, Rostrup 2004) Biomass provides an abundant, renewable source of carbon-containing molecules, thereby representing a potential alternative source of liquid transportation fuels (Klass 1988). The conversion of biomass-derived feedstocks to fuels in a manner that is cost-competitive with the refining of petroleum requires the development of new approaches that simplify the processing steps, thereby reducing capital and operating costs associated with separation and purification (Ragauskas et al. 2006). In this respect, the production of bio-diesel is an attractive option in that the processing steps are well established and production costs are low, compared to the case for production of liquid fuels from lignocellulose resources. For example, triglycerides can be converted to long-chain esters for diesel

fuel by transesterification reactions with methanol. However, this process leads to the formation of a glycerol co-product stream, representing an opportunity to convert a low-value waste stream to a more valuable product.

In the present paper, we explore the option of converting the waste glycerol stream from bio-diesel production to long-chain, linear hydrocarbons for use as liquid fuel. The approach we follow is to utilize the coupling of glycerol reforming to syngas over a PtRe-based catalyst, with syngas conversion to liquid alkanes by Fischer-Tropsch synthesis. This integrated process can potentially improve the economics of "green" Fischer-Tropsch process by reducing costs associated with syngas production, for example, by eliminating the need for an O_2-blown auto-thermal reformer or biomass. Also, our process presents the opportunity for reducing the size of the Fischer-Tropsch synthesis reactor by producing an undiluted synthesis gas stream and for eliminating subsequent cleaning steps required for synthesis gas produced from biomass gasification (Hamelinck et al. 2004, Bartholomew et al. 2006, Soares et al. 2006, Spath et al. 2003). Thus, our integrated process potentially allows for smaller scale Fischer-Tropsch synthesis plants to produce liquid fuels from biomass, as would be required for utilization of the glycerol stream derived from bio-diesel production.

2. General considerations

The main objective of this work is the synthesis and economic evaluation of a process for the transformation of crude glycerol, generated as a byproduct in the bio-diesel industry, into liquid hydrocarbons that could be used as transportation fuel. The process synthesis study was conducted using a detailed simulation-evaluation model created in ASPEN PLUS ® - ASPEN ICARUS PROCESS EVALUATOR 2006. This approach allows the use of accurate thermodynamic property calculations, detailed process unit modeling as well as detailed sizing-costing procedures.

2.1. Raw material characterization and plant capacity

The quality of crude glycerol appears to be highly dependent on the facility in which it is produced. A typical crude glycerol mixture coming from the bio-diesel industry includes: 84.5wt% glycerol, 0.3t% methanol, 12.2wt% moisture and 3.0wt% NaCl. The design presented here processes 16,000 ton/y of such mixture.

2.2. Glycerol reforming reactor

The glycerol reformer is an isothermal fixed bed multi-tube reactor. Simonetti et al. (2007) studied the reforming action of a Pt-Re/C catalyst using aqueous solution of glycerol with concentrations between 50wt% and 80wt%, at a temperature of 548 K and pressures between 5 and 17 bar, obtaining good catalyst activities. The associated stoichiometry and kinetic expressions are presented in Table 1. This non-standard ASPEN kinetic model was included in the ASPEN PLUS simulation model as a FORTRAN subroutine.

2.3. Fischer-Tropsch reactor

The FT unit is also an isothermal fixed bed multi-tube reactor. For this study we used the kinetic model for a cobalt-based catalyst as reported by Hamelinck (2003). This kinetic model combines an expression for the total rate of consumption of CO and a product distribution function in the form of the Anderson-Schulz-Flory formula. As before, a FORTRAN subroutine was used to include this non-standard ASPEN PLUS kinetics in the model.

Table 1: Glycerol reforming kinetic model.

$$C_3H_8O_3 \rightarrow 3 \cdot CO + 4 \cdot H_2 \qquad (rxn1)$$

$$H_2O + CO \rightarrow CO_2 + H_2 \qquad (rxn2)$$

$$r_1 = k_1 \cdot \alpha \cdot P_{C_3H_8O_3} \cdot \theta_*^2 \quad , \quad r_2 = k_2 \cdot P_{H_2O} \cdot \theta_* - k_{-2} \cdot \theta_{OH} \cdot \sqrt{P_{H_2}}$$

$$\theta_* = \frac{-\left(K_{CO} \cdot P_{CO} + \sqrt{K_{H_2} \cdot P_{H_2}}\right) + \sqrt{\left(K_{CO} \cdot P_{CO} + \sqrt{K_{H_2} \cdot P_{H_2}}\right)^2 + 8 \cdot \alpha \cdot P_{C_3H_8O_3}}}{4 \cdot \alpha \cdot P_{C_3H_8O_3}}$$

$$\theta_{OH} = \frac{k_2 \cdot P_{H_2O} \cdot \theta_* + k_{-3} \cdot P_{CO_2} \cdot \sqrt{P_{H_2}} \cdot \theta_*^2}{k_3 \cdot K_{CO} \cdot P_{CO} \cdot \theta_* + k_{-2} \cdot \sqrt{P_{H_2}}} \qquad k_{-3} = k_2 \cdot k_3 \cdot K_{CO} / \left(k_{-2} \cdot K_{WGS}\right) \qquad K_{WGS} = 10^{\left(2073/T[K] - 2.029\right)}$$

$k_i = A_i\, e^{-Eai/(RT)}$	A_i	Ea_i
$k_1 = A_1.e^{-Ea1/(RT)}$	60·E10 [min^{-1}]	85.6406·E3 [kJ/kmol]
$k_2 = A_2.e^{-Ea2/(RT)}$	60·E7 [$atm^{-1}.min^{-1}$]	78.5441·E3 [kJ/kmol]
$k_{-2} = A_{-2}.e^{-Ea-2/(RT)}$	60·E13 [$atm^{-1/2}.min^{-1}$]	96.3409·E3 [kJ/kmol]
$k_3 = A_3.e^{-Ea3/(RT)}$	60·E13 [min^{-1}]	55.8972·E3 [kJ/kmol]
$\alpha = A_\alpha.e^{-Ea\alpha/(RT)}$	1·E-7 [atm^{-1}]	-83.7304·E3 [kJ/kmol]
$k_i = A_i\, e^{\Delta Hi/(RT)}$	A_i	ΔH_i
$K_{CO} = A_{CO}.e^{\Delta Hco/(RT)}$	1E-7 [atm^{-1}]	96.4060E3 [kJ/kmol]
$K_{H2} = A_{H2}.e^{\Delta Hh2/(RT)}$	1E-6 [atm^{-1}]	68.8776E3 [kJ/kmol]

3. Process diagram

The proposed process diagram is shown in **Figure 1**. In this process, crude glycerol is fed to a low pressure evaporation unit to remove inorganic salts that could affect downstream operations (e.g., catalytic reactors). The generated vapor stream is mixed with water to adjust its composition and it is finally condensed. This processed glycerol stream is then compressed and heated before being fed to the glycerol reformer where it is transformed into syngas. The syngas stream is later mixed with a recycle stream of light hydrocarbons and fed to the Fischer-Tropsch reactor to produce a heavier hydrocarbon mixture. Operating conditions were chosen to allow a coupled operation of the FT and GR reactors, eliminating intermediate compression units. The resulting hydrocarbon stream is finally separated into three different cuts: light components, a C_4+ fraction that can be incorporated to transportation fuels, and finally, an aqueous waste stream.

The novelty of the presented process comes from a rhenium-based catalyst (Simonetti et al. 2007) that allows the glycerol reformer to be coupled to the Fischer-Tropsch reactor. This glycerol reformer produces a highly concentrated syngas stream, ready to be used by the FT unit. Operational and capital costs of the reformer are significantly lower than those of biomass gasifiers used in traditional green Fischer-Tropsch processes, leading to a competitive process alternative.

From the thermodynamic point of view, high glycerol conversion and a proper H_2/CO ratio in the syngas produced by the GR reaction are favored by low pressures and high temperatures. On the other hand, high C_{5+} selectivity in the FT reaction is favored by high pressures and low temperatures. These trends do not allow for full heat integration between the two reactors, because thermal coupling of the exothermic FT reactor with the endothermic GR reactor would require the temperature of in the GR reactor to be lower than the temperature in the FT reactor.

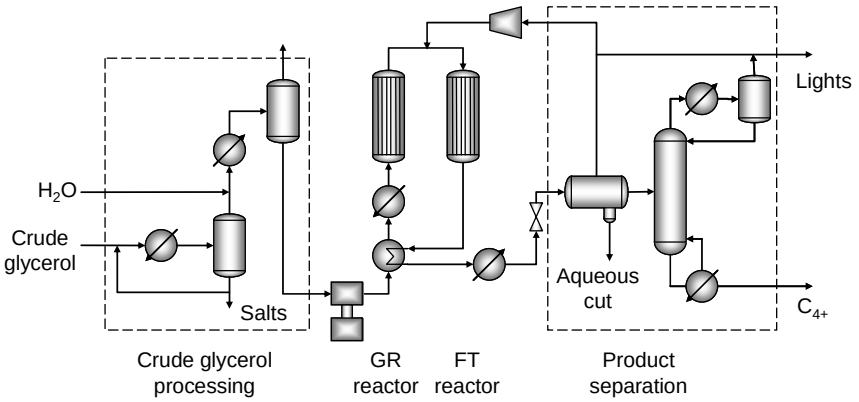


Figure 1: Preliminary glycerol-to-fuel process flow diagram.

However, we found that operating the two reactors at similar pressures while allowing the GR reactor temperature (T_{GRR}) to be higher than the FT reactor temperature (T_{FTR}) is the best option. In this way, an intermediate compression step between the reactors is not required, and the gains from a higher C_{5+} selectivity (resulted from allowing $T_{GRR}>T_{FTR}$) far outweigh the potential savings coming from a reactor heat integration.
For the preliminary process design presented in this study, both GR and FT reactors are operated at approximately 17 bar, while the operating temperatures are 573 and 548 K, respectively. In the FT reactor, these conditions lead to high C_{5+} selectivity. In the GR reactor, these conditions favor the water gas shift reaction, an effect that can be controlled by limiting the water content in the glycerol stream fed to the GR reactor, leading to the proper H_2/CO syngas ratio.

4. Economic evaluation

Detailed economic evaluation of the proposed alternative was conducted using ICARUS PROCESS EVALUATOR 2006 based on the aforementioned detailed process simulation model. Raw material costs are assumed to come solely from crude glycerol, while revenue is considered to come from the sale of light and intermediate liquid hydrocarbon fractions, assuming their values are linked, respectively, to those of natural gas and gasoline. The standard material prices considered here are as follows:

- *Crude glycerol price*: Some reports indicate the high availability of crude glycerol, coming from biodiesel production and other processes, has made its price plunge to levels near 0.02 USD/lb (Yoder et al., 2007).
- *Intermediate hydrocarbon mixture price*: The value of this mixture was considered equivalent to 80% of the current retail prices of gasoline (i.e. 1.80 USD/gal) (http://www.eia.doe.gov/oil_gas/petroleum/data_publications/wrgp/mogas_home_page. html). The current price of oil is 40 USD/bbl.
- *Light hydrocarbon mixture price*: The value of this mixture was assumed to be 30% of the current wellhead price of natural gas (http://tonto.eia.doe.gov/dnav/ng/ng_pri_sum_dcu_nus_m.htm).

The major parameters for this study and the main economic indicators are given in Table 2.

Table 2: Economic evaluation parameters and process evaluation results.

Economic Life [yr]	30	Cap. Cost [USD]	3,899,000
Working Cap./Cap. Expense	5%	Op. Cost [USD/y]	1,436,000
Op. Charged /Op. Labor	15%	Raw Material Cost[USD/y]	644,000
Overhead/Op. Labor	40%	Utility Cost [USD/y]	26,000
Desired RoR [%/yr]	8%	Product Sales [USD/y]	1,957,000
Tax Rate[%/yr]	40%		
Salvage Value/Cap. Cost	20%		
Depreciation	S.L.		
Cap. Escalation [%/yr]	5%		
Material Escalation [%/yr]	1.5%		
Product Escalation [%/yr]	5%		
Utility Escalation [%/yr]	3%		

The Net Present Value (NPV) of the project using the above nominal values is positive ($6.5 \cdot 10^6$ USD). The profitability of the process depends primarily on the price of crude glycerol, which is the sole raw material, and the price of gasoline, which directly determines the price of the major product of the process. To account for uncertainty in these parameters, we carried out sensitivity analysis studies. The goal in these studies was to determine an upper bound on the price of crude glycerol and a lower bound on the price of gasoline that would still allow the processes to be economically feasible (break-even point).

To this end, detailed NPV calculations where performed modifying, one at the time, the two aforementioned prices while keeping all other parameters at their nominal values presented above. The results of these analyses are presented in **Figure 2**. The break-even (maximum) price for glycerol is 4.5¢/lb, well above the current price. Most interestingly, the break-even (minimum) price of gasoline is 1.12$/gal, which is again lower than the current price, indicating that the integrated GR-FT process can be economically attractive even with relatively low oil prices. Thus, we conclude that the proposed process has the economic potential that justifies further development.

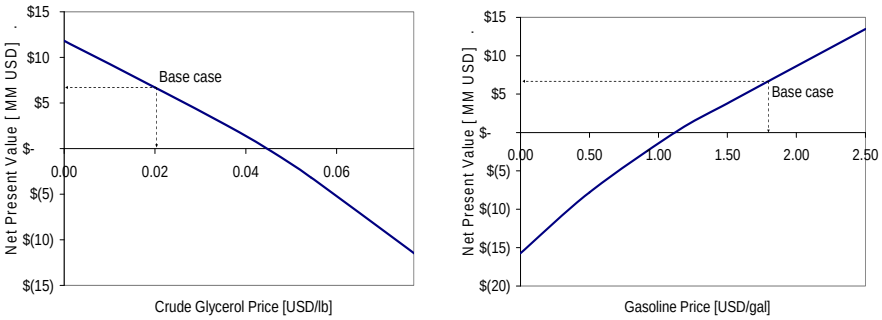


Figure 2: NPV sensitivity analysis for crude glycerol prices and gasoline prices. Nominal values (base case) lead to positive NPV.

5. Conclusions

A novel process for the conversion of crude glycerol to liquid hydrocarbons has been presented and its economic potential evaluated. The proposed design has a number of advantages. First, glycerol reforming (GR) can be carried out at high pressures (up to 17 atm), eliminating the syngas compression step otherwise required to reach FT reaction conditions. Second, glycerol can be reformed at temperatures below 600 K, thus allowing some heat integration with the exothermic FT synthesis. Third, compared to traditional green FT processes (e.g., processes combining biomass gasification and FT), the proposed process features a substantially lower capital cost because it does not require expensive mass gasification and syngas cleaning units. This implies that small- or medium-scale plants can be built closer to biomass production areas. Finally, it allows the use of glycerol coming from other sources such us the fermentation of carbohydrates. Sensitivity analysis on the prices of raw materials and products indicates that the process has the economic potential to justify further development. The proposed process diagram constitutes the result of a preliminary synthesis study. The process can be improved through the development of better catalysts and the use of systematic process synthesis and optimization tools.

References

V.H. Bartholomew, R. J. Farrauto, 2006, Fundamentals of Industrial Catalytic Processes, , Wiley, Hoboken, NJ.

J.J. Bozell, 2006, Feedstocks for the Future using Technology Development as a Guide to Product Identification, ACS Symp. 921, 1-12.

J.N. Chheda, G.W. Huber, J.A. Dumesic, 2007, Liquid-phase Catalytic Processing of Biomass-derived Oxygenated Hydrocarbons to Fuels and Chemicals, Angew. Chem. Int. Ed. 46, 7164-7183.

C. A. Corma, G.W. Huber, L. Sauvanaud, P. O'Connor 2007, Processing Biomass-derived Oxygenates in the Oil Refinery: Catalytic Cracking (FCC) Reaction Pathways and Role of Catalyst, J. Catal, 247, 307-327.

C. N. Hamelinck, A. P. Faaij, H. Uil, H. Boerrigter 2004, Production of FT transsportation fuels from biomass; technical options, porocess analysis and optimisation, and development potential, Energy, 29, 11, 1743-1771.

G.W. Huber, P. O'Connor, C. A. Corma, 2007, Processing biomass in conventional oil refineries: Production of high quality diesel by hydrotreating vegetable oils in heavy vacuum oil mixtures, Appl. Catal. A-Gen, 329, 120-129.

J. Klass, 1998, Biomass for Renewable Energy, Fuels, and Chemicals, Academic Press: San Diego, CA.

A.J. Ragauskas, C.K. Williams, B.H. Davison, et al. 2006, The path forward for biofuels and biomaterials, Science, 311, 484-489.

D. A. Simonetti, J. A. Dumesic 2007, Coupling of Glycerol Processing with Fishcher-Tropsch synthesis for production of liquid fuels, Green Chemistry, 9, 10, 1029-1144.

R.R. Soares, D.A. Simonetti and J.A. Dumesic, 2006, Glycerol as a Source for Fuels and Chemicals by Low-Temperature Catalytic Processing, Angew. Chem. Int. Ed., 45, 3982-3985.

P.L. Spath, D. C. Dayton, 2003, Preliminary Screening--Technical and Economic Assessment of Synthesis Gas to Fuels and Chemicals with Emphasis on the Potential for Biomass-Derived Syngas, United States Department of Energy, National Renewable Energy Laboratory, NREL/TP-510-34929, 2003, 1-142.

J. Yoder, P. Wandschneider (2007). Economics and Policy for Washington State Biofuel Markets: Interim Report. Pg 33.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Dynamic modeling and simulation of CO_2 chemical absorption process for coal-fired power plants

Adekola Lawal[a], Meihong Wang[a], Peter Stephenson[b], Hoi Yeung[a]

[a]*Process Systems Engineering Group, School of Engineering, Cranfield University, Bedfordshire, MK43 0AL, United Kingdom*
[b]*RWE npower, Windmill Park, Swindon, SN5 6PB, UK*

Abstract

Post combustion capture via chemical absorption is viewed as the most mature CO_2 capture technique. The effects of the addition of CO_2 chemical absorption process on power plant performance have been studied using various steady-state models. However, there are several gaps in the understanding of the impact of post combustion capture on the operability of the power plant. These questions could be addressed by studying the dynamic behavior of such plants. In this study, dynamic models of the CO_2 chemical absorption process were developed and validated. Dynamic analyses of the process reveal that absorber performance is sensitive to L/G ratio and that changes in reboiler duty significantly affect the regenerator performance.

Keywords: Post combustion, CO_2 capture, Chemical absorption, dynamic modeling.

1. Introduction

Power generation from fossil fuel-fired power plants is the largest single source of CO_2 emissions [1]. With growing concerns about the environmental impact of such plants, effective CO_2 emission abatement strategies such as Carbon Capture and Storage (CCS) are required for their continued use. One approach to CCS is post combustion capture which involves the separation of CO_2 from the flue gas stream after combustion occurs. Chemical absorption is well suited for separating CO_2 from streams with low concentration of CO_2.

Chemical absorption involves the reaction of CO_2 with a chemical solvent to form a weakly bonded intermediate compound which may be regenerated with the application of heat [2]. Figure 1 describes one of the most popular technologies proposed for post combustion capture. The facility consists of two main units – the absorber and regenerator. Several studies have shown that the energy requirement for solvent regeneration would have adverse effects on power plant efficiency [3]. However, there are several gaps in the understanding of the impact of post combustion capture on the operability of the power plant. For instance, would such power plants be able to effectively operate at varying loads or what modifications are required during start-up [4]? These questions can be addressed by studying the dynamic behaviour of such plants. To achieve this, accurate dynamic models of the power plant and the CO_2 capture facility are required.

2. Developments in modeling chemical absorption of CO_2

Post combustion capture with MEA is a reactive absorption process. Two main phenomena are involved: mass transfer of CO_2 from the bulk vapour to the liquid

solvent and the chemical reaction between CO_2 and the solvent. A number of studies have employed steady state models of the chemical (or reactive) absorption process at different levels of complexity. Kenig et al describes the different levels of complexity of these models [5]. The equilibrium-based approach assumes theoretical stages in which liquid and vapour phases attain equilibrium while the rate-based approach estimates actual mass transfer rates. These models may assume the reactions are at equilibrium or may consider the reaction kinetics. Reaction kinetics could be considered by including an enhancement factor to estimate actual absorption rates (with chemical reactions) from known physical absorption rates. Otherwise, reaction kinetics could be modelled directly [5].

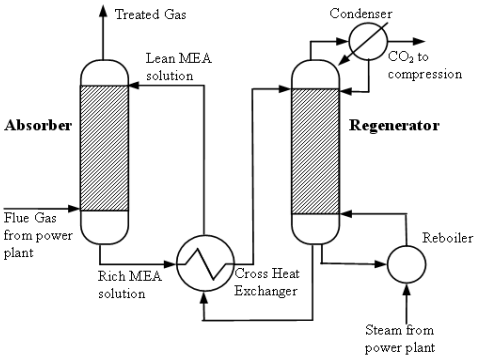


Figure 1 Chemical absorption process for post combustion capture from [4]

The dynamic behavior of the CO_2 absorption process for post combustion capture using MEA has not been extensively studied. Kvamsdal et al considered the dynamic simulation of only the absorber of the process using dynamic models of reduced complexity and also assumed a constant value for the heat of absorption of CO_2 and vaporization of water [4]. Lawal et al considered the dynamic simulation of only the absorber. Mass transfer was based on Maxwell-Stefan formulation and expressions were developed for the heat of absorption of CO_2 [6]. This paper extends the study to the simulation and analysis of both absorber and regenerator columns.

3. Model Development

This section describes the model development of the absorber and regenerator using the rate-based approach for mass transfer. The physical property method used for both approaches is the Electrolyte Non-random-two-liquid (NRTL) model. MEA electrolyte solution chemistry is used to predict the equilibrium mass fractions in the liquid and vapour phases [7]. Mass transfer is described using the two-film theory (Figure 2) using the Maxwell-Stefan formulation. Heat and mass transfer resistances are modelled in the liquid and vapour films. The rate-based model was developed from the Gas-Liquid Contactor model in Process Systems Enterprise's[1] Advanced Model Library using their process modelling tool, gPROMS. Modifications made include the diffusivity (χ) of CO_2 in the liquid phase which was based on expressions provided by Vaidya et al [8]. The diffusivity (χ) of CO_2 and other components in the vapour phase was estimated using the Fuller's equation [9]. Mass transfer coefficients in the liquid and vapour films were determined by correlations given by Onda et al [10]. Expressions for the heat of absorption and the heat lost to the surrounding were obtained from literature [11,12].

[1] Process Systems Enterprise (PSE) Ltd.

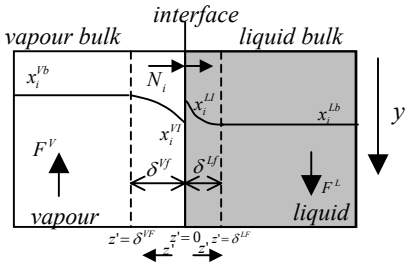


Figure 2 Liquid and vapour bulks, films and interface

3.1. Model Assumptions

The following assumptions were used in developing this dynamic model:

- Plug flow regime and linear pressure drop along the column
- Phase equilibrium at interface between liquid and vapour films
- Negligible solvent degradation

3.2. Model Equations

Mass Balance: $$\frac{dM_i}{dt} = \frac{-1}{L \cdot A}\frac{\partial F_i^L}{\partial y} + N_i \cdot Sp \cdot MW_i \cdot \omega \quad (1)$$

Energy Balance: $$\frac{dU}{dt} = \frac{-1}{L \cdot A}\frac{\partial F_H^L}{\partial y} + Sp \cdot \omega\left(H_{liq}^{cond} + H_{liq}^{conv} + H_{abs}\right) + HL \quad (2)$$

Heat of absorption (or desorption): $$H_{abs} = N_{CO_2} \times h_{abs} \quad (3)$$

Maxwell-Stefan formulation: $$\frac{1}{\delta}\frac{\partial x_i^M}{\partial z'} = \frac{1}{c_t}\sum_{k=1}^{n}\left(\frac{x_i^M N_k - x_k^M N_i}{\chi_{i,k}}\frac{\mu^R}{\mu}\frac{T}{298.15}\right) \quad (4)$$

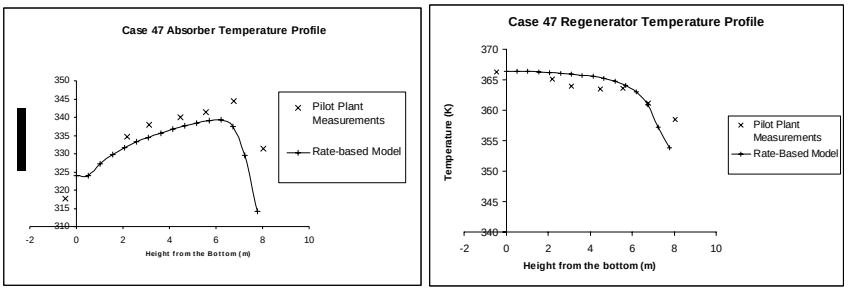
The physical property estimation models were set up in Aspen Properties through the CAPE-OPEN Thermo interface.

4. Model Validation

The models developed were validated using data from a pilot plant study [12]. Both absorber and regenerator columns of the pilot plant are packed columns with diameters of 0.427m and total packing heights of 6.1m [12]. Out of the 48 experimental cases carried out in the study, two cases (Cases 32 and 47) were selected for steady state validation purposes. These two cases were selected because of their relatively high and low liquid to gas (L/G) ratios respectively. Simulation results were validated using the temperature profile of both columns measured in the pilot plant [12]. Both columns were simulated separately for validation. In addition the measured CO_2 loading of the amine solvent taken at different positions was compared with values obtained from simulation.

4.1. Case 47

This case involved a relatively low liquid to gas (L/G) ratio in the absorber thus a lower CO_2 capture level. Because of the reported inaccuracy in the flue gas flow measurement [4,12], its value was adjusted to match reported capture levels (Table 1). The temperature profiles in the absorber and regenerator were used to validate the two models as shown in Figure 3(a) and (b). The rate-based model gives fairly good predictions of temperature profiles.

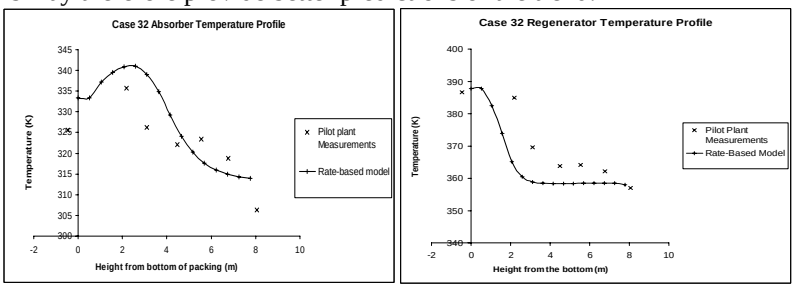


(a) (b)

Figure 3 (a) Absorber and (b) Regenerator temperature profile for Case 47

4.2. Case 32

This case involved a relatively high liquid to gas (L/G) ratio thus a high CO_2 capture level. The inlet flue gas flow rate to the absorber was reduced by about 15% to give better predictions of the temperature profile (Figure 4a). However, this change implies higher CO_2 capture levels than what was measured in the pilot plant (Table 1). This discrepancy may be due to the assumption that the reactions between CO_2 and MEA are at equilibrium as calculated by the electrolyte solution chemistry. Kinetically controlled reactions may therefore provide better predictions of the trend.



(a) (b)

Figure 4(a) Absorber and (b) Regenerator temperature profile for Case 32

Table 1 Comparing CO_2 loading at various locations measured in pilot plant and predicted by rate-based model

	Case 32		Case 47	
	Pilot Plant	Rate Model	Pilot Plant	Rate Model
Absorber Capture Level (%)	95	99.5	69	69.2
Absorber rich MEA loading (mol CO_2/mol MEA)	0.428	0.456	0.539	0.487
Regenerator lean MEA loading (mol CO_2/mol MEA)	0.272	0.260	0.286	0.262

5. Dynamic Analysis

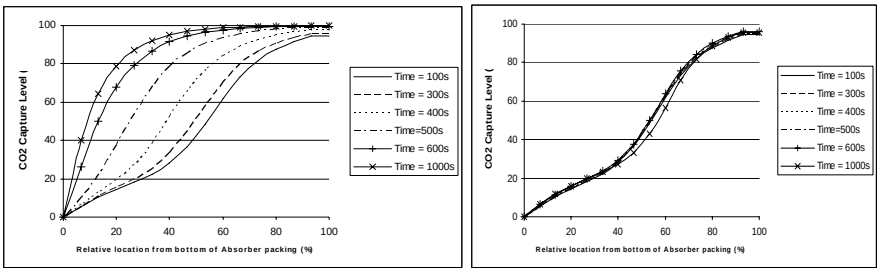
These analyses consider the effect of disturbances on the performance of the columns.

5.1. Reducing Power Plant Load

A 50% reduction in power plant load occurs. Two cases were considered:

- Case-A: Change of flue gas flow rate without changing liquid (solvent) flow rate
- Case-B: Change of flue gas flow rate with corresponding decrease in liquid solvent rate to maintain CO_2 capture level

In Case A, the process was simulated with the base-load conditions (Case 32) for three minutes after which the above changes were implemented in ten minutes. Finally conditions were maintained for eight minutes. From Figure 5a, the 100s curve represents the profile before dropping load. The other curves show a trend of increasing absorption levels with time. Since the flue gas flow rate is ramped down with time while the solvent flow rate is constant, an increase in L/G ratio occurs. In Case B, by reducing the lean solvent feed rate correspondingly (by 50%), roughly the same capture level (Figure 5b) could be maintained through the period of change. This suggests that the absorption process is more sensitive to the L/G (liquid solvent to flue gas) ratio than their actual flow rates. Since the amount of steam required for regeneration corresponds to the amount of lean MEA circulated, the energy requirement of the regenerator could be correspondingly reduced.



(a) Case A (b) Case B

Figure 5(a) and (b) Reducing Power Plant Load

5.2. Reducing Regenerator Reboiler Duty

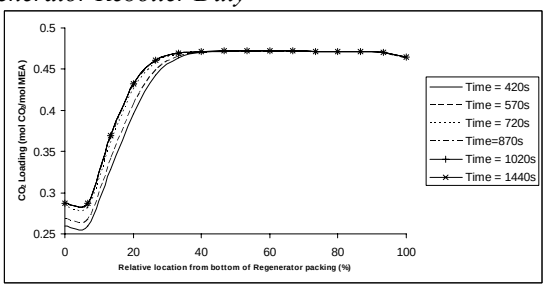


Figure 6 Changes in CO_2 loading profile in regenerator with reducing reboiler duty

The regenerator was simulated with the base-load conditions (Case 32) for seven minutes. A 10% reduction in reboiler duty was implemented over five minutes. Conditions were maintained for 12 minutes. Significant changes in the CO_2 loading profile are observed from the start of the disturbance (time = 420s) onwards especially towards the bottom of the column where CO_2 loading increases (Figure 6). The lean MEA solvent to the absorber would therefore have reduced absorption capacity.

6. Conclusions

This paper presents a study of CO_2 capture using chemical absorption based on the dynamic modelling of the process. Validation results show that the model predicts the absorber and regenerator temperature profiles and CO_2 loadings fairly well. Dynamic analyses show that the absorber performance is more sensitive to the L/G ratio than the actual flow rates of the solvent and flue gas. The performance of the regenerator is significantly affected by the reboiler duty.

Nomenclature

A	Cross sectional area (m^2)	μ	Viscosity (Pa.s)
c_t	Total molar concentration (mol/m^3)	ω	Wetted area ratio
F_i	Component mass flow rate (kg/s)	χ	Diffusivity (m^2/s)
F_H	Enthalpy flow rate (J/s)	*Subscripts*	
H	Heat flux (J/m^2)	*abs*	Absorption
h	Specific Enthalpy (J/kg)	*H*	Enthalpy
HL	Heat loss to surroundings (J/m^2)	*i*	Component number
L	Length of column section (m)	*Liq*	Liquid
L/G	Liquid to gas	*Vap*	Vapour
M	Mass Holdup (kg/m^3)	*Superscripts*	
MW	Molecular weight (kg/mol)	*Cond*	Conduction
N	Molar flux ($mol/m^2.s$)	*Conv*	Convection
n	Number of components	*I*	Interface
Sp	Specific area (m^2/m^3)	*L*	Liquid
U	Energy Holdup (J/m^3)	*Lb*	Liquid bulk
x	Mass fraction	*Lf*	Liquid film
x_i^M	Molar fraction	*R*	Reference
y	Axial position	*V*	Vapour
z'	Film position	*Vb*	Vapour bulk
Greek Symbols		*Vf*	Vapour film
δ	Film thickness (m)		

References

[1] P. Freund, 2003, Making deep reductions in CO_2 emissions from coal-fired power plant using capture and storage of CO_2, Proc. Inst. Mech. Eng. Part A J. Power Eng.,217(1):1-8.

[2] IPCC, 2005, IPCC special report on carbon dioxide capture and storage.

[3] J. Davison, 2007, Performance and costs of power plants with capture and storage of CO_2, Energy, 32, 7, 1163-1176.

[4] H.M. Kvamsdal, J.P. Jakobsen, K.A. Hoff, 2009, Dynamic modeling and simulation of a CO_2 absorber column for post-combustion CO2 capture. Chemical Engineering and Processing: Process Intensification, 48, 1, 135-144.

[5] E.Y. Kenig, R. Schneider, A. Górak, 2001, Reactive absorption: Optimal process design via optimal modelling, Chemical Engineering Science, 56, 2, 343-350.

[6] A. Lawal, M. Wang, P. Stephenson, H. Yeung, 2008, Dynamic modelling of CO_2 absorption for post combustion capture in coal-fired power plants, Fuel, doi:10.1016/j.fuel.2008.11.009.

[7] Aspentech, 2008, Rate-based model of the CO_2 capture process by MEA using Aspen Plus, Available at: http://support.aspentech.com/. Accessed May, 2008.

[8] P.D. Vaidya, V.V. Mahajani, 2005, Kinetics of the Reaction of CO_2 with Aqueous Formulated Solution Containing Monoethanolamine, N-Methyl-2-pyrrolidone and Diethylene Glycol, Ind. Eng. Chem. Res, 44, 6, 1868-1873.

[9] R.C. Reid, J.M. Prausnitz, T.K. Sherwood, 1977, The properties of gases and liquids, Third ed, New York, McGraw-Hill.

[10] K. Onda, H. Takeuchi, Y. Okumoto, 1968, Mass transfer coefficients between gas and liquid phases in packed columns, Journal of Chemical Engineering of Japan, 1, 1, 56-62.

[11] B.A. Oyenakan, 2007, Modeling of Strippers for CO_2 capture by Aqueous Amines, Doctor of Philosophy Thesis, University of Texas at Austin.

[12] E.R. Dugas, 2006, Pilot plant study of carbon dioxide capture by aqueous monoethanolamine, M.S.E. Thesis, University of Texas at Austin.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Transition to Clean Coal Technologies in India

Anish C. Patil

Delft University of Technology, Faculty of TPM, Delft 2628BX, Netherlands

Abstract

India has the third largest proven reserves of coal in the world and it is obvious that coal will play a major role in its future development. It is well documented that coal is not an environmentally friendly fuel, unless it is utilized in cooperation with clean coal technologies. This paper uses the Transition Management approach to devise a framework for the transition towards clean coal technologies in India. The concept of Transition Management is firmly rooted in the traditions of system thinking that highlights the co-evolution of the socio-technical systems. In order to make technological change sustainable, technical changes alone are not sufficient – changes in the social dimension – such as user practices, governance and institutions are necessary. The role of the government during the transition process is interactive policy formulation – lowering of legal, institutional and policy obstacles and facilitating knowledge transfer and collaboration between the actors.

Keywords: Sustainability, Clean Coal technologies, Transition Management.

1. Introduction

Energy plays a crucial role in the sustainable development of the society – clean, affordable, and uninterrupted supply of energy is the engine for future growth. It is a key ingredient in all sectors of modern economies as they are dependent on an uninterrupted supply of energy. Currently, fossil fuels cater to about 80% of the World's primary energy demand (IEA 2007). The constant dependence on fossil fuels to energize our lives presents a new dilemma – how to satisfy this constantly growing demand for energy given the fact that the fossil fuel resources are finite? The challenge is not just satisfying the increasing energy demand, but at the same time doing it in an environmentally friendly way. When coal, gas and oil are burnt, they release harmful greenhouse gases (GHG) that trap heat in the atmosphere and cause global warming. Because of the long residence times of GHGs in the atmosphere, measured in centuries, climate change would still constitute a risk, as a result of past emissions, even if human-induced emissions suddenly ceased (Kaygusuz 2009). Hence, at this point we cannot stop global warming but only can strive at minimizing its effect for the future generations.

2. Energy demand – India

A rapidly developing country such as India needs greater access to power if they are to achieve the high growth rates required to raise living standards and reduce poverty. On these lines to promote social and economic development, India has invested heavily into building new power plants. As India have large coal reserves, naturally Indian investment in the power sector come along the lines of coal power plants. Coal-based electricity production has driven much of the growth in the Indian electricity sector over the past few decades (COAL 2008). As the future demand for power increases in India,

so will the coal consumption. Although, coal is one of the most abundant energy sources available, it is also the most polluting energy source with very high carbon intensity (Marland and Boden 2001).

The current Indian policies are geared towards increasing the generation capacity that results into the installation of the least risky and cheapest technology. Advanced technologies are necessary for increasing efficiency, reducing environmental pollution and lowering carbon emissions from the Indian power sector (Chikkatur 2008; COAL 2008). India's coal based power generation is primarily based on the subcritical pulverized coal technology – mainly because of the need to rapidly install generation capacity in the country using domestic resources while building up indigenous manufacturing. Indian coal based power plants have low efficiency compared to their counterparts in the developed world. The poor efficiency of India's power plants is usually blamed on a variety of technical and institutional factors such as poor quality of coal, bad grid conditions, degradation due to age, lack of proper operation and maintenance at power plants, ineffective regulations, and lack of incentives for efficiency improvements (Chikkatur 2008). Given the right framework and incentives – Indian power sector too can employ and enjoy clean coal technologies.

The problem at hand is – how can the Indian power sector transition towards clean coal technologies?

Indian power sector is a large socio-technical system that includes, besides the technical subsystem (including the energy carriers, conversion and storage facilities) as well as the people, organizations, and authorities that plan, build, operate, use, and regulate the system (Kaijser 1994; Weijnen and Bosgra 1998). The challenge of transition within a socio-technical system entails joint optimization, at the least coherence, of the social and technical elements of the system (Neij and Åstrand 2006; Ajah, Patil et al. 2007).

Clean coal technologies have the potential to increase the power plant efficiency and are environment friendly, but are not employed on a large scale in India due to the high costs associated with them. Hence, there needs to be proper framework in place that determines which technology is best suited? Who pays for the technology? how long it will take to pay for the technology? etc. For a large country like India how do the policy makers come up with a framework for the transition towards clean coal? This paper suggests that a large developing country like India will need much more than just tweaking around the existing market structure. Just tweaking here and there will not give the long term transition framework that is required for sustainable development. This research proposes transition management approach to draw up a framework for transition towards clean coal for India. 'Transition management' approach can be characterized by the following items:

1. Long-term thinking (at least 25 years) as a framework for shaping the short-term policy thinking, thus trying to bring about system innovation alongside system improvement (Rotmans, Kemp et al. 2001)
2. Multi-level and multi-actor perspective (Geels 2002)
3. A focus on learning and a special learning philosophy (learning-by-doing and doing by-learning) achieved through Bounded Socio Technical Experimentation at the Niche level (Brown, Vergragt et al. 2003)

2.1. Problem background and research question

It is obvious that large developing countries such as India will require huge energy resources to fuel their development further, but at the same time it is paramount to balance this development without compromising the environment. Coal will play a

major part in India's development, but with the threat of climate change looming around it is wise that India invest in clean coal technology otherwise the negative effects on the environment will wipe out the positive economic development. Research question for this paper can be formulated as:
How can the Transition Management approach help to devise a framework for transition towards Clean Coal in India?

3. Transition Management Approach

Transition management aims at understanding socio-technical transitions and at finding methods that facilitate change directed towards achieving societal goals (Geels 2001). The concept "Niches" is of importance for this research, since it defines the level where transitions originate. In order to make changes in the socio-technical system sustainable, mere technological improvements are not sufficient (Rotmans, van Asselt et al. 2000) – changes in the social dimension – such as user practices, rules and regulations are inevitable (Hekkert, Suurs et al. 2007).
The `multi-level' perspective upon which transition management revolves around is shown in figure 1 below. Transition within a socio-technical system distinguishes between the macro-level of the Socio-technical landscape, the meso-level regime, and the micro-level niches.

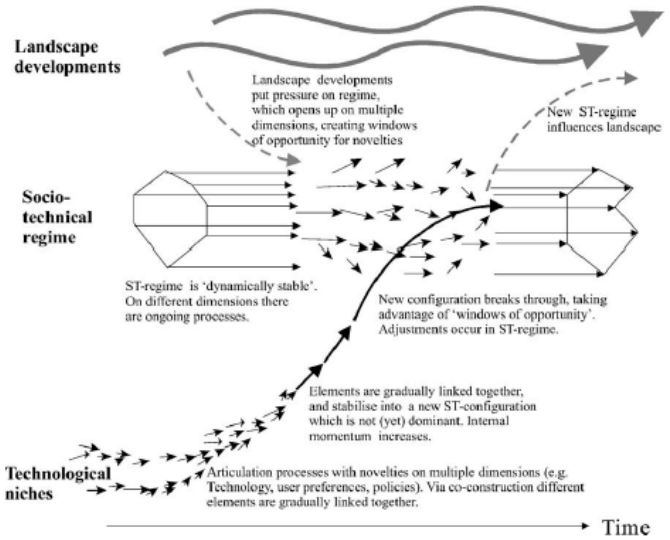


Figure 1: Multi-Level Perspective in Transition Management approach (source: (Geels 2002))

The multi-level approach is centered around the *socio-technical regime* – the socio-technical regime, e.g. energy system, accounts for stability of the socio-technical system through the coordinated and aligned activities of all the actors that are part of the system. The key feature of the 'multi-level perspective' is that system innovations occur through the interplay between the dynamics at these three levels: niche; regime; and landscape (Rotmans, Kemp et al. 2001; Geels 2002). Geels (Geels 2001; Geels 2002) shows that this multi-level model is a nested hierarchy, which can account for the stability in a regime. Geels argues that novelties emerging in technological niches may,

under certain conditions, break through to the regime and come to complement or substitute it – for example, the transition from the usage of horses to cars.

At the highest level of the multi-level model is the socio-technical landscape, which forms an exogenous macro level environment that influences developments in niches and regimes (Geels 2002). The socio-technical landscape tends to change only very slowly (for example, demographic changes, macro-economics, and cultural change). Socio-technical landscape developments refer mainly to national and international developments.

At the lowest level is the Niche level. This is the level that is most dynamic and where novelties emerge. These novelties are initially unstable socio-technical configurations with low performance. Hence, niches act as 'incubation rooms' protecting novelties against mainstream market selection (Kemp 1994; Geels 2002). Studies of past transitions (Kemp 1994; Geels 2002) suggest that a new technology will typically first commercialize in niche markets – where the advantages for this particular technology are the strongest (Utterback 1974; Christensen 1997). Niches help to create virtuous cycles that allow a new technologies to escape lock in, by helping the technology to overcome initial barriers of high costs; the non-availability of complementary technologies; institutional rigidities; and the nonalignment of a new technology to the external environment during the infancy period (Mulder, Reschke et al. 1999).

Niches act as 'incubation rooms' protecting novelties against mainstream market selection (Kemp 1994; Geels 2002). Niches are Bounded Socio-Technical Experiments (BSTE) that exists through large demonstration projects (Brown, Vergragt et al. 2003; Agnolucci and McDowall 2007) – BSTE implies experimentation with a new technology or service, on scale bounded by space and/or time. The time dimension is around five years, but may be longer or shorter, while the space dimension is defined either geographically (a community) or by a number of users (small) (Brown, Vergragt et al. 2003). Studies of past transitions (Kemp 1994; Geels 2002) suggest that a new technology will typically first commercialize in niche markets – where the advantages for this particular technology are the strongest (Utterback 1974; Christensen 1997). Niches help to create virtuous cycles that allow a new technologies to escape lock in, by helping the technology to overcome initial barriers of high costs; the non-availability of complementary technologies; institutional rigidities; and the nonalignment of a new technology to the external environment during the infancy period (Mulder, Reschke et al. 1999). Niches are important for the societal learning and experimentation of a new technology and corresponding social arrangements. The understanding of the underlying mechanisms during the introduction of new technologies in these Niches is of paramount importance for the societal learning.

4. Discussion

Transition towards clean coal technologies in India will encounter many uncertainties. The first step in understanding and managing the transition is to understand the uncertainties in the coal and power sector. To understand these uncertainties extensive literature review was employed (Van der Veen and Wilson 1997; Philibert and Podkanski 2005; Chikkatur 2008; COAL 2008). Figure 2a below classifies the uncertainties at the niche, regime and landscape level on the Y axis, and the level of its impact is shown on the X axis. Such framework is beneficial in understanding the uncertainties at different levels and also the impact they will have on the coal sector. At the macro level oil prices and global energy demand will determine the investment in coal technologies. For example, if oil prices are low then the attention will shift away

from the usage of coal, thus adversely affecting the investments and diffusion of clean coal technologies. At the meso level, environmental standards will play a major role in clean coal technology transition as standards are known to promote technology development (Patil, A.N et al. 2007). At the niche level uncertainties related to the qualities of the technology (such as the price, effectiveness, and availability) are important.

Governance at different levels (macro, meso and niche) for a successful transition towards clean coal is shown in figure 2b below. At the macro level India can participate in the international emissions trading, this will give incentives for the coal sector to reduce emissions. At the meso level, incentives for managers can be provided to improve the productivity of their enterprises and reduce emissions and is thus a precondition for sustainable environmental improvements. The governance should encourage competition and create incentives to improve the efficiency of each sector in the coal energy chain, as well as the efficiency of the chain as a whole (Van der Veen and Wilson 1997). At the niche level there should be policies that will allow easy technology transfer and allow the coal sector to experiment with new clean coal technologies.

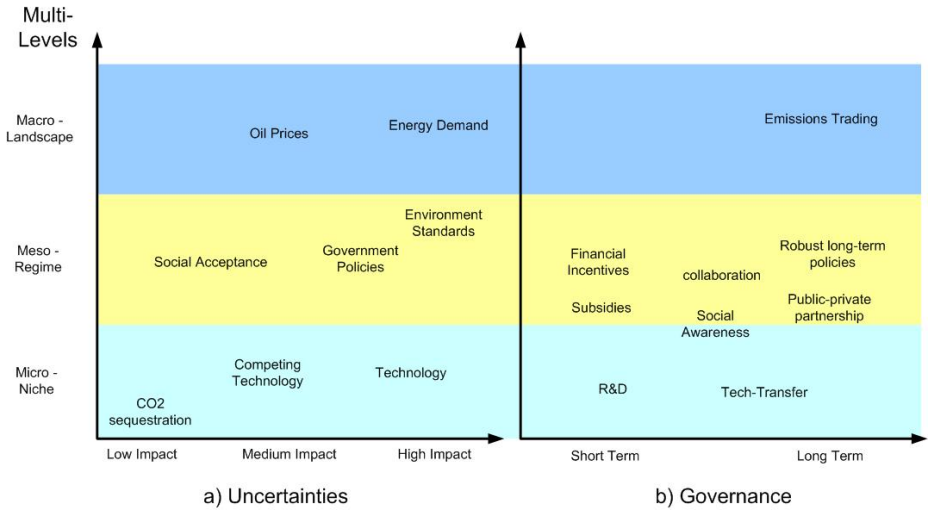


Figure 2: Uncertainties & Governance during the transition towards clean coal

5. Conclusion

The paper concludes that with the coordinated efforts at the national level and good practices at local level the transition towards clean coal technologies can be achieved. Clean coal technologies at the beginning will need protection via subsidies and financial incentives so that they are sustainable in the long run. This paper suggests that the technological improvements alone will not guarantee transition towards clean coal, but such efforts should be complemented by robust and stable policies at the national level. For example: Indian government should sort out the financial incentives for CO2 sequestration or reduction – this will encourage investments in the environment friendly technologies. To devise the incentives, Indian government could employ the concept of Niches. Niches could entail experimentation at local or project level in order to understand how clean coal technologies can overcome initial barriers of high cost, such

understanding is paramount in formulating a framework for the transition towards clean coal technologies in India.

References

Agnolucci, P. and W. McDowall (2007). "Technological change in niches: Auxiliary Power Units and the hydrogen economy." Technological Forecasting and Social Change 74(8): 1394-1410.

Brown, H. S., P. Vergragt, et al. (2003). "Learning for Sustainability Transition through Bounded Socio-technical Experiments in Personal Mobility." Technology Analysis & Strategic Management 15(3): 291 - 315.

Chikkatur, A. (2008). A Resource and Technology Assessment of Coal Utilization in India, Pew Center on Global Climate Change.

Christensen, C. (1997). The Innovator's Dilemma. Cambridge, Harvard Business School Press.

COAL (2008). Annual Report Ministry of Coal - Government of India.

Geels, F. (2001). Technological Transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. Nelson and Winter Conference, Aalborg, Denmark.

Geels, F. (2002). Understanding the Dynamics of Technological Transitions. Twente, Twente University Press.

Hekkert, M. P., R. A. A. Suurs, et al. (2007). "Functions of innovation systems: A new approach for analysing technological change." Technological Forecasting and Social Change 74(4): 413-432.

IEA (2007). Renewables in Global Energy Supply - An IEA Factsheet. Paris, International Energy Agency.

Kaijser, A. (1994). I fädrens spår: Den svenska infrastrukturens historiska utveckling och framtida utmaningar. Stockholm, Sweden, Carlsson.

Kaygusuz, K. (2009). "Energy and environmental issues relating to greenhouse gas emissions for sustainable development in Turkey." Renewable and Sustainable Energy Reviews 13(1): 253-270.

Kemp, R. (1994). "Technology and the transition to environmental sustainability : The problem of technological regime shifts." Futures 26(10): 1023-1046.

Marland, G. and T. Boden (2001). The Increasing Concentration of Atmospheric CO2: How Much, When, and Why? Erice International Seminars on Planetary Emergencies, 26th Session, Sicily, Italy.

Mulder, P., C. H. Reschke, et al. (1999). Evolutionary Theorising on Technological Change and Sustainable Development. European Meeting on Applied Evolutionary Economics. Grenoble, France, Institute for Energy Politics and Economics.

Patil, A. C., A. A.N, et al. (2007). Trends in Emission Standards for Public Transport Buses: How Low should they go. IAEE Asian Conference: Asian Energy Security and Economic Development in an Era of High Oil Prices, Taiwan.

Philibert, C. and J. Podkanski (2005). Clean Coal Technologies. International Energy Technology - Collaboration and Climate Change Mitigation, International Energy Agency.

Rotmans, J., R. Kemp, et al. (2001). "More evolution then revolution : transition management in public policy." Foresight - the journal of foresight studies, strategic thinking and policy 3(1).

Rotmans, J., M. van Asselt, et al. (2000). "Visions for a sustainable Europe." Futures 32(9-10): 809-831.

Utterback, J. M. (1974). "Innovation in Industry and the Diffusion of Technology." Science 183(4125): 620-626.

Van der Veen, P. and C. Wilson (1997). "A new initiative to promote clean coal." Finance and Development 34(4): 36-39.

Weijnen, M. and O. Bosgra (1998). An Engineering Perspective on the Design and Control of Infrastructure - Explorations into a Generic Approach to Infrastructure Scenario Analysis. The Infrastructure playing field in 2030. M. P. C. Weijnen and E. F. t. Heuvelhof. Noordwijk, Delft University Press.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Storage logistics of fruits and vegetables in distribution centers

Daniela F. Borghi[a], Reginaldo Guirardello[a], Lúcio Cardozo Filho[b]

[a] *School of Chemical Engineering, State University of Campinas – UNICAMP – P.O. Box 6066, 13083-970, Campinas – SP, Brazil*
[b] *Department of Chemical Engineering, University of Maringá – UEM – Colombo Avenue, 5790 – 87020-900, Maringá – PR, Brazil*

Abstract

This study has as main objective to minimize the losses of fruits and vegetables that happen during its storage in distribution centers, using mathematical models to optimize the products distribution in deposit and to minimize costs referring to the storage of these products. Data of keeping quality of some horticultural products were taken from the literature. The mathematical model for optimization used a mixed integer linear programming (MILP) formulation. It was implemented in the software GAMS® and the solver used was CPLEX®. Simulations were conducted considering various temperatures of storage and the same cost of U$ 1.00 for every analyzed products (celery, lettuce, leek, plum, beetroot, onion, carrot, chicory, cauliflower, spinach, papaya, strawberry, turnip, cucumber, peach, bell pepper, radish, parsley, tomato and grape), aiming to minimize the costs relating to its storage. Starting from values of $KQ_{ij(real)}$ (keeping quality of product *i* in zone *j* in real storage temperature) and $KQ_{i(ideal)}$ (keeping quality of product *i* in its ideal storage temperature) and considering two zones (refrigerating chamber and external deposit) it was possible to proceed an analysis in relation to keeping quality of each product in function of the temperature. The results showed that the more distant the real storage temperature of a product is of its ideal temperature, the greater the costs with the loss of quality of this product. It was verified that the cost relating to the loss of quality of products has great influence on the total storage cost. It shows the importance of optimizing the storage logistics of these products, so that these costs are not still more expressive, what certainly occur without a tool for that. It was also verified that the order of products priority to occupy the chamber depends on the zones temperatures combinations, but fruits as grape and plum have larger priority in majority combinations of temperatures, followed by peach and strawberry. Papaya, bell pepper, tomato and cucumber have great variation in its priority and for smaller chamber temperatures (T_c), the products with smaller priorities were cucumber and tomato and for larger T_c were cauliflower and celery. In that way, the proposed model comes as a tool that can help the distribution centers managers in taking decisions referring to the logistics of storage of fruits and vegetables in order to minimize total costs and losses of the sector.

Keywords: fruits and vegetables, keeping quality, cost of storage, optimization.

1. Introduction

The storage of fruits and vegetables has a great importance to the establishments that commercialize it, because they are very perishable and sometimes become a detriment to the establishment, due to the great quantity of product waived because of loss of

quality. To get an idea of the size of this problem in Brazil, in 2001, were harvested 15 million tons of vegetables, and more than 5 million tons were lost, enough to supply 53 million people and generate a loss more than US$ 1 million (Vilela et al., 2003).
Fruits and vegetables are susceptible to loss of keeping quality due to injuries or loss of moisture. The keeping quality of a product is the time until it becomes unacceptable for consumption, from the sensory, nutritional or safety point of view (Fu and Labuza, 1993).
The storage temperature is one of the most important factors that influence the keeping quality of fruits and vegetables. Low temperatures reduce the respiration rate, the speed of growth of pathogenic microorganisms and loss of moisture, and delay the senescence and changes in color and texture (Borghi, 2008).
It is known that the management of a deposit of fruits and vegetables is more complex than a deposit of processed products as the first ones need rapid commercialization and special conditions of storage to reduce its loss of quality, because it reflects in reduction of its commercial value.
The main objective of this study is to verify the influence of temperature on the storage logistics of fruits and vegetables in distribution centers, using a mathematical model to optimize the distribution of the products and to minimize the costs related to its storage.

2. Mathematical Model

To develop the mathematical model were needed some data describing the physical space available, the volume of product to be stored, the conditions under which products must be kept and the location of these within the deposit. Knowing this information, it was possible to determine which area in the accommodation of the deposit, with specific conditions of storage that a product should take. This area was called "zone" by Broekmeulen (1998). It was considered that in a distribution center deposit there are basically two zones, the external deposit, subject to room temperature, and the refrigeration chamber, maintained at a pre-established temperature.
Three assumptions were considered. The first one is that the storage time and temperature are the only storage conditions that influence the reduction of products keeping quality. The second is that the storage temperature depends on the zone. And the last one is that the stock levels and the models of keeping quality change are known for all products.
Define the decision variable x_{ij} (binary variable) for all $i \in P$ and $j \in Z$:

$$\begin{cases} x_{ij} = 1 \text{ if the product } i \text{ is designated to zone } j \\ x_{ij} = 0 \text{ in any other case} \end{cases}$$

Thus, the problem may be placed in the form of a mixed integer linear programming (MILP) (Weng, 2008). Then, the problem is to find a plan for distribution of fruits and vegetables that minimizes the cost function (C) given by:

$$C = \sum_{i \in P} \sum_{j \in Z} \frac{c_i . s_i}{d_{ij}} . x_{ij} + \sum_{i \in P} \sum_{j \in Z} \frac{b_{ij} . c_i . s_i}{d_{ij}} x_{ij} + \sum_{i \in P} \sum_{j \in Z} \frac{q_{ij} . c_i . s_i}{d_{ij}} . x_{ij} + \sum_{j \in Z} a_j . S_j \quad (1)$$

and fulfilling the following restrictions:
1. Each product can be assigned to only one specific zone.
2. The storage capacity of the zone can not be exceeded.

3. The time that the product *i* stays in zone *j* can not exceed the keeping quality of this product.

The cost function (*C*) depends on several factors, the parameters of the model. The set *P* includes all fruits and vegetables studied and each receives an index *i*. The set of zones (*Z*) encompasses all zones *j*, each with specific conditions and a physical capacity of storage (S_j).

The parameter a_j is the storage cost of zones. Keep a cold chamber requires an expense of extra energy compared to the deposit. For this reason, the costs to storage of the products in the deposit and the cold chamber are different. The percentage of products loss (b_{ij}) depends on many factors such as transport, handling of the products and inadequate storage, which added, lead to loss of the products. The acquisition cost of the products (c_i) is the price paid per kilo or unit of the product. The parameter d_{ij} is the time that the product is stored, which was considered equal to a day for all products.

Another parameter of the model is q_{ij} (fraction of the original quality that is lost when compared with the storage in ideal conditions for a product). This parameter is obtained through the following relationship:

$$q_{ij} = 1 - \frac{KQ_{ij(real)}}{KQ_{i(ideal)}} \tag{2}$$

where $KQ_{ij(real)}$ and $KQ_{i(ideal)}$ is the keeping quality of product *i* in the real storage temperature and the keeping quality of product *i* in its ideal temperature, respectively.

Tijskens and Polderdijk (1996) developed a useful formulation for calculating the keeping quality of fruits and vegetables. They showed that the keeping quality is proportional to the inverse of reaction rate *k*, independent of the kinetic mechanism of the loss of product quality. Thus, it is possible to describe the behavior of keeping quality as a function of temperature. To Labuza (1984), the reactions of loss of quality of food can be described by the Arrhenius Law, where specific reaction rates depend on temperature. The reaction rate *k* can be approximated by this Law:

$$k = k_{ref} e^{\frac{E_A}{R}\left(\frac{1}{T_{ref}} - \frac{1}{T_{abs}}\right)} \tag{3}$$

In many fruits and vegetables, the quality attribute which limits consumers acceptance, changes at a different temperature. This can be seen in cold sensitive products. In tomatoes, for example, kept at constant temperatures below 8 °C, the limiting factor is usually color, but when the temperature is above 13 °C, the limiting factor is firmness.

The keeping quality proposed by Tijskens and Polderdijk (1996) under constant conditions, may be represented by:

$$KQ = \frac{KQ_{ref}}{\sum_{n=1}^{N} k_{ref(n)} e^{\frac{E_{A(n)}}{R}\left(\frac{1}{T_{ref}} - \frac{1}{T_{abs}}\right)}} \tag{4}$$

where *N* is the process number that contribute to the keeping quality, which normally is not greater than two.

Assuming $k_{ref(1)}$ (reaction rate of the first component in the reference temperature) equal to one, statistical analysis was made possible without the need for information on the

type of quality function. Moreover, the expression for the keeping quality becomes almost independent of the type of kinetic mechanism involved.
The loss of product quality is usually measured by analysis of one or more parameters which may be physical, chemical, microbiological or sensorial (Taoukis and Labuza, 1989). It is possible to determine the keeping quality of a product at a reference temperature (KQ_{ref}) using methods for assessing quality, which may be measurement of color and firmness, for example (Mitcham et al., 1996).
This method is generic and applicable to products of moderate and tropical areas. This model can be used, therefore, to predict the keeping quality of fruits and vegetables stored in a supermarket. The proposed equation was implemented in the software GAMS® and optimizations were made leading into account different temperatures of the cold chamber and the deposit and variability in quantities of products stored in the distribution center, using the CPLEX® solver.

3. Results and Discussion

Optimizations were performed using Equation (1) and the restrictions, taking into account different temperatures for zones and different quantities of products purchased by the distribution center. Initially, the values of $KQ_{ij(real)}$ and $KQ_{i(ideal)}$ were obtained using Equation (4) and data estimated by Tijskens and Polderdijk (1996). It was found that for products with a number of process equal to one (N = 1) the keeping quality increases as temperature decreases. However, for products with number of process equal to two (N = 2), both in low as in high temperature, the keeping quality is small and increases as it approaches the ideal temperature of storage.
From these data was obtained the change of the quality factor, q_{ij}, with the temperature by Equation (2). The calculated results are shown in Figure 1.
The analysis of results shows that when the real storage temperature of each product is equal to its ideal temperature, q_{ij} = 0, there is no cost to the loss of quality of this product.
The maximum value of q_{ij} is 1 and the higher its value, more distant of ideality a product is stored and the greater the costs for its loss of quality.
The priority order of the studied products to stay in the cold chamber was also evaluated. For that, it was considered that the amount of each product was equal a box and the capacity of the chamber was increased one by one box to see what would the first product in the chamber. This procedure was performed for many combinations of chamber temperature (T_c) and stock temperature (T_e), with T_c of 5 °C to 15 °C and T_e of 10 °C to 33 °C, both ranging from one by one degree. The price paid for each product used in these optimizations was U$ 1.00 for all products.
Analyzing the results, it was observed that the grape, the first in order of priority in lower T_c, loses its place for the papaya, as T_c increases. The plum, peach and strawberry have its priority reduced because papaya and tomatoes pass to enter before in the chamber in higher T_c (T_c is approaching the T_{ideal} of these fruits).
The radish loses priority for the onion, bell pepper and cucumber, which pass to enter before in the chamber in higher T_c. The spinach lost its place to turnip and parsley to spinach, turnip and beetroot in higher T_c.
The results described above do not take into account the third restriction. According to this restriction, the keeping quality of the product i in zone j must be greater than a minimum period of residence of this product in this zone. So, for some values of T_e some products should necessarily be stored in cold chamber, as quickly ruined in stock, taking its keeping quality lower than d_{ij}. Optimizations were carried out taking into

account the third restriction. The results show that for a T_e up to 17 °C, independent of T_c, there is no product that is unable to remain in stock. In these situations, the results are very similar to those obtained in the optimizations that do not consider this restriction. For other cases the results are presented in Table 1.

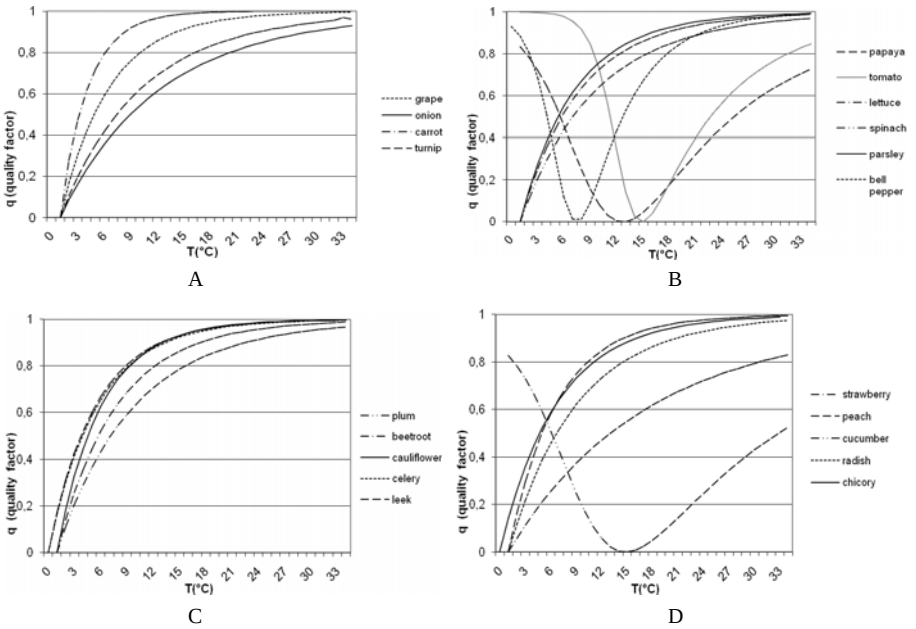


Figure 1. Change in quality factor of products with temperature. For all the Figures (A, B, C and D), when the product is stored exactly at its ideal temperature, q_{ij} is equal to zero, and it value increases as more distant is the storage temperature of ideal temperature. Figures A and C: the storage ideal temperature of all the products is 0 °C. Figure B: the storage ideal temperature of papaya, tomato and bell pepper is 12 °C, 14 °C and 7.5 °C, respectively, and for the other products it is equal to 0 °C. Figure D: the storage ideal temperature of cucumber is 14 °C and for the other products it is equal to 0 °C.

Table 1. Products that should be stored in cold chamber due to the third restriction.

T_e	Products that should be stored in cold chamber
18	peach, lettuce, carrot, radish
20	peach, lettuce, carrot, spinach, radish, celery, leek
22	peach, lettuce, carrot, cauliflower, spinach, radish, parsley, celery, leek
24	peach, lettuce, beetroot, carrot, cauliflower, spinach, radish, parsley, celery, chicory, leek, bell pepper
26	strawberry, peach, lettuce, beetroot, carrot, cauliflower, spinach, radish, parsley, celery, chicory, leek, bell pepper
28	strawberry, peach, lettuce, beetroot, carrot, cauliflower, spinach, radish, parsley, celery, chicory, leek, bell pepper
30	plum, strawberry, peach, lettuce, beetroot, carrot, cauliflower, spinach, radish, parsley, celery, chicory, leek, bell pepper
32	plum, strawberry, peach, lettuce, beetroot, carrot, cauliflower, spinach, radish, parsley, celery, chicory, leek, bell pepper
33	plum, strawberry, peach, grape, lettuce, beetroot, carrot, cauliflower, spinach, radish, parsley, celery, chicory, leek, bell pepper

Looking at the table above and the Figure 1, it is possible to observed that the order of entry of products into the cold chamber is affected by the third restriction and by the temperatures T_c and T_e. The importance of Restriction 3 is perceived, as if it were ignored many products remain in stock, even taking its keeping quality in that zone ($KQ_{ij\ (real)}$) less than the minimum keeping quality necessary for this product (d_{ij}).
In the optimizations performed in this work, the parameter d_{ij} was considered equal to one for all products, this means that the minimum time of residence of all the products in any of the zones should be one day. It was noted that, for example, lettuce, cauliflower and celery, which have the keeping quality less than a day in stock even for T_e not too high, are some of the latest products in order of priority to remain in the cold chamber when optimizations were performed not considering this restriction.

4. Conclusions

The proposed model is a tool that helps reduce the total cost of the establishment related to the storage of fruits and vegetables.
The equation proposed by Tijskens and Polderdijk (1996) was a useful tool to predict the keeping quality of the products in different storage temperatures.
It was found that the temperature in the chamber and the temperature of the stock significantly affect the order of priority to the products be in the cold chamber and that is very important to consider the keeping quality of each product in a certain temperature before deciding on the logistics of these products within the deposit.
The method is generic and therefore can be used as a tool to assist in decision making related to the logistics of storage of fruits and vegetables in supermarket.

References

D. F. Borghi, 2008, Logística de armazenamento de frutos e hortaliças em supermercado, 203p. Dissertation (Masters in Chemical Engineering) – School of Chemical Engineering, State University of Campinas, Campinas.

R. A. C. M. Broekmeulen, 1998, Operations Management of Distribution Centers for Vegetables and Fruits. International Transactions in Operational Research. vol. 5, n. 6, 501-508.

B. Fu; T. P. Labuza, 1993, Shelf-life Prediction: Theory and Application. Food Control, Butterworth-Heinemann, vol. 4, n. 3, 125–133.

T. P. Labuza, 1984, Application of Chemical Kinetics to Deterioration of Foods. Journal of Chemical Education, vol. 61, n. 4.

B. Mitcham; M. Cantwell; A. Kader, 1996, Methods for Determining Quality of Fresh Commodities. Perishables Handling Newsletter Issue, n. 85.

P. S. Taoukis; T. P. Labuza, 1989, Applicability of time-temperature indicators as shelf-life monitors of food products. Journal of Food Science, vol. 54, n. 4, 783–788.

L. M. M. Tijskens; J. J. Polderdijk, 1996 A Generic Model for Keeping Quality of Vegetable Produce During Storage and Distribution. Agricultural Systems, vol. 51, n. 4, 431–452.

N. J. Vilela; M. M. Lana; E. F. Nascimento; N. Makishima, 2003, O peso da perda de alimentos para a sociedade: o caso das hortaliças. Hortic. Bras. vol. 21, n. 2, 142–144.

M. Weng, Integer Programming. In: A. R. Ravindran, *Operations Research and Management Science Handbook.* CRC Press, 2008.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal Control Law Development in a Sequential Batch Reactor through Mixed Integer Particle Swarm Dynamic Optimization

Adrián Ferrari[a], Soledad Gutierrez[a], Evaristo C. Biscaia[b]

[a]*IIQ - Department of Reactors Engineering, Engineering School - University of the Oriental Republic of Uruguay, Julio Herrera y Reissig 565, CP 11300, Montevideo, Uruguay.*
[b]*LMSCP/PEQ/COPPE/UFRJ, Av. Horácio Macedo, 2030 - Centro de Tecnologia - Bloco G - Sala G-115 - CEP 21941-914 - Rio de Janeiro – RJ, Brazil.*

Abstract

Dynamic optimization in Sequential Batch Reactors for wastewater treatment represents an enormous challenge in order to improve the time and energy management in real industries. The non convex behaviour presented by these systems limits the application of deterministic techniques to optimize this kind of equipment. Although any real example has been found in the open literature, stochastic contributions to meet global optimization goals in the SBR technology appear to be a promising strategy. A Particle Swarm Optimization (PSO) algorithm in order to minimize the total aeration demand (energy management) in a real dairy SBR process (operating in an Intermittent Aeration Extended Filled configuration) for carbon and nitrogen removal was here developed. The batch reactor network size, sequencing and stages duration, were assumed as the decision variables for the dynamic MINLP problem. Two kinds of PSO algorithms (relaxed; and mixed-integer) were here tried in order to find the best way to take into account the mixed-integer nature of the system. In addition, PSO results were compared with the ones obtained by a sequential shooting method/NLP Code, considering different initial guesses. Stochastic optimization allowed improving the obtained results from deterministic techniques with a very reasonable computational cost, and mixed-integer PSO resulted in the best structure solving the MINLP problem. Despite of that, and in order to assure the most robust and reliable results for this optimization problem (high swarm convergence and good optimization efficiency), the assessment of both PSO formulations must be considered. The PSO results have given an optimal control law that minimizes the total aeration demand along the SBR operation; this law can be easily implemented (switching times).

Keywords: SBR, Dynamic Optimization, Particle Swarm Optimization.

1. Introduction

Sequential Batch Reactors (SBR) represents one of the most important activated sludge technologies for wastewater (WW) treatment. Carbon, nitrogen and phosphorous removal can be carried out simultaneously in this kind of equipment (Artan and Orhon; 2005). The extreme flexibility presented by these systems [feeding pattern and reaction network establishment (Ferrari et. al[(1)]; 2008)] encourages the use of optimal control laws even during its operation at industrial scale. This feature is rarely observed for the rest of WW treatment technologies. Figure 1 shows an example of a usual SBR process treating continuous WW for carbon and nitrogen removal.

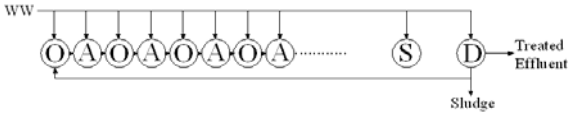


Figure 1: Intermittent Aeration Extended Filled (IAEF) Reactor (Artan and Orhon, 2005). Stages: **O**, Oxic (Aerobic); **A**, Anoxic; **S**, Settling; and **D**, Draw.

Dynamic optimization in Sequential Batch Reactors represents an enormous challenge in order to improve the time and energy management in real WW treatment plants. The non convex behaviour presented by these systems limits the application of deterministic techniques to optimize this kind of equipment. Although any real example have been found in the open literature (Souza et. al; 2008, Coelho et. al; 2001), stochastic contributions to meet global optimization goals in the SBR technology appears to be a promising strategy. Very few papers related with energy optimization in this kind of equipment could be found. Despite the time management represents the usual objective function for the optimization problem in Sequential Batch Reactors (Souza et. al; 2008, Coelho et. al; 2001, Ferrari et. al[(2)]; 2008), new insights of optimal control laws can be obtained through the analysis of their energy management. The main purpose of this contribution is to develop an optimal control law in order to minimize the total aeration demand (energy management) along the operation of a real SBR process, through the application of Particle Swarm Optimization (PSO) (Kennedy and Eberhart; 1995). In addition, a comparison between PSO results with the ones obtained from a sequential shooting method/NLP Code, in order to verify the non convex behaviour of the system, is presented in this paper.

2. Mathematical Modelling and Applied Methods

2.1. SBR Model and Kinetic Description

The dimensionless transient mass balance applied to each component in the system is presented below (lumped structure):

$$\frac{dc_i(t)}{dt} = Fe_{SBR} \cdot \frac{q(t)}{v(t)} \cdot \left[c_{f,i}(t) - c_i(t)\right] - Da_i \cdot r_i\left[\boldsymbol{c}(t)\right]$$

Where $\boldsymbol{c_i}$ is the component *i* concentration; $\boldsymbol{c_{f,i}}$, the component *i* feed concentration; $\boldsymbol{t}$, the time variable; $\boldsymbol{q}$, the volumetric wastewater flow rate; $\boldsymbol{v}$, the reactor volume; $\boldsymbol{r_i}$, the component *i* consumption rate; $\boldsymbol{Da_i}$, the Damköhler number for component *i*; $\boldsymbol{Fe_{SBR}}$, the Feeding number; and $\boldsymbol{c}$, the column vector of the concentration of all components. More details regarding dimensionless SBR models are presented in Ferrari et. al[(1)] (2008). Aerobic/anoxic carbon removal and autotrophic nitrification represents the main processes involved in this equipment. The biomasses growth kinetics adopted in this work are in accordance to ASM1 Model (Orhon and Artan; 1994) and their complete description is presented below.

Anoxic Carbon Removal (Denitrification)

$$r_{X_H} = (\mu_{hN} - kd_{hN}).X_H \qquad \mu_{hN} = \mu_{hNmax}\left(\frac{S_C}{K_{C1}+S_C}\right)\left(\frac{S_{NO}}{K_{NO}+S_{NO}}\right) \qquad r_{S_{NO}} = \frac{1}{Y_{\frac{X}{N}}}.\mu_{hN}.X_H + kd_{hN}.X_H.f_{\frac{N}{X}}.\left(\frac{S_{NO}}{K_{NO}+S_{NO}}\right)$$

$$r_{S_C} = \frac{1}{Y_{hN}}.\mu_{hN}.X_H$$

Aerobic Carbon Removal

$$r_{X_H} = (\mu_h - kd_h).X_H \qquad \mu_h = \mu_{hmax}\left(\frac{S_C}{K_{Ch}+S_C}\right)\left(\frac{S_O}{K_O+S_O}\right) \qquad r_{S_C} = \frac{1}{Y_H}.\mu_h.X_H$$

Autotrophic Nitrification

$$r_{X_A} = (\mu_a - kd_a).X_A \qquad \mu_a = \mu_{amax}\left(\frac{S_{NH}}{K_{NHa}+S_{NH}}\right)\left(\frac{S_O}{K_{Oa}+S_O}\right) \qquad r_{S_{NH}} = \frac{1}{Y_{NH}}.\mu_a.X_A \qquad r_{S_{NO}} = -\frac{1}{Y_{NO}}.\mu_a.X_A$$

Table 1: Nomenclature Description for Kinetic Expressions

Symbol Components	Name	Value	Dimension
S_C	Soluble Chemical Oxygen Demand ($COD_{Soluble}$)	-	$gCOD.m^{-3}$
S_{NO} / S_{NH}	Nitrate (NO_3^-) / Ammonium (NH_4^+) Concentration	-	$gN.m^{-3}$
S_O	Dissolved Oxygen (DO)	-	$gO_2.m^{-3}$
X_H / X_A	Heterotrophic / Autotrophic Biomass Concentration	-	$gVSS.m^{-3}$
Kinetic Parameters			
μ_{hmax}	Maximum Specific Aerobic Growth Rate for X_H	1	d^{-1}
μ_h	Specific Aerobic Growth Rate for X_H	-	d^{-1}
K_{Ch}	Aerobic Saturation Constant for S_C	75	$gCOD.m^{-3}$
K_O / K_{Oa}	Saturation Constant for S_O	~ 0	$gO_2.m^{-3}$
kd_h / kd_a	Aerobic Decay Coefficient for X_H / X_A	0.05 / 0.025	d^{-1}
μ_{hNmax}	Maximum Specific Anoxic Growth Rate for X_H	0.75	d^{-1}
μ_{hN}	Specific Anoxic Growth Rate for X_H	-	d^{-1}
K_{C1}	Anoxic Saturation Constant for S_C	75	$gCOD.m^{-3}$
K_{NO}	Saturation Constant for S_{NO}	15	$gN.m^{-3}$
kd_{hN}	Anoxic Decay Coefficient for X_H	0.05	d^{-1}
μ_{amax}	Maximum Specific Aerobic Growth Rate for X_A	0.5	d^{-1}
μ_a	Specific Aerobic Growth Rate for X_A	-	d^{-1}
K_{NHa}	Aerobic Saturation Constant for S_{NH}	1	$gN.m^{-3}$
Yield Parameters			
Y_H	Aerobic Heterotrophic Biomass Yield Coefficient	0.42	$gVSS.gCOD^{-1}$
Y_{NH}	Aerobic Autotrophic Biomass Yield Coefficient	0.170	$gVSS.(gN\text{-}NH_4^+)^{-1}$
Y_{NO}		0.174	$gVSS.(gN\text{-}NO_3^-)^{-1}$
Y_{hN}	Anoxic Heterotrophic Biomass Yield Coefficient	0.42	$gVSS.gCOD^{-1}$
$Y_{X/N}$		1.67	$gVSS.(gN\text{-}NO_3^-)^{-1}$
$f_{N/X}$	Chemical Nitrate Demand for X_H	0.36	$(gN\text{-}NO_3^-).gVSS^{-1}$
Reaction Rates			
r_{XH} / r_{XA}	Growth Rate for X_H / X_A	-	$gVSS.m^{-3}.d^{-1}$
r_{SC}	Removal Rate for S_C	-	$gCOD.m^{-3}.d^{-1}$
r_{SNO} / r_{SNH}	Removal Rate for S_{NO} / S_{NH}	-	$gN.m^{-3}.d^{-1}$

2.2. Operational Data

The table 2 summarizes the conditions presented by the SBR process. They correspond to the mean wastewater composition of a real dairy cheese factory with a sole reactor operating in an IAEF configuration with carbon and nitrogen removal purposes.

Table 2: Plant Conditions

WW Composition & Flow Rate [SBR Inlet](*)	2000 $mgCOD_{Total}/L$; 140 $mgN\text{–}NH_4^+/L$; 20 $mgN\text{-}NO_3^-/L$ & 800 m^3/d		
Cycle Time // Total Cycles by Day	12 h. // 2		
Reactor Sequencing for Each Cycle [Fixed]	Stage 1: Reaction Time 11 h.	Stage 2: Settling 0.5 h.	Stage 3: Draw 0.5 h.
Reactor Volume at Cycle Beginning	3000 m^3		
X_H & X_A at Cycle Beginning	2500 & 25 mgVolatile Suspended Solids(VSS)/L [experimental data]		
Sludge Retention Time	> 20 days		
DO Concentration	Not growth limiting (DO > 2 mgO_2/L)		

(*) Anaerobically pre-treated effluent

2.3. Constrained Dynamic MINLP Problem

The total aeration demand in the SBR cycle was assumed as the scalar performance index to be minimized. The batch reactor network size (integer variable), sequencing (binary variable: oxic/anoxic) and the stages duration (continuous variables), were assumed as the decision variables for the dynamic MINLP problem. These inputs allow constructing an optimal control law that can be easily implemented at industrial scale (switching times in the equipment). In addition to some physical path constraints, the environmental restrictions respect to the draw effluent composition defines the feasible region of the optimization problem. In this sense, the following inequalities terminal constraints must be satisfied: $COD_{Soluble}$<60 mg/L, NH_4^+<5 mgN/L, NO_3^-<10 mgN/L. The complete constrained dynamic MINLP problem to be solved is presented below:

$$\min_{u(t)} \; J = \sum_{k} t_{Oxic,k}$$

$$\text{s.t.:}$$

$$\dot{c} = F(c,u) \qquad c(0) = c_0$$

$$S(c,u) \le 0 \qquad T\left[c(t_f)\right] \le 0$$

Where **J** is the objective function; $\mathbf{t_{Oxic,k}}$, the aeration time for the *k* stage in the SBR; **c**, the five-dimensional vector of states (COD, NH_4^+, NO_3^-, X_H, X_A) with known initial conditions $\mathbf{c_0}$ (60, 5, 10, 2500, 25); **u**, the m-dimensional vector of inputs (decision variables); **S**, the z-dimensional vector of path constraints; **T**, the three-dimensional vector of terminal constraints in accordance to the environmental regulations; **F**, the vector function who defines the ODE system (see *SBR Model and Kinetic Description*); and $\mathbf{t_f}$, the cycle time assumed fixed for this equipment (see Table 2). The function **u(t)** is parameterized in accordance to a piecewise constant control (Srinivasan et. al; 2003). The following three main inputs were chosen for such parameterization: **N**, the reaction network size (integer variable defining the number of O/A sub-sequences in the network); **I**, a binary variable who defines the first kind of reaction appearing in the process [oxic (**I = 1**); anoxic (**I = 0**)]; and $\boldsymbol{\tau_{First}}$, the duration of the stage located first at each O/A sub-sequence along the reaction network (continuous variable). To simplify the implementation of the optimal control law in the real equipment (PLC programming), every O/A sub-sequence in the process was assumed presents the same switching times. Thus, the objective function for this work can be represented as: $\mathbf{J = N.[I.\tau_{First} + (1\text{-}I).(Reaction\ Time/N - \tau_{First})]}$, and the following path constraints were assumed for the inputs: $\mathbf{1 \le N \le 20}$; $\mathbf{0 < \tau_{First} \le}$ **Reaction Time/N**; and $\mathbf{I.(Reaction\ Time/N - \tau_{First}) + (1\text{-}I).\tau_{First} \le 30\ min.}$. The last constraint must not be eluded in order to avoid the biomass settling during each anoxic stage (reactor not provided with an independent mixer). An adequate mixing in the system by short times once the aeration stopped (maximum ~ 30 minutes), can be achieved as a consequence of the N_2 generation (gas bubbles) during the denitrification process.

2.3.1. Deterministic Methods

Sequential shooting techniques/NLP Codes [Levenberg-Marquardt, Conjugate-Gradient, and Quasi-Newton (with steepest descent/linesearch algorithms)] (Srinivasan et. al; 2003, Biegler and Grossmann; 2004) were here employed to solve the optimal control law. It should be pointed out that the mixed-integer nature in the optimization problem was reached through the inclusion of additional constraints for **N** and **I** inputs. As a MINLP problem must be solved, the use of variational dynamic optimization methods (Pontryagin's Minimum Principle; Optimality Principle of Hamilton – Jacobi - Bellman; etc.) is not allowable (Srinivasan et. al; 2003, Biegler and Grossmann; 2004).

2.3.2. Stochastic Methods

Two kinds of PSO structures were here developed in order to calculate the optimal inputs: **mi-PSO** (mixed-integer PSO) and **r-PSO** (relaxed PSO). The difference resides in the way the mixed-integer nature in the optimization problem is accounted for. While **mi-PSO** restricts the input nature at each stage in the swarm movement, **r-PSO** permits a relaxed particle motion introducing the mixed-integer nature just when the objective function is computed. In order to meet the path and terminal constraints of the dynamic MINLP problem, the following augmented objective function is assumed for both PSO algorithms.

$$J = \begin{cases} N.\left[I.\tau_{First} + (1-I).\left(\dfrac{\text{Reaction Time}}{N} - \tau_{First} \right) \right] & \text{if } (N, I, \tau_{First}) \in \text{Feasible Region} \\ J_{Maximum} = \text{Reaction Time (Fixed Value)}_{\text{(see Table 2)}} & \text{Otherwise} \end{cases}$$

3. Results and Discussion

3.1. Deterministic Optimization

Table 3 summarizes the optimal input values and the total aeration demand corresponding to four reaction networks obtained from different initial guesses and NLP codes. These results reveal the non convex behaviour presented in the Sequential Batch Reactor technology. Consequently, the application of deterministic techniques to optimize this kind of equipment will be only helpful if they are preceded by the use of an adequate global optimization method (heuristic, stochastic, etc.). In addition, if a powerful technique is employed to search for the global optimum [Genetic Algorithms, Scatter Search, Ant Colony Optimization, Particle Swarm Optimization, etc.], further deterministic contributions may not be necessary.

Table 3: Sequential Shooting Results

Shoot	NLP Code(*)	N	I	Reaction Network	τ_{First} (min)	Total Aeration Needs [J/J_{Max}] (%)
1	Conjugate-Gradient	10	0	$AO(AO)_8AO$	30.0	54.5
2		13	0	$AO(AO)_{11}AO$	30.0	40.9
3		16	1	$OA(OA)_{14}OA$	16.9	40.9
4		19	1	$OA(OA)_{17}OA$	14.1	40.6

(*) – Method that Optimize the Computational Cost.

3.2. Stochastic Optimization

As a constrained optimization problem must be solved, a large scale particle swarm should be generated in order to assure in the zero iteration, the random allocation of a minimal population sample (at least ~ 10 particles) inside the feasible region. As a result, and once the algorithm in course, the whole particle swarm would be gradually pushed inside the feasible region improving progressively their exploratory capabilities. Populations with 500 individuals, departing from repose and along 30 iterations were here used for both developed PSO algorithms. Table 4 shows the stochastic optimization results for **mi-PSO** and **r-PSO**.

Table 4: Particle Swarm Optimization Results (after 30 iterations)(*)

Method	N	I	Reaction Network	τ_{First} (min)	Total Aeration Needs [J/J_{Max}] (%)	Particles Inside Feasible Region [PIFR] (%)	Swarm Convergence Index [% of Population](**)
mi-PSO	15	0	$AO(AO)_{13}AO$	27.2	38.3	100	89.6
r-PSO	17	0	$AO(AO)_{15}AO$	23.9	38.4	100	98.6

(*) – Swarm Temporal Step **(Δt)** = 1; and Swarm Deadening Constant (**K**) = 2 for both PSO structures.

(**) **– Swarm Convergence Index (SCI) = 100.Total Converged Particles/Swarm Size**. It is assumed a Converged Particle when $|J_{Best,Particle} - J_{Leader,swarm}| < SCT.J_{Leader,swarm}$; ***SCT*** = Swarm Convergence Tolerance (assumed 0.1 in this table).

These results show that stochastic optimization improved the obtained results from deterministic techniques with a very reasonable computational cost (no further optimization improvement is needed). In addition, both PSO structures could allocate the complete swarm inside the feasible region (**PIFR = 100%**) demonstrating the enormous potential that this technique has. For identical computational efforts (after 30 iterations), the mixed-integer algorithm resulted slightly better than the relaxed one in order to optimize the constrained dynamic MINLP problem presented in this paper ($J_{miPSO} < J_{rPSO}$). However, **r-PSO** presented a higher swarm convergence index (**SCI**) than **mi-PSO** revealing very different natures in their exploratory capabilities. The **SCI** comparison strictly depends on the ***SCT*** value adopted for the calculations. Figure 2 shows a more exhaustive contrast between both PSO structures analyzing their complete response (**SCI & PIFR** indexes) along the course of the iterations. This data reveals very interesting additional results. Firstly, the higher **SCI** value for **r-PSO** is also observed during the whole course of the algorithm. Moreover, **mi-PSO** seems never reach a complete swarm convergence potential (**SCI = 100%**) because this performance index, even for different ***SCT*** values, always tend to a low asymptotic flat value as the iterations increases.

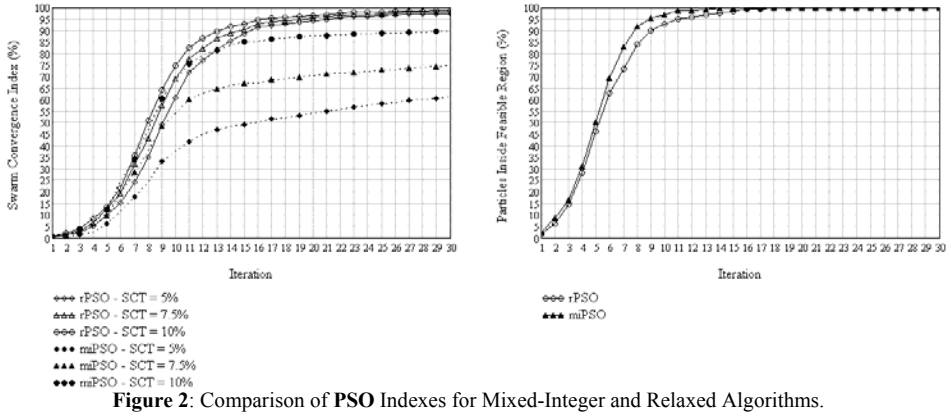


Figure 2: Comparison of **PSO** Indexes for Mixed-Integer and Relaxed Algorithms.

Finally, in order to reduce the computational cost for both structures, it can be observed that 15 – 20 iterations could be acceptable enough in order to get good optimization results (**PIFR** ~ **100%**; and **SCI** profiles already stabilized). In order to assure the most robust and reliable results for the MINLP problem presented in this work (high swarm convergence and good optimization efficiency), the assessment of both algorithms (**r-PSO** and **mi-PSO**) must be considered.

4. Conclusions

Optimal control laws obtained from deterministic and stochastic techniques were here developed for a real SBR process. New insights were obtained through the analysis of the energy management (aeration demand) in the equipment. Deterministic optimization results exposed a strong non convex behaviour for this system demonstrating the unavoidable need of stochastic contributions in order to search for the global optimum. The use of Particle Swarm Optimization allowed improving the obtained results from deterministic techniques with a very reasonable computational cost (no further optimization improvement was needed). In order to solve the constrained dynamic MINLP problem presented in this work and despite higher swarm convergences were obtained with a relaxed PSO formulation, better optimization results were achieved with a mixed-integer particle swarm structure, introducing the MI nature of the inputs at each stage in the algorithm. The assessment of both PSO formulations should be considered in order to guarantee the most robust and reliable results for this optimization problem. An optimal control law who minimizes the total aeration demand along the SBR operation can be easily programmed (switching times) in accordance to PSO results.

References

Artan, N.; Orhon, D. (2005). *Mechanism and Design of Sequencing Batch Reactors for Nutrient Removal*. Scientific and Technical Report No. 19, IWA Publishing.

Biegler, L.T. and Grossmann, I.E. (2004); Retrospective on Optimization. *Computers and Chemical Engineering*, 28, 1169–1192.

Coelho, M.A.Z.; Russo, C. and Araújo, O.Q.F. (2001); Optimization of a Sequencing Batch Reactor for Biological Nitrogen Removal. *Water Research*, 34(10), 2809-2817.

(1)Ferrari, A.; Biscaia Jr., E.C.; Melo, P.A. (2008); Sequential Batch Reactor and Plug Flow Reactor Network Comparison under Dynamic Conditions for Wastewater Treatment. *Proceedings of the **10th International Chemical and Biological Engineering Conference CHEMPOR 2008*** (Braga – Portugal), 138 – 146.

(2)Ferrari, A.; Gutiérrez, S.; Benítez, A.; Canetti, R. (2008); Modelado Cinético y Desarrollo de Ley de Control para Reactores Batch Secuenciales Lácteos; *Memorias del **IV Encuentro Regional de Ingeniería Química/2º Feria de la Industria Química y de Procesos EXPOQUIM 2008*** (Montevideo – Uruguay). Paper in Spanish.

Kennedy, J. and Eberhart, R. (1995); Particle Swarm Optimization. *Proceedings of **IEEE International Conference on Neural Networks*** (Perth - Australia), 1942-1948.

Orhon, D.; Artan, N. (1994). *Modelling of Activated Sludge Systems.* Technomic Publishing Company Inc.

Souza, S.M.; Araújo, O.Q.F. and Coelho, M.A.Z. (2008); Model-based optimization of a sequencing batch reactor for biological nitrogen removal. *Bioresource Technology*, 99(8), 3213-3223.

Srinivasan, B.; Palanki, S. and Bonvin, D. (2003); Dynamic optimization of batch processes. I. Characterization of the nominal solution. *Computers and Chemical Engineering*, 27, 1–26.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Addressing Long-Term Biorestoration in Eutrophic Lakes as an Optimal Control Problem, under Different Scenarios

Vanina Estrada [a], Elisa R. Parodi [b], M. Soledad Diaz [a]

[a]*Planta Piloto de Ingeniería Química (PLAPIQUI), Universidad Nacional del Sur-* CONICET, *Camino La carrindanga Km 7, Bahía Blanca 8000, Argentina, sdiaz@plapiqui.edu.ar*
[b]*Instituto Argentino de Oceanografía (IADO), Universidad Nacional del Sur-CONICET, Camino La carrindanga Km 7, Bahía Blanca 8000, Argentina,*

Abstract

In this work we propose an ecological water quality model for a reservoir, based on first principles, to study the response of main phytoplankton groups (algae) to temperature rise due to global warming and different nutrient (phosphate) loading scenarios in the middle and long term. An optimal control problem is formulated to determine optimal combination of restoration strategies, including nutrient loading reduction through inflows remediation within a wetland and in-lake biorestoration techniques. The dynamic optimization problem is formulated within a simultaneous approach.

Keywords: Eutrophication, Optimal control, Dynamic optimization, Climate change, Nutrient loading.

1. Introduction

Lakes and reservoirs over-enrichment by plant nutrient due to urban, agricultural, and industrial development has accelerated the natural process that takes place in water bodies and is referred to as eutrophication. This problem is associated to excessive growth of phytoplankton (algal blooms), especially cyanobacteria, which may produce hepato- and neurotoxins that can severely compromise human and animal health. Cyanobacterial blooms take place during warm months because their maximum growth rates are attained at temperatures above 25 ºC. Temperature is a key parameter in most biological systems and it is suggested that global warming could directly affect several processes in freshwater bodies, e. g. growth, respiration and death rate of phytoplankton, nutrients release from lake sediments (Komatsu et al., 2007). Since cyanobacteria grow on high temperatures, climate change may encourage their growth and they could increasingly dominate the community. Biorestoration in water bodies is a technique based on trophic chain theory and has been one of the most applied to control phytoplankton growth in lakes and reservoirs. The basic idea is to keep a high grazing pressure on the phytoplankton community by the herbivore zooplankton by performing zooplanktivorous fish removal (Sondegaard et al., 2007). However, the middle and long term effect of restoration is not well documented and has had different experimental results depending on the specific water bodies reported in the literature (Søndergaard et al., 1997). A few papers have addressed simulation of the evolution of water bodies under biomanipulation (Sagehashi et al., 2001, Gurkan et al., 2006).

Estrada et al. (2008) have formulated an optimal control problem for algae growth control along a one year horizon, based on a hybrid ecological water model.
In this work, we propose a hybrid eutrophication model within an optimal control problem to determine restoration policies for water quality improvement, taking into account two climate change scenarios for Argentina up to 2013 and two scenarios of nutrient loading, with special attention to the response of cyanobacteria. The hybrid biogeochemical model has been formulated within a simultaneous dynamic optimization approach (Kameswaran and Biegler, 2006) whose solution provides limiting nutrient inflow profiles to the lake, as well as in-lake bio-manipulation profiles.

2. Process description and forecast scenarios

Paso de las Piedras Reservoir is located in the south of Buenos Aires Province (Argentina) at 38° 22´ S and 61° 12´ W. It has 60 km of coastline perimeter, 36 km^2 surface area and a mean depth of 8.2 m. It supplies drinking water to two cities (population around 450,000) and to a petrochemical complex. The reservoir has two tributaries, Sauce Grande River and El Divisorio Stream (6.69 and 0.64 m^3/s mean flowrates, respectively), which run through an important agricultural area. High nutrient concentration in the water body is due to loading from tributaries, as the reservoir is within a protected area (Paso de las Piedras Provincial Park). Its current trophic state is eutrophic and there is an urgent need for restoration. Weekly data on concentrations, temperatures, inflows, outflows, solar radiation etc, collected throughout an entire year in 2004 at four sampling stations in the reservoir, have made possible the development of a first principles hybrid water quality model (Estrada et al., 2009). Inlet variable profiles for the model have been obtained as time correlations for temperature, solar radiation, inflows flowrate and concentrations, based on collected data. In the present work, we study the effect of climate warming for the next five years upon both main phytoplankton and nutrient concentration in Paso de las Piedras reservoir and restoration policies. We have selected two scenarios that represent the lowest and highest temperature forecasts for Argentina, based on greenhouse gases emission models (http://www.cru.uea.ac.uk/~mikeh/research/wwf.argent.pdf). Furthermore, regarding nutrient loading, phosphate concentration profiles in the two tributaries have been reduced to a half and increased twice their current annual mean value (Sauce Grande River 18.53 mg/m^3 and El Divisorio Stream 90.26 mg/m^3) to study its influence on remediation strategies.

3. Water quality modeling

The proposed water quality ecological model is based on first principles, with parameters that have been tuned with collected data from the specific reservoir under study (Estrada et al., 2009). It includes dynamic mass balances for three phytoplankton groups (cyanobacteria, diatoms and chlorophyta) and main nutrients. Input variable profiles are represented as sinusoidal functions in time. Detailed description of the model equations are given in Estrada et al. (2009). The main simplifying assumption is the consideration of horizontally averaged concentrations. This assumption has been introduced in the literature in most mechanistic eutrophication models for lakes and reservoirs (Arhonditsis et al., 2006, Zhao et al., 2008). To decide the inclusion of this assumption, we analyzed collected data at the four existing sampling stations. The small relative differences found (below 10%) justify the assumption of horizontally averaged concentrations in our model. Additional simplifying assumptions in the present model include constant transversal area in the reservoir and constant water density. The

resulting partial differential equations system (considering concentration gradients along the water column) has been discretized into two water layers and main equations for each layer are given below.
Total mass balance

$$\frac{dh_T}{dt} = \frac{1}{A}\sum_{k=1}^{NIN} QIN_k - \frac{1}{A}\sum_{m=1}^{NOUT} QOUT_m + Q_{rain} - Q_{evap} \quad (1)$$

Component mass balances for horizontal layers (U: upper layer, L: lower layer: j: Cyanobacteria, Diatoms, Chlorophyta, Nitrate, Ammonium, Organic Nitrogen, Phosphate, Organic Phosphorus, Biochemical Demand of Oxygen, Dissolved Oxygen)
Upper layer

$$\frac{dC_{Uj}}{dt} = \sum_{k=1}^{NIN} \frac{QIN_{U,k}}{V_U} CIN_{Ujk} - \sum_{m=1}^{NOUT} \frac{QOUT_U}{V_U} C_{Uj} + r_{Uj} - \frac{kdA}{\Delta h_U h_U}(C_{Uj} - C_{Lj}) - \frac{C_{Uj}}{h_U}\frac{dh_U}{dt} \quad (2)$$

Lower layer

$$\frac{dC_{Lj}}{dt} = \sum_{m=1}^{NOUT} \frac{QOUT_L}{V_L} C_{Lj} + r_{Lj} + \frac{kdA}{\Delta h_L h_L}(C_{Lj} - C_{U,j}) - \frac{C_{Lj}}{h_L}\frac{dh_L}{dt} \quad (3)$$

Phytoplankton: Rate equations for phytoplankton groups take into account production and losses due to respiration, natural death, settling and grazing. The growth rate of the three phytoplankton groups is calculated with a multiplicative model as function of solar radiation, water temperature and nutrients availability, as follows:

$$R_{ij,growth} = k_{i,growth} f(T)_{ij} f(I)_{ij} f(N)_{ij} C_{ij} \quad (4)$$

$$f(T)_{ij} = -\frac{(T_j - T_{opt_i})^2}{T^2_{opt_i}} + 1 \quad f(I)_{ij} = \frac{I_{oi}}{I_{opt\,j}} exp\left(1 - \frac{I_{oi}}{I_{opt\,j}}\right) \quad f(N)_{ij} = \frac{C_{PO_{4j}}}{C_{PO_{4j}} + K_{P_i}} \quad (5)\ (6)\ (7)$$

Phytoplankton respiration, natural death, settling rates to the sediment and herbivorous zooplankton grazing rate are calculated as:

$$R_{ij,resp} = k_{j,resp}\,\theta_r^{(T-20)} C_{ij}, \quad i = UL, LL;\ j = cyano, diatom, chlorophyte \quad (8)$$

$$R_{ij,death} = k_{j,death}\,\theta_m^{(T-20)} C_{ij}, \quad i = UL, LL;\ j = cyano, diatom, chlorophyte \quad (9)$$

$$R_{ij,settling} = k_{j,settling}\frac{C_{ij}}{h_i}, \quad i = UL, LL;\ j = cyano, diatom, chlorophyte \quad (10)$$

$$R_{ij,graz} = k_{j,graz}\frac{C_{ij}}{C_{ij} + K_{graz}} C_{i,Zoo}, \quad i = UL, LL;\ j = cyano, diatom, chlorophyte \quad (11)$$

Nutrient cycles (phosphorus and nitrogen): State variables describing phosphorus cycle are phosphate and organic phosphorus. Phosphorus is uptaken by phytoplankton in as phosphate. As phytoplankton biomass is composed of carbon, nitrogen and phosphorus, upon death, phytoplankton increases both phosphate and organic phosphorus and organic nitrogen pool. Organic phosphorus reacts to phosphate, which is then available for phytoplankton generation again. Three state variables describe nitrogen cycle: ammonium, nitrate and organic nitrogen. Phytoplankters uptake both ammonium and nitrate for growth. Organic nitrogen is hydrolyzed at to ammonium, which is in turn oxidized to nitrate. For the lower layer, the model includes a term of ammonium release from sediments. Nitrate can be reduced to molecular nitrogen by denitrification. Detailed rate equations are given in Estrada et al. (2009).

4. Biorestoration and nutrient loading reduction in the middle and long term as an optimal control problem

The DAE system, partially described in the previous section, has been formulated within a dynamic optimization problem whose objective function is the minimization of the weighted sum of the offset between cyanobacteria concentration in the upper water layer and a desired value of 0.10 mg/l and the offset between phosphate concentration and a desired value of 0.02mg/l (both desired values below eutrophication limits) ($\min\left[0.7*\int_0^{tf}\left(\sum_{j=phyto}C_j(t)_{,}-0.10\right)^2 dt+0.3*\int_0^{tf}\left(C_{PO4}(t)-0.02\right)^2 dt\right]$) throughout a time horizon of five years. Not only nutrient loading reduction is required for algae growth control, but in-lake restoration techniques are additionally required due to the complex nutrient recycles previously described (Sondegaard et al., 2007). The proposed model has two time dependent degrees of freedom to represent these strategies: flowrate of tributary that is derived to a nearby wetland for remediation (*Fwetland*) and zooplankton concentration ($c_{U,zoo}$) within the reservoir. The first optimization variable represents handling of nutrient loading to the reservoir through the fraction of inflows that is derived to an existing wetland for nutrient absorption. The wetland size and retention capacity have been included in the model to impose bounds on the maximum derivation flowrate (50% of El Divisorio Stream flowrate) and nutrient concentration reduction (50% global retention of phosphate in the wetland). The second optimization variable ($c_{U,zoo}$) stands for one of the most efficient in-lake restoration strategies, based on the food chain concept: removal of zooplanktivorous fish to increase the possible top-down control of zooplankton on the phytoplankton (Sondegaard et al., 2007). Zooplankton concentration along the time horizon is considered a measure of fish removal rate by including fish biomass data from the specific lake.

The dynamic optimization problem is formulated within a simultaneous dynamic optimization framework by transforming it into a large-scale nonlinear program (NLP) by collocation on finite elements. We have used an Interior point method within program IPOPT (Biegler et al., 2000) with reduced SQP techniques for solving the large-scale NLP. Extensions of this approach are described in Kameswaran and Biegler (2006).

5. Case study: Control of algae growth in Paso de las Piedras under different nutrient scenarios and global warming

As a first step, we have solved the DAE system representing water quality evolution along a time horizon of five years, on a daily basis, for the three temperature profiles shown in Fig. 1, corresponding to current, lower and upper forecasts calculated with green house emission models for Argentina. The DAE model has also been solved for different nutrient (phosphate) scenarios, including current situation and half and double its current inflows, shown in Fig. 2. Simulation results show that there is an increase in cyanobacteria biomass (mgC/l) annual mean of 0.31% in the first year, 8.26% in the second year and 10.4% in the last three years when phosphate loading is increased twice its current profile, while there is no change in the annual mean of cyanobacteria concentration when the model was run with half phosphate inflows concentration (~10^{-5}%) and with the different temperature scenarios (less than 0.01%). In this sense, numerical results show that the change in water temperature predicted for the next five years has much less significant effect on annual cyanobacteria biomass than the increase in phosphate loading.

We have solved the optimal control problem described in Section 4 for the different scenarios. The dynamic optimization problem for two water layers has 23 differential equations and 60 algebraic ones. Within a simultaneous optimization approach, it has been discretized in time with 100 finite elements and two collocation points rendering a nonlinear problem with 19523 variables, which has been solved with program IPOPT (Biegler et al., 2002) that implements an Interior Point method with reduced Successive Quadratic Programming. The optimization variables are the fraction of inlet stream that is derived to the wetland and the concentration of zooplankton in the lake to control phytoplankton growth, along the five years time horizon. Numerical results show that tributary deviation through a wetland for nutrient loading reduction is required throughout the entire time horizon, at its maximum allowed flowrate (50% of the inflows, with 50% nutrient retention) when the second term (minimization of phosphate offset) is included in the objective function. Figures 3 and 4 show cyanobacteria concentration profiles before and after biomanipulation, as well as zooplankton concentration profiles for the model run in normal conditions and for double phosphate inflows concentration. Nutrient loading reduction requires even longer time horizons to show noticeable effects on concentration (Fig. 5) due to internal nutrient recycle whereas increases in phosphorus loading has an immediately effect on cyanobacteria concentration.

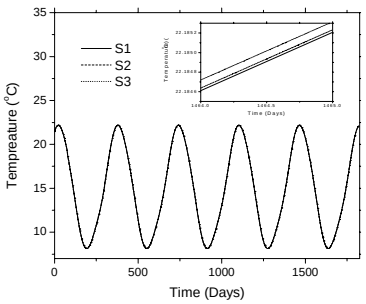


Figure 1. Temperature scenarios. S1: current temperature, S2: lower predicted temperature, S3: upper predicted temp.

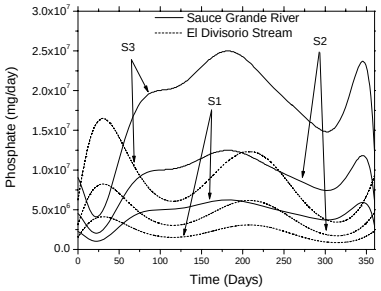


Figure 2. Phosphate inflows scenarios. S1: current PO_4 inflows, S2: half PO_4 inflows conc., S3: twice PO_4 inflows concentration

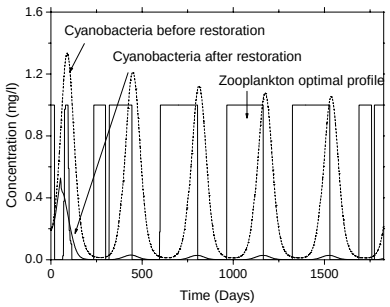


Figure 3. Comparison between cyanobacteria concentration profiles before and after restoration and optimal profile for zooplankton conc. with a current conditions of PO_4 inflows.

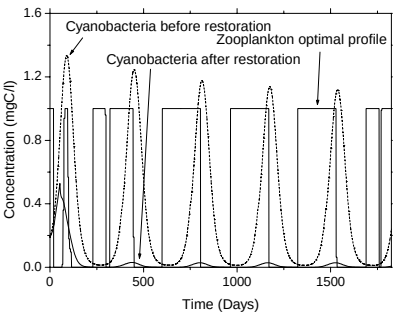


Figure 4. Comparison between cyanobacteria concentration profiles before and after restoration and optimal profile for zooplankton conc. with a twice PO_4 inflows concentration.

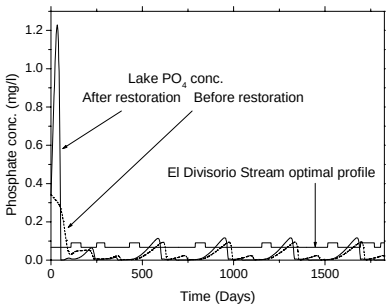


Figure 5. Comparison between phosphate concentration profiles in Paso de las Piedras reservoir before and after restoration and optimal PO_4 conc. profile from El Divisorio Stream.

6. Conclusions

Advanced dynamic optimization techniques integrated to first principles hybrid models for water bodies have given useful insights on the determination of biorestoration policies. The formulation of an optimal control problem for a long term horizon has provided quantitative information on the combined application of nutrient loading reduction strategies by partial remediation of inflows within an existing wetland and the biomanipulation technique of zooplanktivorous fish removal for top-down control of zooplankton on phytoplankton. Numerical results have also shown that global warming along the next five years has no noticeable effect on algal growth. On the other hand, increases in nutrient loading, mainly due to nonpoint sources (agricultural activities) will certainly affect lake quality if no restoration action is taken.

References

G. B. Arhonditsis, C. A. Stow, L. J. Steinberg, M. A. Kenny, R. C. Lathrop, S. J. McBride and K. H. Reckhow, 2006, Exploring ecological patterns with estructural equation modeling and Bayesian analysis, Ecol. Model., 192, 385-409.

L. T. Biegler, Cervantes, A. and Waechter, A., 2002, Advances in simultaneous strategies for dynamic process optimization, Chem. Eng. Sci., 57, 575-593.

V. Estrada, E. R. Parodi and M. S. Diaz, 2008, Addresing the control problem of algae growth in water reservoirs with advanced dynamics optimization approaches, FOCAPO 2008, USA.

V. Estrada, E. R. Parodi and M. S. Diaz, 2009, Determination of biogeochemical parameters in eutrophication models with simultaneous dynamic optimization approaches, Comp. & Chem. Eng., in press.

Z. Gurkan, J. Zhang, and S. E. Jørgensen, 2006. Development of a structurally dynamic model for forecasting the effects of restoration of Lake Fure, Denmark, Ecol. Model., 197, 89-102.

S. Kameswaran and L. T. Biegler, Simultaneous dynamic optimization strategies: Recent advances and challenges, 2006, Comp. & Chem. Eng., 30, 1560-1575.

E. Komatsu, T. Fukushima and H. Harasawa, 2007, a modelin approach to forecast the effect of long-term climate change on lake water quality, Ecol. Model., 209, 351-366.

M. Sagehashi, A. Sakoda and M. Suzuki, M., 2001, A mathematical model of a shallow and eutrophic lake (The Keszthel Basin, Lake Balaton) and simulation of restorative manipulations, Wat. Res., 35, 1675-1686.

M. Søndergaard, E. Jeppesen, T. L. Lauridsen, C. Skov, E. H. van Nes, R. Roijackers, E. Lammens, and R. Portielje, 2007, Lake restoration: successes, failures and long-term effects. J. Appl. Ecol., 44, 1095–1105.

J. Zhao, M. Ramin, V. Cheng and G. B. Arhonditsis, 2008, Plankton community patterns across a trophic gradient: The role of zooplankton functional groups, Ecol. Model. 213, 417-436.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Design of Stable Large-Scale Metabolic Networks

Jimena Di Maggio, Aníbal M. Blanco, J. Alberto Bandoni, M. Soledad Diaz

Planta Piloto de Ingeniería Química (PLAPIQUI),
Universidad Nacional del Sur-CONICET,
Camino La Carrindanga Km 7, Bahía Blanca 8000, Argentina

Abstract

In this work we propose an eigenvalue optimization approach to ensure steady state stability of the Embden-Meyerhof-Parnas pathway, the pentose-phosphate pathway and the phosphotransferase system of *Escherichia coli*. The model consists of eighteen differential equations that represent dynamic mass balances for extracellular glucose and intracellular metabolites and thirty kinetic rate expressions. The nonlinear optimization problem including stability constraints has been solved with reduced space Successive Quadratic Programming techniques within program IPOPT (Waechter and Biegler et al., 2006). Numerical results provide useful insights on the stability properties of the studied kinetic model.

Keywords: Metabolic networks, Stability, Eigenvalue optimization.

1. Introduction

Metabolic networks design can be formulated as an optimization problem aimed at optimizing a given objective, for example the production of a certain metabolite, subject to mass balance equations that represent the network. Kinetic models allow the analysis of stability of the predicted states, which is of fundamental importance because biological systems may exhibit monotonic stable states, bistable switching threshold phenomena, oscillations and chaotic behavior. Due to nonlinear kinetics of the biochemical reactions and their coupling through common metabolites, biological systems may undergo drastic changes in their qualitative behavior when a variation on the enzyme level occurs. If no stability constraints are included in the formulation, the optimal operating point might be unstable, making the metabolic network vulnerable to external disturbances. In other words, in spite of the presence of modest disturbances an unstable network will reach physiological constrains and collapse. Several authors have addressed the analysis if biological systems of small to moderate size (Hatzimanikatis and Bailey, 1997, Haddad and Chellaboina, 2005). The design-for-stability problem, an important sub problem of the general design-for-operability problem, has also motivated many contributions from the process systems engineering community. Different strategies have been proposed to include stability considerations within the design problem (Chang and Sahinidis, 2004; Blanco et al., 2004).

In this work we propose an eigenvalue optimization approach (Blanco and Bandoni, 2007) to ensure steady state stability of the glycolysis, the pentose-phosphate pathway and the phosphotransferase system of *Escherichia coli* K-12 W3110 (Chassagnole et al., 2002). The nonlinear optimization problem, corresponding to steady state equations and stability constraints, has been solved with reduced space Successive Quadratic Programming techniques within program IPOPT (Waechter and Biegler, 2006). Optimization results provide an improved metabolic network for the maximization of

serine production in *Escherichia coli*, within the steady state stable region, based on a detailed kinetic model.

2. Optimization model description

In this work, we have studied an adaptation of the dynamic model for the Embden-Meyerhof-Parnas pathway, the pentose phosphate pathway and the phosphotransferase system, as shown in Fig. 1, of *Escherichia coli* K-12 W3110 (Chassagnole et al., 2002). It comprises eighteen differential equations that represent dynamic mass balances of extracellular glucose and intracellular metabolites, thirty kinetic rate expressions and it involves one hundred and sixteen parameters. Most influential parameters have been determined through previous work on global sensitivity analysis of the proposed model (Di Maggio et al., 2008a,b).

Equations (1) to (6) correspond to main mass balances on metabolites involved in the metabolic network shown in Fig. 1. A detailed description of the remaining twelve balances is given in Chassagnole et al. (2002).

$$\frac{dC_{glc}^{ext}}{dt} = D(C_{glc}^{a\lim} - C_{glc}^{ext}) + f^{pulso} - \frac{C_X r_{PTS}}{\rho_X} \tag{1}$$

$$\frac{dC_{g6p}}{dt} = r_{PTS} - r_{PGI} - r_{G6PDH} - r_{PGM} - \mu C_{g6p} \tag{2}$$

$$\frac{dC_{f6p}}{dt} = r_{PGI} - r_{PFK} + r_{TKb} + r_{TA} - 2r_{MurSynth} - \mu C_{f6p} \tag{3}$$

$$\frac{dC_{gap}}{dt} = r_{ALDO} + r_{TIS} - r_{GAPDH} + r_{TKa} + r_{TKb} - r_{TA} + r_{TrpSynth} - \mu C_{gap} \tag{4}$$

$$\frac{dC_{pep}}{dt} = r_{ENO} - r_{PK} - r_{PTS} - r_{PEPCxylase} - r_{DAHPS} - \mu C_{pep} \tag{5}$$

$$\frac{dC_{pyr}}{dt} = r_{PK} + r_{PTS} - r_{PDH} + r_{MetSynth} + r_{TrpSynth} - \mu C_{pyr} \tag{6}$$

Equations 7 to 13 show kinetic expressions for phosphotransferase system, glucose-6-phosphate isomerase, 6-phosphogluconate dehydrogenase, phosphoglycerate kinase and phosphoglycerate mutase, the Serine synthesis pathway and the Chorismate and Mureine synthesis pathway, respectively. The remaining rate equations involved in the present metabolic network can be found in Chassagnole et al. (2002).

$$r_{PTS} = \frac{r_{PTS}^{\max} c_{glc}^{extracelul\ \ lar} \frac{c_{pep}}{c_{pyr}}}{\left(K_{PTS,a1} + K_{PTS,a2}\frac{c_{pep}}{c_{pyr}} + K_{PTS,a3} c_{glc}^{extracelul\ \ lar} + c_{glc}^{extracelul\ \ lar}\frac{c_{pep}}{c_{pyr}}\right)\left(1 + \frac{c_{g6p}^{N_{PTS,g6p}}}{K_{PTS,g6p}}\right)} \tag{7}$$

$$r_{PGI} = \frac{r_{PGI}^{\max}\left(C_{g6p} - \frac{C_{f6p}}{K_{PGI,eq}}\right)}{K_{PGI,g6p}\left(1 + \frac{C_{f6p}}{K_{PGI,f6p}\left(1 + \frac{C_{6pg}}{K_{PGI,f6p,6pginh}}\right)} + \frac{C_{6pg}}{K_{PGI,g6p,6pginh}}\right) + C_{g6p}} \tag{8}$$

$$r_{PGDH} = \frac{r_{PGDH}^{\max} c_{6pg} c_{nadp}}{\left(c_{6pg} + K_{PGDH,6pg}\right)\left(c_{nadp} + K_{PGDH,nadp}\left(1 + \frac{c_{nadph}}{K_{PGDH,nadph,inh}}\right)\left(1 + \frac{c_{ATP}}{K_{PGDH,ATP,g6pinh}}\right)\right)} \tag{9}$$

$$r_{PGK} = \frac{r_{PGK}^{\max}\left(C_{adp}C_{pgp} - \frac{C_{atp}C_{3pg}}{K_{PGK,eq}}\right)}{\left(K_{PGK,adp}\left(1 + \frac{C_{atp}}{K_{PGK,atp}}\right) + C_{adp}\right)\left(K_{PGK,pgp}\left(1 + \frac{C_{3pg}}{K_{PGK,3pg}}\right) + C_{pgp}\right)} \tag{10}$$

$$r_{PGluMu} = \frac{r_{PGluMu}^{\max}\left(C_{3\,pg} - \frac{C_{2\,pg}}{K_{PGluMu,eq}}\right)}{K_{PGluMu,3\,pg}\left(1 + \frac{C_{2\,pg}}{K_{PGluMu,2\,pg}}\right) + C_{3\,pg}} \tag{11}$$

$$r_{SerSynth} = \frac{r_{SerSynth}^{\max} C_{3\,pg}}{K_{SerSynth,3\,pg} + C_{3\,pg}} \tag{12}$$

$$r_{Synth1} = \frac{r_{Synth1}^{\max} C_{pep}}{K_{Synth1,pep} + C_{pep}} \tag{13}$$

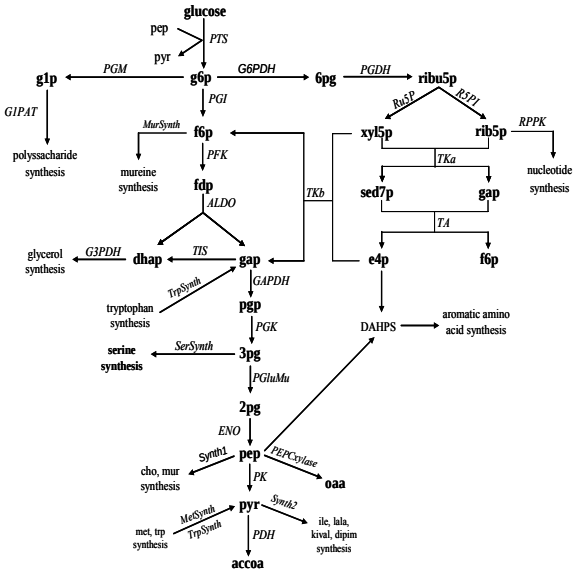


Figure 1. Metabolic network of glycolysis and pentose-phosphate pathway in *Escherichia coli*.

The maximum reaction rates, r_i^{max}, are associated to the corresponding enzyme concentration, so they could be tuned to maximize or minimize the production of a given metabolite. We have formulated an optimization problem in which the maximum reaction rates for certain pathways are the degrees of freedom. As changes in these parameters imply that the levels of enzymes will be modified, the total proteins concentration will vary and physiological changes that are not taken account in the dynamic model could be appear. To account for processes such as redistribution of limited mRNA contents and homeostasis (Mauch et al., 2001, Nikolaev et al., 2005), we have included equations (14) and (15). Eqn. (14) imposes that an increase in certain enzyme levels is compensated by a decrease in the remaining ones. Eqns. (15) ensure that gene expression rates of non-modulated enzymes be maintained at ratios equal to the ones at the reference steady state.

$$\frac{1}{M}\sum_{j=1}^{M}\frac{r_j^{\max}}{r_j^{\max,0}} = 1 \tag{14}$$

$$\frac{r_{j_1'}^{\max}}{r_{j_1'}^{\max,0}} = \ldots = \frac{r_{j_K'}^{\max}}{r_{j_K'}^{\max,0}} = \gamma \tag{15}$$

where $r_j^{max,0}$ is the maximum reaction rate at the reference state, M is the number of enzymes in the metabolic network, $j_1' \ldots j_K'$ are the indices of non-modulated enzymes,

$K = M-L$ and L is the number of modulated enzymes (corresponding to maximum reaction rates being optimization variables).
Equations (14) and (15) can be re-written as (16) and (17), respectively, to be included in the optimization problem formulation.

$$\frac{r_{j_1}^{\max}}{r_{j_1}^{\max,0}} + \ldots + \frac{r_{j_L}^{\max}}{r_{j_L}^{\max,0}} + K\gamma = M \tag{16}$$

$$r_{j_S}^{\max} = \gamma r_{j_S}^{\max,0} \quad , S = 1,\ldots,K \tag{17}$$

3. Optimization under stability constraints

In order to assess asymptotic stability of dynamic systems, eigenvalue analysis is usually performed. For an asymptotically stable equilibrium point, the eigenvalues of the dynamic system Jacobian matrix lie on the left half of the complex plane. In an eigenvalue optimization problem (Blanco and Bandoni, 2007), the real parts of the eigenvalues of the Jacobian matrix of the dynamic system under study are forced to be strictly negative, ensuring that way asymptotic stability of the resulting equilibrium point. The eigenvalue optimization problem can be stated as follows:

$$\begin{aligned} & \min_{y} \Phi(y) \\ \text{s.t} \quad & A^T(y)P + PA(y) + I = 0 \\ & \det(P_i^{-1}) \geq \xi \qquad i=1,\ldots,n \\ & \xi > 0 \\ & h(y) = 0 \\ & g(y) \leq 0 \\ & y \in Y \end{aligned} \tag{18}$$

where $\boldsymbol{y}$ represents the optimization variables which comprises both, design and operating variables. $\Phi(y)$ is the objective function, $h(y)$ is the set of equality constraints and $g(y)$ is the set of inequality constraints. $\boldsymbol{A(y)}$ is the Jacobian matrix of the dynamic system and $\boldsymbol{P}$ is a real symmetric matrix defined in (18) through the Lyapunov equation. Det($\boldsymbol{P}_i^{-1}$) stands for determinants of the principal minors of the inverse of $\boldsymbol{P}$, $\boldsymbol{I}$ is the identity matrix and ξ a user defined positive constant. For details on the derivation and solution strategy of (18) see Blanco and Bandoni, (2007).

4. Discussion of results

In this study, and based on previous work on global sensitivity analysis on the dynamic metabolic network to main parameters (Di Maggio et al., 2008a,b), we have selected four maximum reaction rates as design variables, corresponding to serine synthesis ($\boldsymbol{r_{SerSynth}^{max}}$), glucose-6-phosphate isomerase ($\boldsymbol{r_{PGI}^{max}}$), phosphotransferase system ($\boldsymbol{r_{PTS}^{max}}$) and Chorismate and Mureine synthesis ($\boldsymbol{r_{Synth1}^{max}}$), respectively. The design problem to maximize serine production ($\Phi(y)=r_{SerSynth}$), for the steady state kinetic model of the Embden-Meyerhof-Parnas pathway, the pentose phosphate pathway and the phosphotransferase system comprises mass balances for eighteen metabolites and thirty rate equations associated to thirty enzyme levels (partially shown as Eqns. (7) to (13), as well as twenty seven equations (K+1) standing for Eqns. (16) and (17).
When formulating the design problem under stability constraints, Eqn. (18), additional equality constraints (eighteen) and inequalities (eighteen) are included in the optimization problem formulation, standing for Lyapunov's equation and nonnegativity on the determinants of the principal minors of P, to ensure its positive definiteness,

respectively. Nonlinear optimization models have been implemented in a Fortran 90 environment and they have been solved with reduced space Successive Quadratic Programming techniques within program IPOPT (Waechter and Biegler, 2006). The design problem has been first solved without stability constraints (Case 1) and it has been extended to include the proposed equations to ensure a stable optimal network design (Case 2).

Table 1. Main optimization variable values.

Optimization variables	Nominal values	Case 1	Case 2
C_{glcext} (mM)	1,7222	0,9946	0,2599
C_{g6p} (mM)	3,4800	5,1859	4,2877
C_{f6p} (mM)	0,6000	0,8941	0,7393
C_{fdp} (mM)	0.2720	0,2211	0,3280
C_{gap} (mM)	0,2180	0,1818	0,2349
C_{dhap} (mM)	0,1670	0,1485	0,1838
C_{pgp} (mM)	0,0080	0,0035	0,0059
C_{3pg} (mM)	2,1300	0,8866	1,5158
C_{2pg} (mM)	0,3990	0,1656	0,2837
C_{pep} (mM)	2,6700	1,0491	1,7945
C_{pyr} (mM)	2,6700	3,1879	2,9868
C_{6pg} (mM)	0,8080	0,9927	0,8572
C_{ribu5p} (mM)	0,1110	0,1475	0,1385
C_{xyl5p} (mM)	0,1380	0,1713	0,1637
C_{sed7p} (mM)	0,2760	0,5147	0,3624
C_{rib5p} (mM)	0,3980	0,5089	0,4823
C_{e4p} (mM)	0,0980	0,1005	0,1110
C_{g1p} (mM)	0,6530	0,9661	0,7910
r_{PGI}^{max} (mM/sec)	495,870	442,316	495,842
$r_{SerSynth}^{max}$ (mM/sec)	0,0203	0,2000	0,0985
r_{Synth1}^{max} (mM/sec)	0,0148	0.0000	0.0000
Serine Production (mM/sec)	0,0138	0,0939	0,0594

Case 1*: Optimization for maximization of Serine production*

Numerical results show that serine production could be increased from 0,01381 mM/sec in a reference steady state (experimental) to 0,09396 mM/sec when the maximum serine synthesis reaction rate ($r_{SerSynth}^{max}$) is at its upper bound (=10* $r_{SerSynth}^{max,0}$). Main variable values are shown in the third column of Table 1. The eigenvalues of matrix *A(y)* show that the metabolic network is stable, being the real part of the largest eigenvalue -3.7E-4. While stable, the system is close to critical stability and small changes in some parameters could easily lead to an unstable equilibrium. Taking into account the difficulty to modulate enzyme levels with precision, this implies that a relative deviation on the enzyme levels from their desired value may produce unstable behavior in the system. Therefore, the inclusion of stability constraints in the design problem becomes necessary.

Case 2*: Optimization for maximization of Serine production under stability constraints*

In order to modify the spectrum of matrix ***A(y)*** different trials with parameter ξ were performed. For ξ= 1E-4, the largest eigenvalue in real part is -1.5E-3. In this case, maximum serine production can be increased 330% from its reference steady state value (0,01381 mM/sec) to 0,05936 mM/sec. It can be noted that some robustness of the

network regarding dynamic stability was achieved at the expense of a worsening in the adopted objective function. Furthermore, serine synthesis reaction rate ($r_{SerSynth}^{max}$) is not at its upper bound, as constraints on the determinants of the inverse of P become active. The fourth column in Table 1 shows main variable values in this case. In both Case 1 and 2, optimal values for the maximization of Serine production implies $r_{Synth1}^{max} = 0$. This corresponds to deletion of Chorismate and Mureine synthesis pathway.

5. Conclusions

The proposed approach in the formulation of a design problem for a large-scale kinetic model for a metabolic network including stability constraints, allows the determination of an improved network for serine production by ensuring its stability at the design level.

6. Acknowledgment

The authors gratefully acknowledge financial support from the National Research Council, Universidad Nacional del Sur and ANPCYT, Argentina.

References

A.M. Blanco and J.A. Bandoni, 2007, Eigenvalue Optimization Based Formulations for Dynamics and Control Problems, Chemical Engineering Process, 46, 1192-1199

A.M. Blanco, J.A. Bandoni, L. Biegler, 2004, Re-design of the Tennessee Eastman Challenge Process: An Eigenvalue Optimization Approach, Proceedings FOCAPD 2004, Princeton, 517-520

Y. Chang, N. Sahinidis, 2004, Optimization of metabolic pathways under stability considerations, Computers and Chemical Engineering, 29, 467-479

C. Chassagnole, N. Noisommit-Rizzi, J. Schmid, K. Mauch, M. Reuss, 2002, Dynamic Modeling of the Central Carbon Metabolism of *Escherichia coli*, Biotechnology and Bioengineering, 79, 1, 53-73

J. Di Maggio, J.C. Diaz Ricci, M.S. Diaz, 2008a, Global Sensitivity Analysis: Estimation of Sensitivity Indices in Metabolic Network Dynamic Models, submitted to Metabolic Engineering.

J. Di Maggio, J.C. Diaz Ricci, M.S. Diaz, 2008b, Ranking of most influential kinetic parameters in metabolic networks through Global Sensitivity Analysis, Engineering Conferences International-Metabolic Engineering VII, September 14-19, Puerto Vallarta, Mexico

W. Haddad and V. Chellaboina, 2005, Stability and dissipativity theory for nonnegative dynamical systems: a unified analysis framework for biological and physiological systems, Nonlinear Analysis-Real World Applications, 6, 33-65

V. Hatzimanikatis and J. Bailey, 1997, Studies on glycolysis.1.Multiple steady states in bacterial glycolysis, Chemical Engineering Science, 52, 2579-2588

K. Mauch, S. Buziol, J. Schmid, M. Reuss, 2001, Computer Aidded Design of Metabolic Networks, Chemical Process Control-6 Conference, Tucson, Arizona

E. Nikolaev, P. Pharkya, C. Maranas, A Armaou, 2005, Optimal selection of enzyme levels using large-scale kinetic models, Proceedings of 16th I.F.A.C. World Congress, Prague, Czech Republic

A. Waechter, L. Biegler, 2006, On the Implementation of an Interior Point Filter Line Search Algorithm for Large-Scale Nonlinear Programming, Mathematical Programming, 106, 1, 25-57

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Multi-resolution Multi-scale Computational Approach for Characterization and Analysis of Nanostructured Surfaces

Rajib Mukherjee[a], Ahmet Palazoglu[b], Jose A Romagnoli[a]

[a]*Cain Dept. of ChemicalEngineering, Louisiana State University,Baton Rouge 70803, USA*
[b]*Dept. of Chemical Engineering and Materia Science, University of California, Davis, USA*

Abstract
Structures and property of surfaces are very important in different chemical, physical and biological processes. Understanding the surface characteristics in the microscopic level is essential in order to relate the surface characteristics to the performance of the product. Relation of product performance with surface characteristics helps to improve the product performance through optimizing the manufacturing process. Spatial distribution of surface features which defines the surface characteristics can be captured by the multi-resolution capabilities of wavelet transforms (WT) that can provide not only frequency localization but also spatial localization of feature signatures. A multi-scale molecular simulation can help to investigate the physical and chemical mechanism in the surface. Together with the multi-resolution surface feature analysis, the multi-scale molecular simulation will give a better understanding of the surface phenomena and its relation with the performance matrices. In this paper we discuss the application of this approach for surface characterization of Rh(111) in the adsorption desorption of CO. The adsorption on the surface depends on its crystal lattice structures and the presence of defects. In the atomic level a first principle density functional theory (DFT) calculation is used to find the adsorption energy. In the mesoscopic level a kinetic Monte Carlo (KMC) model of the surface is used to simulate the temperature programmed desorption (TPD) from the surface. The on-top adsorption energy increases with surface defects in the form of vacancies which shifts the peak maximum of TPD to a higher temperature. To characterize the surface, fractal dimension of the crystal surface is found using wavelet transformation. The fractal dimension of the surface increases with presence of vacancies.

Keywords: wavelet transform; density functional theory (DFT); kinetic Monte Carlo (KMC); temperature programmed desorption (TPD)

1. Introduction
The performance of various products like heterogeneous catalysts, solar energy conversion panels, semiconductor devices etc. are controlled by the characteristics of the molecular interfaces and nanostructure of their surfaces. The surface structure of these applications, created by molecular deposition on the surface due to surface kinetics and surface diffusion, governs the function of the material. To create an application of desired features, we first have to understand the surface characteristics of these applications in the microscopic level find the relation of product performance with

surface characteristics, thereby improve the product performance through optimizing the manufacturing process towards novel tailored nanostructures. Thus, the key to control fabrication of applications whose performance matrix is governed by its surface feature is to characterize the surface and use those characteristic features as control parameters. In this way we can optimize product performance as well as have a quality control.

The performance of heterogeneous catalysts is controlled by the characteristics of the molecular interfaces and nanostructure of their crystal surfaces. The characterization of these nanostructures often involves surface imaging using Scanning Electron Microscopy (SEM), high resolution electron microscopy (HREM) and different Scanning Probe Microscopy (SPM). A high resolution electron microscopy (HREM) image of the catalyst surface used by Yakaman et al. to obtain the crystal lattice structure of the active metal, revealing the defects on the crystal surface of the crystal lattice structure. Crystal formed in non equilibrium condition will have defects. Such surface defect affects the surface property. The effects of surface defects like steps and kinks in adsorption and subsequent dissociation have been studied for some time In this paper, we have studied the effect of surface vacancies in the adsorption-desorption property. While microscopy images can provide qualitative information about the features of a surface, we have used wavelet transformation on simulated surface with and without vacancies for the quantitative description of the surface morphology. A first principle DFT model of similar surface is used to relate the surface characteristics with its adsorption-desorption property. In our study the adsorption energies of CO in Rh(111) for low surface coverage (p(2×2) for 0.25 monolayer (ML)) were obtained using Car-Parrinello Mdecular Dynamics (CPMD) program. These adsorption energies are then used to find the desorption kinetics by the kinetic Monte Carlo (KMC).

The TPD of CO for a regular Rh(111) surface at low surface coverage of CO was compared with the experimental results obtained by Dubios and Somorjai. A surface with vacancies will have higher surface energies. The effects of the presence of vacancies on adsorption energies and the corresponding effects on TPD were simulated. The changes in fractal dimension with the presence of vacancies are noted. A regular surface will have a fractal dimension of two while a surface with vacancies will have fractal dimension between two and three.

2. Computational details

2.1. Surface Fractal dimension

Fractal dimension can be obtained from the variance of wavelet coefficients. The roughness coefficient or the Hurst exponent, H, is obtained from the slope (s) of the plot of wavelet variance at different scale index (Mukherjee et al.). The slope is related to the Hurst exponent as given in Equation 1.

$$H = \frac{s-2}{2} \tag{1}$$

For a surface the roughness or Hurst exponent is related to the fractal dimension, D, as $D = 3 - H$.

2.2. DFT estimation of adsorption energies

The surface adsorption energy E_{ad} at a specific temperature and pressure is given by the difference in the internal energy of the surface with adsorbate in it $E_{CO/Rh-slab}$ and

sum of the internal energy of the surface $E_{Rh-slab}$ and adsorbate E_{CO} as shown in Equation 2 (Gajdos et al.).

$$E_{ad} = E_{Rh-slab} + E_{CO} - E_{CO/Rh-slab} \quad (2)$$

A DFT estimation of the adsorption energies of CO in Rh(111) is calculated for both on-top and bridge sites. 0.25 ML surface coverage is used. Lateral interactions can be considered to be small as CO adsorbate-adsorbate lateral interactions at surface coverage lower than 0.33 ML is not significant (Gajdos et al.). The wave function was expanded on a basis set of plane wave with cut-off of 70 Ry. We have used a norm conserving pseudo potential of Trouiller-Martins with Becke-Lee-Yang-Parr (BLYP) functional and non linear core correction (NLCC). A geometry optimization is done. The adsorption is supposed to occur in ultra high vacuum (UHV), hence there is no dissociation of CO molecules upon adsorption as found experimentally and theoretically by Dabois et al.

2.3. Kinetic Monte Carlo (KMC) simulation of the surface

The KMC is a stochastic algorithm that will propagate the system from state to state. The steps used in simulation are constructing the surface and adsorbate, determining the kinetics of desorption process taking the lateral interaction into consideration, followed by the simulation of the elementary processes like diffusion and desorption to determine the output. The intrinsic binding energies, activation energies are derived from DFT results and lateral interaction energies are obtained from the values estimated by Oh et al.. The structure of the catalyst is assumed to be rigid and represented by a lattice with each point corresponding to a surface site. Every lattice site is assigned with the type of the site and whether or not adsorbate CO is associated with it.

2.3.1. Surface Kinetics

We have used the transition state theory to get the kinetic parameters from the activation barriers and the rate constant of the surface features like desorption are described by the Arrhenius equation. The overall rate for a particular reaction r_i is given by Equation 3.

$$r_i = A \exp\left[\frac{-(E_{i-des} + Q)}{k_B T}\right] \quad (3)$$

The energetic barrier to the desorption process are computed according to the local environment around each site. E_{i-des} is the energy of desorption without lateral interaction and Q is the interaction energy due to lateral interaction. A is the pre-exponential factor and k_B and T are the Boltzmann constant and absolute temperature respectively. The total rate (r_{tot}) of all surface processes is calculated by summing the rates for all elementary reactions ($\sum_i r_i$) in ascending order. A random number (RN) is drawn to determine the event that will occur in the next time step. The forward time step Δt is then calculated by Equation 4.

$$\Delta t = \frac{-\ln(RN)}{r_{tot}} \quad (4)$$

2.3.2. Surface Diffusion

The diffusion rate for the systems studied was assumed to be fast compared to the other surface process like desorption. Therefore the surface is equilibrated after every time step using a Metropolis algorithm. First an occupied site on the surface is chosen randomly and the binding energy of the initial site Q_i is calculated. Then a vacant location in the vicinity where CO can move is chosen randomly and the CO is moved there. The binding energy of the final site Q_f is calculated. The difference of the two adsorption energy is given by $\Delta Q = Q_f - Q_i$. The probability of acceptance of the move is given by Equation 5.

$$P_{i-f} = \min\left[\exp\left(-\frac{\Delta Q}{kT}\right), 1\right] \qquad (5)$$

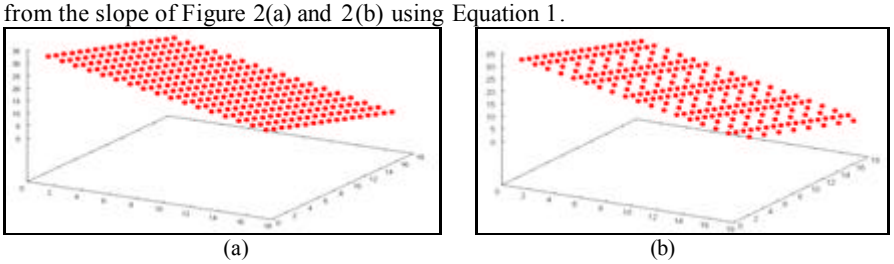
At the initial surface temperature, the surface is populated by placing a given coverage (p(2×2) for 0.25 ML) of CO. The CO molecules are placed on the on-top sites due to the minimum adsorption energy of the sites. The surface diffusion algorithm is used to equilibrate this initial coverage by performing the diffusion step iteratively. A desorption process is then performed according to its local environment and the time step required in this process is calculated. The process of surface diffusion followed by desorption is continued till the total time at a given temperature exceeds one second. After reaching one second, the surface temperature is increased and the process is repeated. In this way, the numbers of molecules per second desorbing at different temperatures are calculated.

3. Results and discussion

The adsorption of CO on Rh(111) has been studied. The on-top adsorption energy increases in surface with vacancies. The temperature programmed desorption (TPD) of CO from regular Rh(111) surface and that with vacancies is modeled using kinetic Monte Carlo (KMC). The peak maximum of TPD from a regular Rh(111) surface is close to that found experimentally by Dubois et al. for low surface coverage of CO. The shift in the peak maximum of TPD has been observed for a rough surface due to the change in activation energies for desorption with the presence of vacancies.

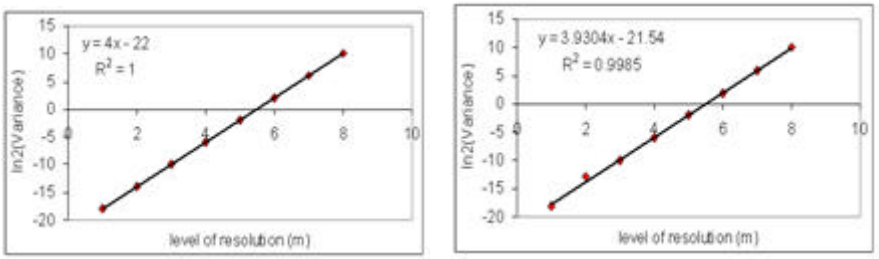
3.1. Surface Fractal Dimension Estimation

Vacancies are caused by missing atoms which makes the surface rough. A fractal dimension of the surface is created for surface characterization by simulating the surface in a periodic manner. This will keep the fractal dimension same as it is based on the self affine nature of the surface. Simulated surfaces without and with defect are shown Figure 1(a) and 1(b) respectively. The corresponding fractal dimensions is calculated from the slope of Figure 2(a) and 2(b) using Equation 1.



(a) (b)

Figure 1 : Simulated (111) lattice plane (a) without/ (b) with defects.



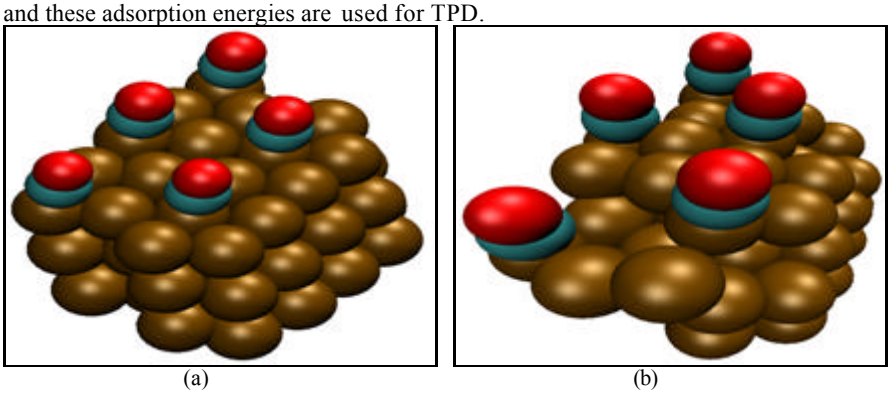
(a) (b)

Figure 2 : Wavelet variance vs. different level of resolution for (a) regular (111) lattice plane and (b) plane with vacancies.

The fractal dimension for a regular surface is 2. The fractal dimension for the surface with vacancy is calculated to be 2.035. This shows that the surface energy increased with its fractal dimension.

3.2. Surface Adsorption Energies

The effect of higher surface energies of rough surface is seen in the adsorption energy. A three layer of Rh(111) surface atoms with CO adsorbing on it are modeled. The adsorption energies for on-top sites on regular surface and surface with vacancies are calculated. Figure 3 shows adsorption on on-top sites. For regular Rh(111) the adsorption energies are found to be 32.8 kcal/mole and that on Rh(111) surface with vacancies are 35.1 kcal/mole. This shows higher adsorption energy when compared with surface without roughness. Adsorption energies on bridge sites are also calculated and these adsorption energies are used for TPD.



(a) (b)

Figure 3: Adsorption on on-top sites in (a) regular Rh(111) and (b) with vacancies.

3.3. Temperature Programmed Desorption (TPD)

The desorption kinetics of CO from the Rh(111) surface is studied under ultra high vacuum (UHV) condition. The low coverage adsorption energies are assumed to be the activation energies for desorption. The CO molecules are arranged in p(2×2) in a regular R(111) surface and the surface temperature is raised at a rate of 25 K/sec. The pre-exponential factor, A, is taken as 5×10^{13} sec^{-1}. The TPD of CO from Rh(111) is shown in Figure 4. The peak maximum is observed at 525 K. This is in agreement with the experimental results obtained by Dubois et al. for low surface coverage indicating that desorption occurs from a single on-top site.

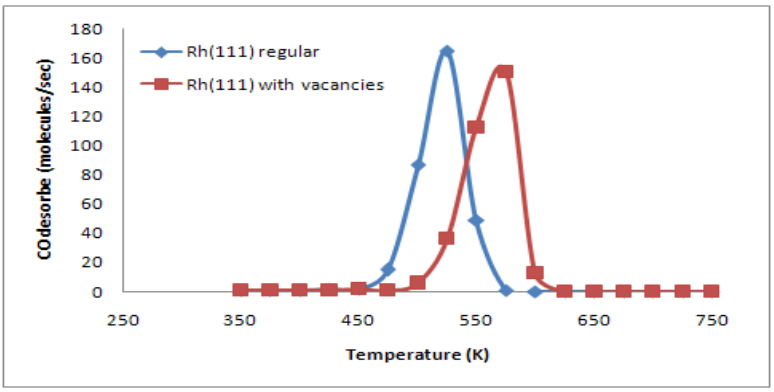


Figure 4: Temperature-programmed Desorption of CO from Rh(111).

The TPD from a surface with vacancies shows a shift in peak maximum from 525K to 575K. The shift in peak maximum to a higher temperature is due to high adsorption energies for a rough surface. We are comparing results from different model surfaces with experimentally created surfaces of similar characteristics.

4. Conclusions

A surface with vacancies has higher energies than regular surfaces. This increases the adsorption activities of the surface. A fractal dimension of the surface estimated in this paper using wavelet transform can be used to characterize the surface in atomic scale. With presence of vacancies, there is an increase in fractal dimension as well as adsorption energies. This increase in adsorption energy shifts the peak maximum from 525 to 575 K in TPD of CO. This shows changes in surface properties with the presence of roughness in form of vacancies. Surface characterization of heterogeneous catalyst in the atomic scale is thus essential for device performance analysis. Similar studies can be made with different lattice surface planes of different active metals as well as defects in terms of dislocations and grain boundaries to see their effects in catalytic activities. Catalytic activities of the surface like dissociation and product formation can also be studied in surfaces with vacancies, dislocations, grain boundaries etc..

References

M.J. Yakaman, G. Diaz and A. Gomez, 1995, Electron Microscopy of Catalysis; the Present, the Future and the Hopes, Catal. Today, 23, pp. 161-199.

CPMD V3.11, copyright INTERNATIONAL BUSINESS MACHINES CORPORATION (1990-2006) and MAX PLANCK INSTTUTE FUER FESTKOERPERFORSCHUNG STUTTGART (1995-2001).

L.H. Dubois and G.A. Somorjai, 1980, The Chemisorption of CO and CO2 on Rh(111) Studied by High Resolution Electron Loss Spectroscopy, Surf. Sci., 91, pp. 514-532.

R. Mukherjee, J.C. Flake, A. Palazoglu and J.A. Romagnoli, 2008, A Multiresolution Spatial Correlation Approach For Line-Edge Roughness Characterization, J. of Wavelet Theory and Applications, Accepted.

M. Gajdos, A. Eichler andJ. Hafner, 2004, CO Adsorption on Close-packed Transition and Noble Metal Surfaces: Trends from ab initio Calculations, J. Phys. Condens. Matter, 16, pp. 1141-1164.

S.H. Oh, G.B. Fisher, J.E. Carpenter and D.W. Goodman, 1986, Comparative Kinetic Studies of CO-O2 and CO-NO Reactions over Single Crystal and Supported Rhodium Catalysts, J. Catal., 100, pp. 360-376.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimization of Compressor Networks in LNG Operations

M. M. Faruque Hasan, Md. Shamsuzzaman Razib, I. A. Karimi

Department of Chemical & Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117576

Abstract

Liquefied natural gas (LNG) is the most economic way of transporting natural gas (NG) over long distances. Although LNG is an attractive source of clean fossil fuel, it involves energy intensive liquefaction of NG using refrigeration. Often the compressors that run the refrigerant cycles in an LNG plant operate in suboptimal fashion, which results in higher fuel and energy consumption. To this end, we present a generalized model for the compressor operations in multiple interacting refrigerant cycles in LNG and other cryogenic applications. We determine the optimal load distribution between the cycles to minimize total power consumption of the system for a given plant capacity and operating conditions. We also show the applicability of our model using a case study on the AP-X™ LNG process, which includes three interacting cycles, each with single or multiple compressors.

Keywords: Compressor Networks, Refrigeration, LNG Operation, Cryogenics, Optimization, AP-X™ Process.

1. Introduction

Natural gas (NG) is the 'natural' choice among fossil fuels. It is the cleanest with abundant proven reserves. It contains mainly methane (about 90%), ethane, propane, butane, and trace amounts of nitrogen and carbon dioxide. Currently, NG is the world's fastest growing energy commodity and the third largest primary energy source after crude oil and coal. It is also the fastest-growing and second largest energy source for electric power generation. In 2007, NG consumption was 2637.7 million tons oil equivalent, or about 23.8% of the total primary energy consumed worldwide (BP Statistical Review of World Energy, 2008). The consumption is projected to increase by nearly 52% between 2005 and 2030 (International Energy Outlook, 2008).

Most NG reserves are offshore and away from demand sites. An attractive option is to liquefy NG at the source and then transport it as liquefied natural gas (LNG) by specially built ships. When liquefied, the volume of NG reduces by a factor of about 600 at room temperature, which facilitates the bulk transport. In fact, LNG is the most economic way for transporting NG over long distances. Furthermore, more than 90% of the feed heating value in a modern LNG plant is shipped as product LNG (Liu et al., 2001). These along with concerns about energy price and security are making LNG the fuel of the future, with the demand doubling every ten years. In 2007, 226.41 billion

cubic meters of NG was transported as LNG (BP Statistical Review of World Energy, 2008).

However, LNG production is still expensive compared to oil. Production and transportation represent nearly 85% of the cost of delivering LNG to the customer's jetty. A world-scale LNG plant usually consumes about 5.5–6 kWh energy per kmol of LNG produced. Therefore, energy is an immediate concern in LNG industry. NG is cooled to and liquefied at around -163 ºC to produce LNG using energy intensive liquefaction via costly refrigeration. Such refrigeration system involves some of the largest compressors in the world, usually driven by frame-type gas turbines using NG as fuel (fuel from feed or FFF) or electric motors. About 40% of the total operating cost of an LNG plant is due to energy consumption in the compressors that run refrigerant cycles. Therefore, optimizing compressor operation in LNG plants is crucial for reducing energy cost, consumption of FFF, emission of CO_2, etc.

Aspelund et al. (2007) proposed a methodology based on pinch analysis to utilize pressure based exergy (available energy) for sub-ambient processes, such as LNG. Del Nogal et al. (2008) presented an optimization framework for the design of mixed refrigerant (MR) cycles which are suitable for LNG. Shin et al. (2007) proposed a mixed integer linear programming (MILP) formulation for optimizing boil-off gas (BOG) compressor operations in an LNG receiving and re-gasification terminal. They minimized the total average power consumption of the BOG compressors. Selot et al. (2008) presented operational planning model for NG production systems. However, hardly any literature rigorously considers compressor operations for LNG production.

Power consumption in the turbine driven compressors is highly nonlinear and depends on three major factors. First, ambient and cooling water temperatures determine utility consumption, turbine efficiency, and the performance of inter-stage coolers present in the compressor network. Second, the refrigeration loads and hence the power consumption in an LNG plant vary with feed NG composition, flow, temperature, and pressure. Plant operators either run the compressors at their maximum speeds, or adjust the speeds according to the changes in feed and ambient conditions. However, they do it based on their experience and try-and-see approach. This often results in suboptimal power consumption levels. Third, while a pipeline transportation and distribution system for NG has a single network of compressors, modern LNG plants use multiple interacting and complex refrigerant cycles involving highly nonlinear dynamics. For instance, the propane pre-cooled mixed refrigerant (C3MR) process involves two interacting cycles, Casecade Phillips LNG process has three interconnected cycles, Dual MR process has two cycles in series, and Liquefin Axens process uses two interconnected cycles. The complex interaction between the cycles poses a major challenge in the optimization of compressor operations in an LNG plant.

Only recently, Hasan et al. (2009) addressed this problem. They optimized the compressor operations for C3MR process which involves two highly interacting propane refrigerant (PR) and mixed-refrigerant (MR) cycles. They identified propane evaporating pressure in the PR cycle to be the key decision variable that determines the loads of the two cycles. They also obtained the optimum propane evaporating pressure using a nonlinear program (NLP), with the objective to minimize the total power consumption in PR and MR compressors. However, the network complexities and interdependence of the two refrigerant cycles were not addressed rigorously. Moreover, no guideline was provided for the load distribution between the two cycles. Most of all, their model is applicable to C3MR process only, since they did not present any generalized model for any number of refrigerant cycles.

In this article, we present a generalized model for the compressor operations in multiple interacting refrigerant cycles. We determine the optimal load distribution between the cycles to minimize total power consumption in the compressors for a given plant capacity and operating conditions. We also show the applicability of our model using a case study on the AP-X™ process, which includes three interacting cycles, each with single or multiple compressors.
In what follows, we state the compressor network (CN) problem. Next, we present a mixed integer nonlinear program (MINLP) to optimize the CN operations in LNG industry. Finally, we demonstrate the benefits of such an optimized compressor network versus a base configuration for the AP-X™ process.

2. Problem Statement

A plant has N interacting refrigerant cycles (n = 1, …, N). Each cycle uses network that involves turbines (gas or steam), single-shaft turbo-compressors, valves, exchangers, etc. and steam at different pressure levels. Let i and j denote the compressors (i = 1, …, I_n) and components (j = 1, 2, …, J_n) in the refrigerants respectively. Let P, T, and F be the pressure, temperature, and flow of NG respectively and T_f be the LNG temperature. Therefore, the total change of enthalpy of NG is constant. In cycle n, the refrigerant is compressed using compressor C_{in} with inter-stage/after coolers using air/cooling/sea water. The final discharge pressure corresponds to the vapor pressure of the refrigerant at its evaporating temperature, which depends on the air/cooling/sea water temperature (T_{sw}) and hence is fixed a priori. The refrigerant then undergoes a pressure reduction through the JT valves to reach the evaporating pressure at the corresponding evaporating temperature. It is then used to precool/liquefy/sub-cool NG and/or refrigerant(s) by evaporating itself in exchangers (precoolers, MCHE, etc.), which is then sent to the compressors again. Assuming no pressure drop in the exchangers, the evaporating pressure is also the suction pressure to the first operating compressor in the network. With this, CN operation problem can be stated as follows.
Given:
1. NG flow rate, pressure, in/out temperature, and composition
2. Air or sea/cooling water temperature
3. J refrigeration cycles with I compressors and C components

Obtain:
1. Optimal load distribution between the refrigeration cycles
2. Flow of refrigerants
3. In/out pressures of each compressor
4. NG temperature after each refrigeration

Aiming to:
Minimize the total cost of power in the compressors.
Assuming:
1. No pressure drops in coolers
2. Identical compressors in each cycle
3. Constant turbine efficiency

We now discuss our MINLP formulation for the above CN operation problem, and apply it to the case of AP-X™ process with a base configuration. Unless stated otherwise, all indices such as i, j, c, etc. assume the full ranges of their valid values in all the constraints.

3. MINLP Formulation

The model involves two sets of primary decisions. The first selects the compressor that is operating using the following binary variable.

$$x_{in} = \begin{cases} 1 & \text{if compressor } i \text{ in cycle } n \text{ is operating} \\ 0 & \text{otherwise} \end{cases}$$

The more is the load in one cycle, the less are the loads in other cycles, and vice versa. However, if the evaporating pressure of the refrigerant is low, then the load in the cycle and the duty in the exchanger are more and the compression ratio is high. This results in lower loads and hence lower duties to the exchangers in other cycles which would then operate with higher evaporating pressure. This results in less compression ratio for those cycles and hence, less power consumption. Therefore, the second set of decisions is to select the evaporating pressure of the refrigerant or the inlet pressure to each compressor network. Let $P_{1,n}$ and $P_{2,n}$ be the in and outlet pressure of the compressor network in cycle n. While $P_{1,n}$ is the decision variable, $P_{2,n}$ is known and equal to the vapor pressure of the refrigerant at its evaporating temperature in cycle n.

Furthermore, since the flow and pressure of the refrigerant affects the efficiency and power consumption of the compressors, the compression ratio (or, the suction/inlet pressure to each compressor) is equally important. Therefore, we also need to select the in and outlet pressures for each compressor. Let p_{in} ($p_{(i+1)n}$) be the in (outlet) pressure of compressor i in cycle n. Since the surface areas and capacities of the exchangers and compressors respectively are fixed, they might operate only in a range of pressures. Therefore,

$$p_{in}^{L} \le p_{in} \le p_{in}^{U} \qquad (1)$$

Depending on p_{in}, all I_n compressors in cycle n may or may not operate. However, if compressor C_{in} operates, then $C_{(i+1)n}$ must operate to increase the pressure continuously, until the final discharge pressure is achieved. Moreover, whenever a compressor operates, it increases the pressure up to its maximum capacity. In other words, if compressor C_{in} operates, then $p_{(i+1)n} = p_{in}^{U}$. Therefore, to model the compressor selection, we use,

$$P_{1,n} \ge p_{in}^{U}(1 - x_{in}) \qquad (2)$$

$$p_{(i+1)n}^{U} x_{(i+1)n} \ge p_{in}^{U} x_{in} \qquad (3)$$

$$P_{1,n} \ge p_{in} - P_{2,n}(1 - x_{in} + x_{(i-1)n}) \qquad (4)$$

$$P_{1,n} \le p_{in}^{U} + P_{2,n}(1 - x_{in} + x_{(i-1)n}) \qquad (5)$$

$$p_{in} \le p_{(i-1)n}^{U} x_{(i-1)n} + P_{2,n}(1 - x_{(i-1)n}) \qquad (6)$$

Let W_{in} be the power consumed in C_{in}. We use,

$$W_{in} \ge \omega_n T_{in}^{m} M_n \left(\frac{p_i^{U}}{p_i} - 1 \right) \qquad (7)$$

where, ω_n is a known factor that depends on the compressor efficiency and polytropic constant, T_{in}^m is the known mean operating temperature of compressor C_{in}, and M_n is the variable flow of refrigerant in cycle n.

Let TF_n and TR_n be the temperatures after the exchanger where NG and other refrigerants respectively are precooled / liquefied / subcooled in cycle n. An energy balance around the exchangers must be applied to determine TF_n, TR_n, and M_n. We use Antoine equation to express the vapour pressure and temperature relations. The energy balances and the expressions for pressure, temperature, and enthalpy relations for each cycle are similar to those used by Hasan et al. (2009).

Finally, the objective is to minimize the total power consumption in the system and is given by

$$\min \sum_i \sum_n W_{in} \qquad (8)$$

4. Case Study

We take the AP-X™ process (Figure 1) for LNG as the case study. It has three refrigeration cycles, namely PR, MR, and N2. In the PR cycle, propane is used for pre-cooling NG as well as MR. Propane is compressed using multistage propane compressors and then cooled using sea/cooling water. The final discharge pressure ($P_{2,PR}$) corresponds to the vapor pressure of propane at its evaporating temperature, which depends on T_{sw}. Propane is then flashed to a reduced pressure $P_{1,PR}$ at the corresponding vaporizing temperature. This propane is used to precool NG and MR by evaporating itself, which is then sent to the PR compressors again. MR is further cooled in the MCHE.

After performing a pressure reduction through the JT valves, MR is sent to the shell side of the MCHE from the top and used as the two-phase evaporating refrigerant to liquefy NG. The MR pressure $P_{1,MR}$ can be varied depending on the heat load in the MCHE. Once MR leaves the MCHE, it is compressed in the MR compressors and cooled using sea water. Unlike propane, MR is not totally condensed after passing through sea water aftercooler. However, similar to the PR cycle, sea water temperature influences the MR cycle performance.

In C3MR process, final subcooling of NG is performed in the MCHE and the temperature exiting the exchanger is about -150^{0}C to -162^{0}C. However, in AP-X™ process, the temperature exiting the exchanger is about -115^{0}C and final subcooling is carried out in N2 cycle. In N2 cycle, nitrogen is compressed to a high pressure using nitrogen compressors and then cooled to ambient temperature using sea water. The high pressure nitrogen is then cooled in a N2-N2 heat exchanger by exchanging heat with low pressure nitrogen returning to the compressor after executing cooling duty with LNG. After cooling the high pressure nitrogen is passed through an expander where it is expanded to a lower pressure as well as cooled further to a lower temperature. Then the cooled low pressure N2 executes final subcooling of LNG. The use of nitrogen to subcool LNG offers reduction of total refrigeration load on PR and MR cycles.

We apply our model to a base configuration of AP-X™ process with NG feed at 303 K and 64 bar, desired LNG temperature of 123 K. We use sea water at 303 K as the cooling medium in the interstage- and after-coolers. The compressor networks are consist of 2, 3, 1 compressors for PR, MR and N2 cycles respectively. The computing platform is an AMD Athlon™ 64×2 Dual Core Processor 6000+ 3.00 GHz, 3.00 GB of RAM using CPLEX v.10.0.1 (LP solver), CONOPT v. 3 (NLP solver) and DICOPT

(MINLP solver) in GAMS 22.2. Results show the final load distribution between the PR, MR, and N2 compressors are 22.2%, 56%, and 21.8% respectively. All 6 compressors need to operate at these loads. However, such an optimized operation can achieve more than 7% decrease in total power cost compared to that of the base configuration.

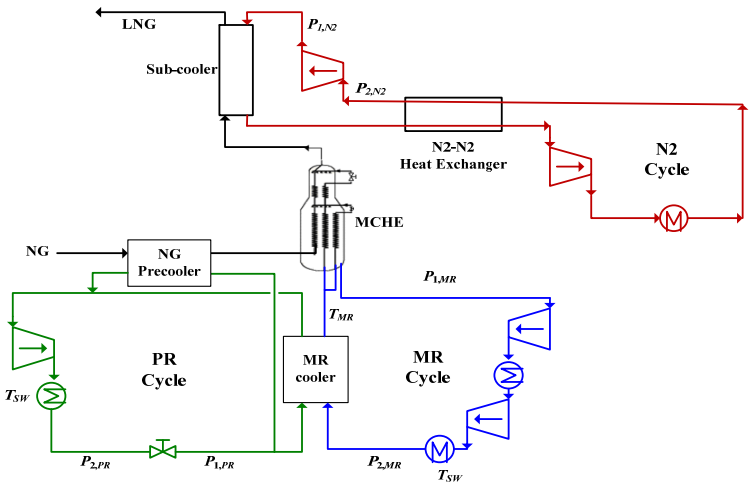


Figure 1. Simplified diagram of AP-X^{TM} process for LNG

5. Conclusion

A generalized model is presented for the optimization of compressor operations in multiple interacting refrigeration cycles which are usually common in LNG, air separation, gas processing, etc. We also determine the load distributions between the refrigeration cycles for minimum power consumption in the compressors. However, the model should be extended to consider variations over time and uncertainties in NG feed flow rate, composition, pressure, and temperature, and seasonal variations in sea/cooling water temperature, etc.

References

A. Aspelund, D. O. Berstad, T. Gundersen, 2007, An Extended Pinch Analysis and Design Procedure Utilizing Pressure based Exergy for Subambient Cooling, Applied Thermal Engineering, 27, 2633-2649

F. Del Nogal, J. K. Kim, S. Perry, R. Smith, 2008, Optimal design of mixed refrigerant cycles, Industrial & Engineering Chemistry Research, 47(22), 8724-8740

M. M. F. Hasan, I. A. Karimi, H. E. Alfadala, 2009, Optimizing compressor operations in an LNG plant, Proceedings of the 1st Annual Gas Processing Symposium, H. Alfadala, G. V. Reklaitis amd M. M. El-Halwagi (Editors), 179-184

Y. Liu, C. E. Lucas, R. W. Bower, J. C. Bronfenbrenner, 2001, Reducing LNG Costs by Better Capital Utilization, Presented in the 13th International Conference and Exhibition on Liquefied Natural Gas (LNG13), Seoul, Korea, 14-17 May, 2001.

A. Selot, L. K. Kuok, M. Robinson, T. L. Mason, P. I. Barton, 2008, A short-term operational planning model for natural gas production systems, AIChE Journal, 54, 495-515

M. W. Shin, D. Shin, S. H. Choi, E. S. Yoon, C. Han, Optimization of the Operation of Boil-Off Gas Compressors at a Liquefied Natural Gas Gasification Plant. Industrial & Engineering Chemistry Research, 46, 6540-6545

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Development of a micro heat exchanger made with ceramic multi-layers (LTCC) and its setup to gas flow measurements

Elsa Vásquez-Alvarez[a], Francisco T. Degasperi[b], Mario R. Gongora-Rubio[c,d] and Reinaldo Giudici[a]

[1]Department of Chemical Engineering, Polytechnic School - University of São Paulo - Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP: 05508-900 – São Paulo – SP – Brazil. (rgiudici@usp.br).
[2]Faculdade de Tecnologia de São Paulo – FATEC-SP – CEETEPS – UNESP – São Paulo, SP – Brazil.
[3]Instituto de Pesquisas Tecnológicas – CTPP - CEP: 05508-901 São Paulo - SP – Brazil.
[4]Laboratório de Sistemas Integráveis, Escola Politécnica da Universidade de São Paulo – CEP: 05508-900 – São Paulo – SP – Brazil.

Abstract

A green ceramic tape micro heat exchanger was developed using LTCC technology. The device was designed by using a CAD software and 2D and 3D simulations using a CFD package (COMSOL Multiphysics) to evaluate the fluid behavior in the micro-channels. The micro heat exchanger is composed of five thermal exchange plates in cross flow arrangement and two connecting plates; heat exchanger dimensions are 26 x 26 x 6 mm^3. Preliminary tests were carried out to characterize the device both in atmospheric pressure and in vacuum. The same techniques used in vacuum technology were applied to check the rotameters and to prevent device leakages. Thermal performance of the micro heat exchanger was experimentally tested.

Keywords: LTCC, micro structured heat exchanger, cross flow, simulation.

1. Introduction

Global competition challenges the chemical industry to search for technology innovations and, for this to happen, relevant organization units are necessary, such as market and sales, research and development, process technology and engineering. In this context of globalization and sustainability, it is evident that process intensification is a step for the future of the area. Process intensification is the development of new equipment and/or techniques which allow a large improvement in some of the production stages in the chemical industry, increasing the area/volume relation of the system and decreasing resistances to the energy and mass transportation in order to increase the correspondent transfer rates.
The advance in Micro fabrication technology opened a new field of research in the Chemical Process area. These techniques applied to microelectronics created huge new markets and are now applied to Chemical Engineering, creating new horizons for extraordinary developments. Low Temperature Co-fired Ceramics (LTCC) technology

enables the possibility of fabricating 3D devices like holes, channels and hollows in a scale from one micron to few millimeters using multiple-layer green ceramic.
The objective of the current work is the development, characterization and performance tests of a green ceramic micro heat exchanger. The device was designed by using a CAD software and simulated in 2D and 3D using a CFD package (COMSOL Multiphysics) to evaluate the fluid behavior in the micro-channels. The construction was made using the LTCC technology. Several Computer Numeric Control (CNC) machining tests, glass ceramic tapes adherence, device lamination and sintering were made before the fabrication of a 3D device was feasible. The micro heat exchanger is composed of five thermal exchange plates in cross flow and two connecting plates with 26x26x6 mm^3. Firstly, the experimental setup was characterized and performed using tests developed according to the gas flow measurement techniques of vacuum technology, more specifically, conductance. Rotameters were used to measure the N_2 gas flow. The leak tests were performed both in the micro heat exchanger and the gas flow setup junctions. The experimental setup is running and the regime of gas flow is viscous (Knudsen < 0.01).

2. Design and Fabrication

Several different geometry projects were analyzed in order to define the device construction, but only four of them are discussed in this work. The geometries were simulated and the results were compared to decide the best configuration for construction. The micro heat exchanger project presents stacked plates with parallel and rectangular shape channels.

2.1. Micro plate designs

Several variations in the fluid distribution were studied in order to assess and determine the best velocity distribution in each one of the plate channels. Among the various options in terms of channel number, length, width and height, four geometries were projected in the CAD. Each plate has two holes: one for the passage to the next plate and another for the hollow outflow. All evaluated geometries presented parallel micro channels with 15,677 mm length, 350 μm width and 500 μm depth.
The main challenge of the project was the hollow that distributes the fluid to the channels. Hence, Figure 1 shows the fluid flow in the proposed geometries. Thus, the first proposed project (A-type) has a side entrance and a triangular shaped fluid distribution hollow (see Figure 1a). In the second project, the B-type, the fluid inflow happens in the central part of the plate and the distribution is also triangular shaped, according to Figure 1b. In the third proposal (C-type), the fluid distribution is in a circular shape, according to Figure 1c. There is a fourth geometry (D-type) which holds two walls in the edges as in Figure 1d.

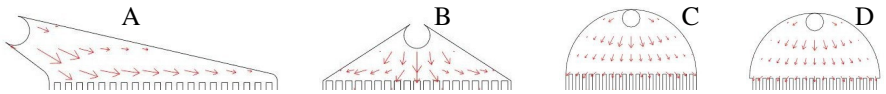


Figure 1. Fluid distribution hollows proposed for the micro heat exchanger plate.

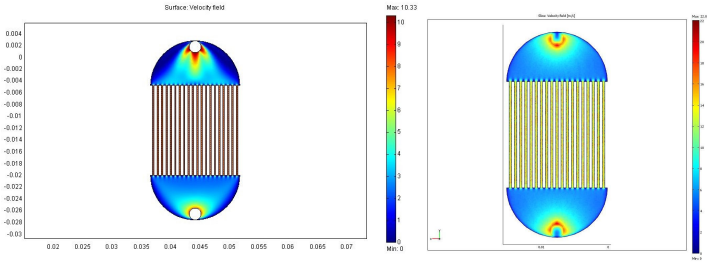
2.2. Fluid behavior in CFD

The geometry project for the device plate involves the fluid dynamics, so simulations of the correspondent processes are necessary to make the device project cheaper, faster, more reproducible and more controllable. In this case, a simplified CFD simulation was

used to assess the effect of the fluid distribution chamber to the channels and determine the best geometry. This simulation was made with the COMSOL Multiphysics v 3.3a, considering nitrogen at 600 K with 1 bar pressure, 0,561 kg/m3 density, 2,95.10-5 pa.s viscosity and 10m/s velocity as the work fluid. To make the simulation easier, it was assumed that the fluid had a top-bottom direction
The assumed boundary conditions were:

- No fluid slipping on the walls
- Zero relative pressure in the outflow and
- Inflow velocities specification,

The incompressibility, the laminar regime and a Knundsen value below 0.01 were adopted as simplifications in the problem, in order for the environment to be treated as continuous (Commenge, 2002). Figure 2 shows the qualitative result of the CFD simulation both in 2D and 3D for the D-type geometry.



a) 2D Simulation b) 3D Simulation

Figure 2. COMSOL Simulations for D-type geometry.

2.3. LTCC Technology

The aluminum vitroceramic in the "green" state (before sinterization) enables the generation of holes, channels and hollows in a simple way up to 200 micrometers. After the individual layer process, they are laminated and a multi-layer system is obtained. When the laminated is co-sinterized, a hard body is generated with the desired application. This technology was initially developed for high reliability, power and frequency applications and it has been an excellent alternative for the silicon in the construction of equipment used in chemical processes, because it allows the generation of 3-D structures with the added possibility of integration of the sensors and thermal actuators. The fabrication process using the LTCC hybrid technology was described in general terms by Gongora-Rubio et al. (2001). Recent researches have shown the integration of microfluidics and microeletronics using the LTCC hybrid technology, such as Martinez-Cisneros et al. (2007) and Ibáñez-García et al. (2008), but a new approach is presented in this work: the development of a micro heat exchanger with multi-plates.

3. Experimental Procedure

3.1. Micro heat exchanger construction

For the fabrication of the micro heat exchanger plates, green ceramic 951 PX tapes with 250 μm thickness from Dupont were used. Since this technology is based on the multilayer approach, the current project considers four stacked layers for a specific geometry. The device construction process follows the sequence presented in Figure 3.

The geometries are projected in the CAD software and, then, they are simulated in a CFD software in order to choose the best fluid dynamics behavior.
In order to stack the tapes, it was necessary to carry out different compaction tests with variable temperatures, time and pressure parameters to prevent delamination, because the channels depth is 500 µm (the same thickness of two tapes). The lamination process was made in a uniaxial hydraulic press with the adaptation of an electric resistance set for the thermal transfer. The microchannel machining was made with a CNC PROTOMAT C100/HF milling cutter, which uses 400 µm cutters and 2.00 mm drills. Then, the green ceramic plates are glued together in a crossed way with a thin layer of an organic fluid (Da Rocha et al. 2004) and they are laminated at a 26 MPa pressure. The micro heat exchanger has three plates for one flow and two plates for the other.

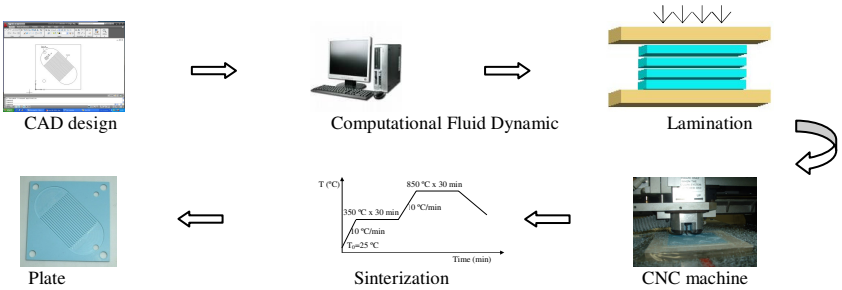
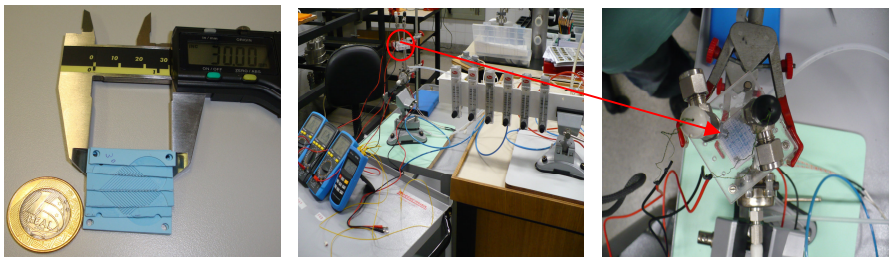


Figure 3. Fabrication process of the micro heat exchanger plates.

3.2. Experimental Setup with Procedures

The cross flow micro heat exchanger (See figure 4a) was previously installed in a metal housing which was specially projected for this experiment, but this material led to huge heat losses. Then, in order to avoid this problem, an acrylic housing was later projected and constructed (see Figure 4b). The micro heat exchanger was attached to the acrylic housing with Kalrez o-rings, which can also be used to avoid heat losses.



a) b) c)

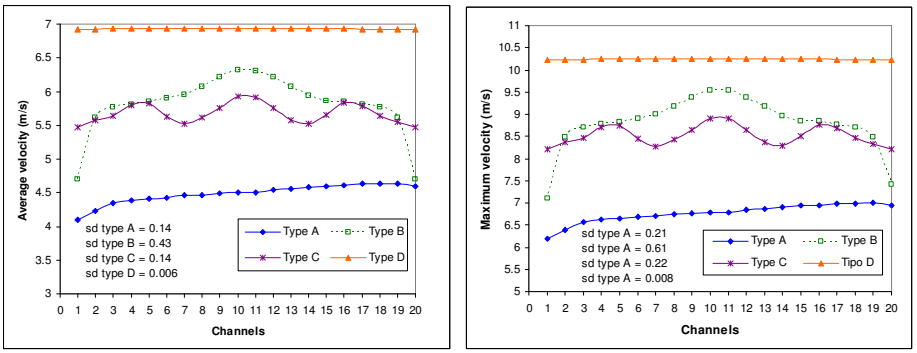
Figure 4. Plates in LTCC and experimental setup of the micro heat exchanger.

The gas flow circuits were considered using N_2 gas, rotameters to measure gas flow, thermocouples to measure the temperatures in each point of gas inlet and outlet in the micro heat exchanger. It was necessary to build a special system of temperature measurement adapted from commercial thermocouples, since the dimensions are extremely reduced. Electrical heater is used to heat up the warm gas stream entering in the exchanger (see Figure 4c) to the desired temperature.

4. Result and discussion

4.1. CFD Simulation

The analysis of the simulations in function of the average and maximum velocities developed in each channel is shown in Figures 5a and 5b, respectively. It indicates that the D-type geometry presents the best flow behavior with a standard deviation in the maximum velocities of 0.0086 m/s and a Re = 54.



a) b)

Figure 5. Comparison of the simulation results in the four proposed geometries.

4.2. Experimental results

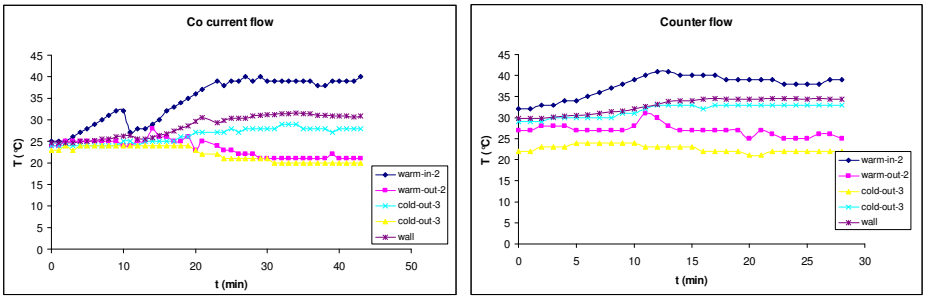
The preliminary tests were made in atmospheric pressure for different flows and configurations in the micro heat exchanger.

The first experimental test was performed for a flow of 40 cm^3/min of each one of the gases, using the device with the following configuration: three plates of warm fluid and two plates for the cold fluid, the system was tested in co current flow. The temperature measures were made with thermocouples installed in the two inlet flows and in the two exiting flows. It is important to emphasize that the warm fluid is pre-heated before entering in the heat exchanger, by using an independent electrically heated device. The experience was executed for a period of 42 min and the result indicates that the steady state was not attained (Figure not shown). The global energy balances were used to both warm and cold fluid as well as the material properties to calculate the heat changes and confirmed that the wall is in transient regime, with a rate of average temperature of 0.0834 ºC/min.

The second preliminary test presented in this work was performed for the N_2 with 10 cm^3/min in both entering gas streams, with the following configuration: two plates for the warm gas and three plates for the cold gas. The system was tested for the input flow in parallel. The temperatures were measured in four points for 42 min, which shows that the steady state was attained, as can be seen in Figure 6a. By checking the experimental result through global energy balance one can observe that there is a heat loss of 3.75%, this possibly occurs due to changes in the external surrounding. It can be observed that; in this case; the residence time of gas is larger than that in the configuration of test 1. This explains the steady state attained faster for the same time period evaluated.

The third experimental test made with the micro heat exchanger was the 10 cm^3/min in both feed streams using the same configuration of test 2, but in counter flow. The experiment had a duration time of 28 min, and it was observed that the system reached the steady state as show in Figure 6b. The global energy balance was performed in each

one of the circuits. It was observed a heat loss even lower (1.05%) compared with that calculated for the test 2.
In the three experimental tests performed, the temperature changes in the warm and cold flows were not exactly the same, even though both streams were air with the same flow rate. This is possibly due to the insufficient system insulation.



a) b)

Figure 6. LTCC device experimental results.

Conclusion

LTCC technology is quite useful for developing novel devices to integrate non-electrical functions for microfluidic, analytical chemistry, microplasma and microreaction applications
In this work, a crossflow multi-plate micro heat exchanger was designed, manufactured and experimentally tested. The CFD simulations for one of the plates show the fluid distribution for every channel and the D-type geometry presented the best behavior of flow distribution when compared with the three other geometries considered.
The preliminary tests indicate that the test 3 (counter flow) was the most efficient, and the steady state has achieved faster and with lower heat losses compared with system in parallel flow (test 2).
Further modifications are ongoing in order to improve electronic instrumentation, temperature sensor measurement, flow measurement and system thermal insulation.

Acknowledgements

This work was supported by CNPq-CTPetro project (502637/2003-0) and the individual grant number 383609/06-2 and it is now supported by FAPESP (2007/03812-5).

References

J-M. Commenge, Falk, L.; Corriou, J-P.; Matlosz, M., 2002, Optimal Design for flow uniformity in Microchannel Reactors. AICHE Journal, v. 48, 2, pp. 345 – 358.

Z. M. Da Rocha, Ibañez-Garcia, N.; Oliveira, N. A.; Matos, J. Do R., 2004, Low Temperature and Pressure Lamination of LTCC Tapes for Meso-Systems, In: IMAPS Conference And Exhibition On Ceramic Interconnect Technology, Denver.

M. R. Gongora-Rubio, Espinoza-vallejos, P.; Sola-Laguna, L.; Santiago-Aviles, J. J., 2001, Overview of Low Temperature Co-Fired Ceramics Tape Technology for Meso-System Technology; Sensors and Actuator A-Physical; v. 89, pp. 222 – 241.

N. Ibáñez-García, Martinez-Cisneros, C. S.; Valdés F.; Alonso J. 2008, Green Tape ceramics. New technological approach for integrating electronics and fluids in Microsystems, v. 27, pp. 24-32.

C. S. Martinez-Cisneros, Ibáñez-García N., Valdés F., Alonso J., 2007, LTCC Microflow analyzers with monolithic integration of thermal control, Sensors and Actuator A-Physical, v. 138, pp. 63-70.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Dynamic Simulation of Nuclear Hydrogen Production

Patricio D. Ramírez,[a] Mujid S. Kazimi,[b] Paul I. Barton[a]

[a]*Department of Chemical Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 20139, USA.*
[b]*Department of Nuclear Engineering, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, USA.*

Abstract

Hydrogen is an important commodity in the chemical industry, but its production is not CO_2-neutral. Currently, 95% of hydrogen is produced from hydrocarbon feedstocks and it releases 12kg of CO_2 per 1kg of H_2 produced. Using nuclear energy to power water-splitting processes is a cleaner alternative to produce H_2 without emitting CO_2. The operation of such a system is difficult and potentially dangerous, and safety needs to be demonstrated using dynamic simulation. However, current approaches to demonstrate safety cannot deal with the complexity added by chemical reactions. A dynamic process simulator such as JACOBIAN® allows the simulation and study of such a system. This paper describes the implementation of a dynamic model of a nuclear hydrogen production plant in JACOBIAN®, its validation and simulation.

Keywords: hydrogen, nuclear energy, high-temperature electrolysis.

1. Introduction

Hydrogen plays an extensive role in the chemical industry; it is used in the refining of crude oils, the production of methanol and the synthesis of ammonia. Roughly 50 million tonnes of hydrogen are produced worldwide every year, and new applications will require higher levels of production. The conversion of tar sands, shale oil and coal to liquid fuels are examples of growing applications. Hydrogen demand will also increase with the conversion of biomass to liquid fuels, in the search for CO_2-neutral fuels. Other future applications will include fuel cell powered vehicles.

Unfortunately, current hydrogen production is not CO_2-neutral, as it emits large amounts of CO_2 and is based on fossil fuels. 95% of hydrogen is currently produced from hydrocarbon feedstocks, and the production processes release around 12kg of CO_2 per kg of hydrogen. In particular, hydrogen produced in this way cannot be used to create CO_2-neutral fuels derived from biomass. Nuclear hydrogen production can overcome these difficulties. The heat and/or electricity from a nuclear reactor can split water into O_2 and H_2, avoiding the production of CO_2 and the use of fossil fuels. High temperature electrolysis and the Sulphur-Iodine process are considered the most promising technologies to do this, and national laboratories in the U.S., Japan and France are working on their development.

However, the operation of a nuclear reactor coupled to a chemical plant becomes more complex, making it harder to demonstrate the safety of the overall system during transients. Demonstrating safety during transients is needed to gain approval from the regulatory agencies, to get public acceptability, and to achieve good economics and plant functionality. The coupled system must behave safely under different transients:

start-up, shut-down, re-start, off-normal response, and accidents. Normally, the safety of nuclear reactors is demonstrated using simulation codes developed specifically for that purpose, for example, RELAP5 (INL (2006)). In this case though, these simulators cannot handle the complexity added by the chemistry and thermophysical properties of the hydrogen production processes.

This problem was solved using JACOBIAN® (NumericaTech (2005)), state-of-the-art dynamic simulation software, and by developing adequate physico-chemical models. JACOBIAN® is an open modeling environment and simulator, which supports the formulation of models of chemical, physical and biological processes. It allows the efficient simulation of processes, because it solves large-scale, stiff, sparse ODEs and DAEs, provides model and sensitivity analysis tools, and can simulate hybrid discrete/continuous models. A model library was developed to represent a 600MWth nuclear reactor, a heat transfer loop and a high-temperature chemical reaction section for hydrogen production. The models allow the study of different transient scenarios such as start-up and off-normal responses. This paper explains the implementation and results of these studies.

2. Description of Process and Dynamic Process Simulator

The nuclear hydrogen production process considered in this paper uses high-temperature electrolysis to split water. The process contains four components: the nuclear reactor, the gas turbine, the heat transfer loop and the high temperature electrolysis cell (Fig. 1). In this process the nuclear reactor provides the heat for the gas turbine and for the electrolysis cell. The gas turbine produces electricity that will be used to carry out the chemical reactions in the electrolysis cell. Part of the heat produced in the nuclear reactor is transferred to the electrolysis cell by the heat transfer loop. This loop guarantees the safety and operability of the process; it physically separates the nuclear reactor from the hydrogen production plant and decouples their operation by introducing auxiliary heat sources/sinks (not shown in figure). Using heat from the nuclear reactor to perform the electrolysis at high temperatures increases the thermodynamic efficiency of the process.

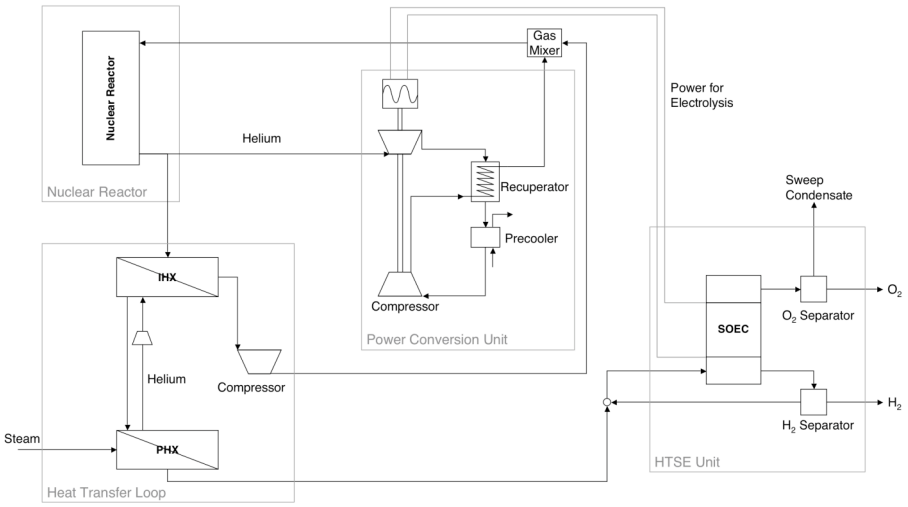


Figure 1. High-temperature steam electrolysis flow diagram.

Equation-oriented process modeling environments can simulate the complexity of such a system in a natural and intuitive way (Pantelides and Barton (1993)). The dynamic behavior of a process can be captured by very detailed models, which can only be constructed in an evolutionary fashion, starting from a simple description and increasing the detail and sophistication when more data and understanding becomes available. Therefore, a dynamic modeling environment, that can offer flexibility in the evolution of models, the possibility of incorporation of discrete and continuous dynamic phenomena simultaneously, and expandability to include other analysis (sensitivity analysis, parameter estimation, optimal control studies) is an enormous help in the design, construction and commissioning of the nuclear hydrogen production process.
JACOBIAN® is an open modelling environment and simulator (NumericaTech (2005)) with these characteristics and it has been selected to simulate nuclear hydrogen production processes. JACOBIAN®, in particular, has state-of-the-art algorithms to identify well-posed systems, to automatically find index, to locate state events, to do hierarchical submodel decomposition, do sensitivity analysis, and to integrate stiff, sparse ODEs and DAEs.

3. Model Formulation

Simplified equations can be used to represent the dynamics of the system, because time scales of less than a second are not important in understanding its operation. This allows to use simpler models with a sufficient level of detail.
In particular, simulating the flow of gas inside the nuclear reactor and inside the heat transfer loop requires a simplified and faster-to-integrate version of the inviscid Navier-Stokes equations. The full system of inviscid Navier-Stokes equations is a hyperbolic system, which means it cannot be integrated efficiently in a dynamic process simulator. However, the fast time scales associated with sound waves can be eliminated from the model, because the focus is the relatively slow transients. This is done by eliminating the transient part in the continuity and momentum equations, and by eliminating the kinetic energy term in the momentum equation, Eq. (1). The resulting system is a parabolic system, which can be integrated in a dynamic process simulator.

$$\frac{\partial \rho v}{\partial x} = 0$$

$$\frac{\partial}{\partial x}(P) = -\frac{2\rho f v|v|}{D} \qquad (1)$$

$$\frac{\partial \rho e}{\partial t} + \frac{\partial}{\partial x}(\rho e v + P v) = -\frac{4U(T - T_{ext})}{D_T}$$

The dynamics of the nuclear reactor are represented using a point kinetics model, a reactivity model and a two-dimensional thermal-hydraulic model (Wang (2003)). These models represent the core neutronics and the heat transfer between the core and the coolant gas. Core neutronics take into account the gross fission power variation due to reactivity change. At the same time, the factors influencing the reactivity are: control rod movement, fuel temperature change and fission product poisoning, as well as external reactivity disturbances. The thermal-hydraulic model represents the heat transfer in the core driven by convection, conduction and radiation.The piping and heat exchangers are also represented by specializing Eq. (1).

The model was built by discretizing the space coordinate. The equations describing the flow of gas in pipes and the heat exchangers were discretized in order to represent the variations along them. 40 nodes were used in the heat exchangers and 20 nodes were used in the pipes. The final model consisted of 1500 variables.

4. Results

The results of the distributed model in JACOBIAN® at the steady-state operating point compare well with the results of the original lumped model design by Davis *et al.* (2005). The values of different variables in the loop showed a discrepancy smaller than 3.5%, and the results of both models are shown in Fig. 2.

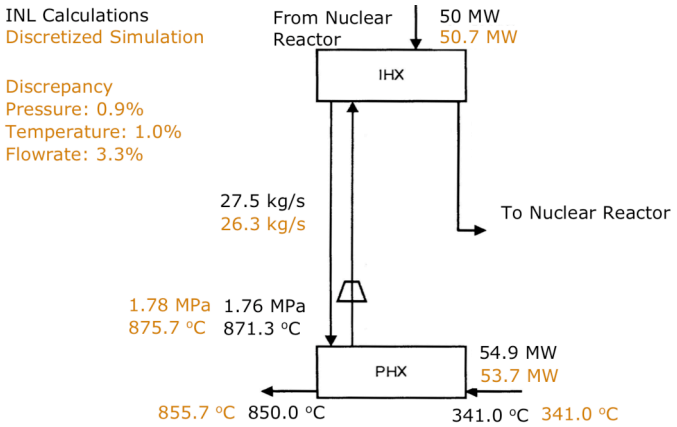


Figure 2. Comparison of results at steady state generated by INL and distributed.

A model in RELAP5 was built to validate the dynamic performance of the distributed model of the heat transfer loop in JACOBIAN®. RELAP5 is a dynamic simulator used to study nuclear reactors, and Fig. 3 shows the structure used to create the model in this platform (based on the work of Davis *et al.* (2005)). At steady state, both models show similar profiles for temperature and pressure.

A simple start-up schedule was simulated as a test for the JACOBIAN® model of the system including the nuclear reactor and the heat transfer loop. The start-up schedule had four stages (Table 1), and the first one was to stabilize the system at a low power level. This was achieved by running the compressors and nuclear reactor at 1% of the designed power level and waiting until the system reached a steady state. Second, the heat removal system was started by increasing the power of the compressors in both loops, and by increasing the flow of gas on the cold side of the process heat exchanger. Third the heat generated in the nuclear reactor was boosted until it reached the design point. Finally, the temperature on the cold side of the process heat exchanger was raised, representing the start-up of the chemical plant.

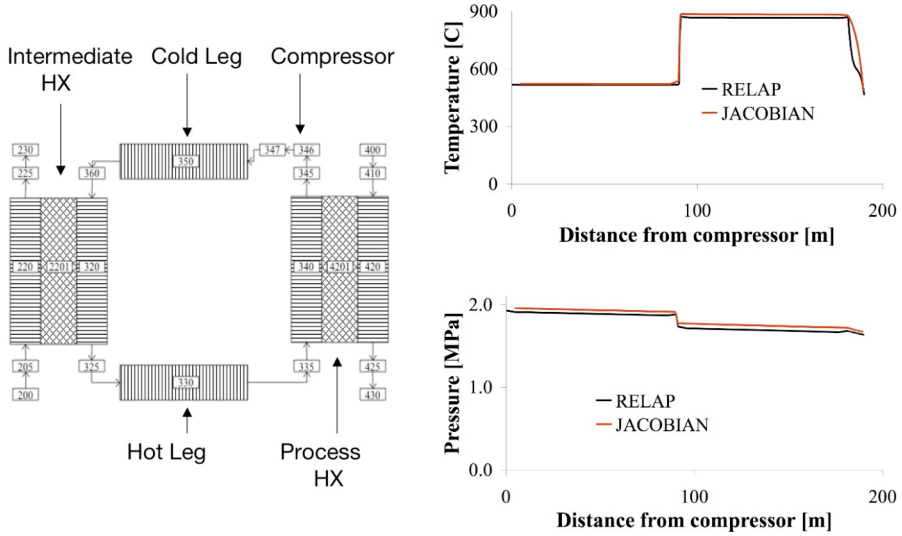


Figure 3. Heat transfer loop for RELAP5 simulation and results at steady state generated by RELAP5 and JACOBIAN® model.

Table 1. Start-up schedule.

Stage	Event	Time
Stabilization	Compressors running at 1%	0 s
Start heat removal	Increase power on compressor in Loop 1 Increase pressure on cold side of PHX Increase power on compressor in Loop 2	300 s 660 s
Increase heat	Add rod reactivity until QFission = 200 MW	986 s
Connection to chemical plant	Increase temperature in the cold side of PHX	6127 s

Fig. 4 shows the temperature of helium at different points in the system. The first point is at the outlet of the nuclear reactor, the second point is at the cold inlet of the intermediate heat exchanger (IHX). The third point is at the hot inlet of the process heat exchanger (PHX). The helium temperatures at the start of the PHX and at the outlet of the nuclear reactor have a similar profile after 1500s. However, these profiles differ before 1500s; the thermal transients generated by the thermal inertia of the IHX explain this. On the other hand, the helium at the inlet of the IHX is one of the coldest points along the loop. These profiles show the temperature gradients transferring the heat from the nuclear reactor to the chemical plant.

5. Conclusions

Dynamic process simulators, such as JACOBIAN®, can be used to study the operation of chemical systems including a nuclear reactor. In the case of nuclear hydrogen production, the JACOBIAN® model shows very good agreement with the original design from INL and with the model created in RELAP5. The model created was able to simulate the start-up of the system and to show the effect of the thermal inertia of the heat exchangers.

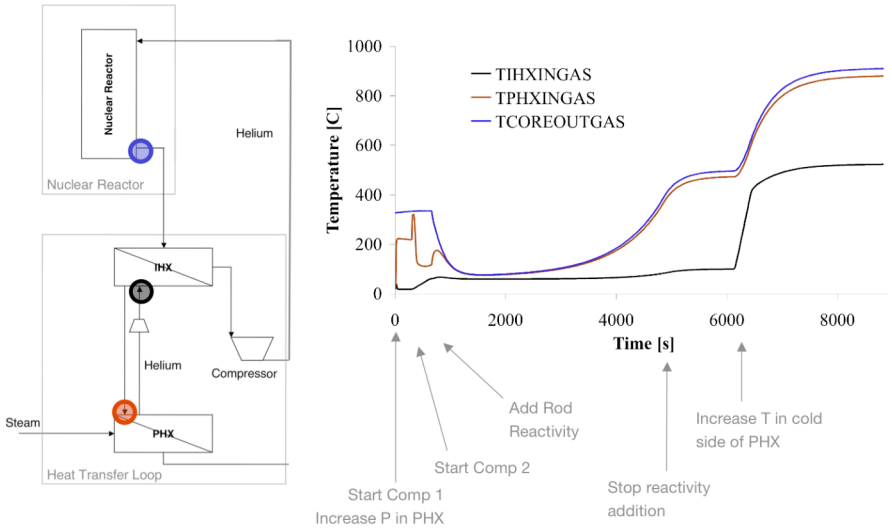


Figure 4. Temperature of helium at various points in the system.

References

C. B. Davis, R. B. Barner, S. R. Sherman, D. F. Wilson, 2005, Thermal-Hydraulic Analyses of Heat Transfer Fluid Requirements and Characteristics for Coupling a Hydrogen Product Plant to a High-Temperature Nuclear Reactor, Technical Report INL/EXT-05-00453, Idaho National Laboratory.

Idaho National Laboratory, 2006, RELAP5-3D Code Manual, INEEL-EXT-98-00834, Rev. 2.4.

C. C. Pantelides, P. I. Barton, 1993, Equation-oriented dynamic simulation current status and future perspectives, Computers and Chemical Engineering, 17, S263-S285.

Numerica Technology LLC, 2005, JACOBIAN dynamic modeling and optimization software, http://www.numericatech.com.

C. Wang, 2003, Design, Anlysis and Optimization of the Power Conversion System for the Modular Pebble Bed Reactor System, PhD Thesis, Massachusetts Institute of Technology.

Acknowledgements

This material is based upon work supported by the Department of Energy Nuclear Energy Research Initiative under Award Number DEFC07-06ID14751.

Disclaimer: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Processing of the Atmospheric Distillation Residue with Supercritical CO_2: Conceptual Project

Ana Mehl[a], Raquel S. Macedo[a], Fernando L. P. Pessoa[a], Silvia M. C. da Silva[a]

[a]*Federal University of Rio de Janeiro – Chemistry School - Av. Athos da Silveira Ramos, 149 - Centro de Tecnologia, Bloco E , Sala 201, Rio de Janeiro/RJ, Brasil CEP: 21941-909*

Abstract

Supercritical fluid extraction is a promising technology for treating heavy feed. The present work proposes an alternative route and uses the processes simulator UNISIN/HONEYWELL to evaluate the supercritical extracting process of the atmospheric distillation residue using carbon dioxide as a solvent. The extraction efficiency is a function of pressure, temperature, oil composition, the ratio between solvent and oil rates and can be improved by cossolvent action. This work presents the results of computer simulation studies in the investigation of the capabilities of supercritical carbon dioxide in the extraction of a light fraction from the atmospheric residua (RAT) for the following range of process variables: 50 - 250 bar (operational pressure), 50 – 150^oC (solvent inlet temperature), a solvent to oil ratio varying from 5:1 to 15:1 and 0-5 mol% of propane as cossolvent. The experimental distillation curve of the feed was used to generate a mixture of pseudo-components representing the RAT. It was observed that the extraction efficiency increases with the solvent to oil ratio and the pressure, whereas it decreases with the increasing of temperature. The RAT vacuum simulation was simulated and the results of both cases were compared.

Keywords: supercritical fluid extraction, CO_2, simulation, heavy oil.

1. Introduction

The current scenario of the global oil production is characterized by the new discoveries and exploration of ultra-heavy and bituminous crude oils. As a result, the refineries should process more and more feeds that exhibit a peculiar behavior in the refining process as well as in the pipeline transportation. Thus, the investigation of technologies to use these fractions, present in greater quantity in the Brazilian oil, is desired. So the development of technology for processing the residue of atmospheric distillation able to withdraw a greater range of products with high added value generating less waste and ensuring that this waste has an appropriate destination is of great importance.

The supercritical fluid extraction may be a viable alternative for the conventional separation process of heavy hydrocarbons in the petroleum industry. The performance of the supercritical fluid extraction in the fractionation of multi-component mixtures depends on several factors among them the solvent density, differences in volatility and intermolecular forces between the solute and supercritical solvent (Hwang et al. (1996)). By use the supercritical fluid extraction technique to treat heavy feeds, products of lower viscosity can be obtained. It represents a cleaner technology wherein it is not necessary to use toxic solvent and the final cuts containing most of the contaminants and asphaltenes can be conveniently discarded (Zhao (2005)).

The atmospheric distillation residuum (RAT) is a fraction with boiling point above 420^{o}C, which is obtained as bottom stream in the atmospheric distillation of the crude oil. It is a residua of high molecular weight, high viscosity, high pour point and low commercial value but it still contains fractions of commercial interest. In the conventional route, RAT is sent to the vacuum distillation followed by the propane desasphaltation in a liquid-liquid extraction. The present work proposes an alternative route and uses the processes simulator UNISIN/HONEYWELL to evaluate the supercritical extracting process of the atmospheric distillation residue using carbon dioxide as a solvent.

The aim of this work is to evaluate the influence of the process variables on the efficiency, using carbon dioxide as the solvent. The effects of the ratio between solvent to oil rate, operational pressure, solvent temperature and the use of propane as an entrainer are investigated.

2. Methodology

The simulation of the RAT extraction with supercritical carbon dioxide was performed using the UNISYM simulator. For comparison purposes the vacuum distillation, which correspond to the conventional refine route, of the same feed was also simulated.

2.1. RAT Characterization

The atmospheric residuum used as the feed of the extraction was analyzed in laboratory and its distillation curve determined by ASTM D2287 is presented in Table 1.

Table 1 – RAT Distillation Curve

Distillation Curve (ASTM D2287) – 760 mmHg, %ml	
IBP	382.2
5/10	410.8/427.2
15/20	441.8/457.5
25/30	475.4/493.9
35/40	512.3/530.9
45/50	549.7/568.2
55/60	587.3/607.1
65/70	627.6/649.1
75/80	673.0/700.8
85/90	726.7/-
95/PFE	-/750.0
(Rec. %m/m)	(87.7)

As in the case of the crude oil, the exact chemical composition of the RAT is unknown, therefore the process simulator uses the distillation curve to define a mixture of pseudo-components to represent the RAT in the simulation. For each of the nine pseudo-components the physical and chemical properties were estimated and the results are presented in Table II. The critical temperature and pressure, the acentric factor and the enthalpy were estimated by the Lee-Kesler correlation and the molecular weight by the Twu correlation. To model the phase behavior of this system the Peng-Robinson cubic equation was selected since it has been extensively used for modeling systems composed by hydrocarbons and light gases. A quadratic mixing rule of van der Walls type was used and the binary interaction parameters were estimated by the simulator. The module chosen to represent the supercritical fluid extraction process was a liquid-liquid extraction column allowing the definition of the number of theoretical stages (N),

where the number of inlet and outlet streams may vary from two (top and bottom) to N. It is necessary to define the pressure profile throughout the column, but in this study the pressure was considered to be constant.

Table II –Pseudo-Componentes Properties

Pseudo-Component	Molecular Weight	Molar Fraction in RAT	Normal Boiling Point(°C)	Density (Kg/m3)	Critical Temperature (°C)	Critical Pressure (kPa)
NBP_470	457.59	0.1105	469.8712	913.3668	618.84	949.5759
NBP_504	504.44	0.1683	503.5198	924.002	645.6333	849.045
NBP_542	562.67	0.1326	541.9196	934.1766	675.166	741.4648
NBP_581	620.71	0.1155	580.8361	948.8383	706.7507	659.6874
NBP_619	685.01	0.1041	619.0583	962.8459	737.4972	587.7264
NBP_657	750.63	0.0914	657.471	976.4508	768.1043	522.5124
NBP_695	815.73	0.0783	695.4081	989.4991	798.0915	464.496
NBP_736	884.89	0.0742	735.5568	1002.996	829.6286	409.4713
NBP_778	962.98	0.1252	777.539	1018.055	863.0186	360.584

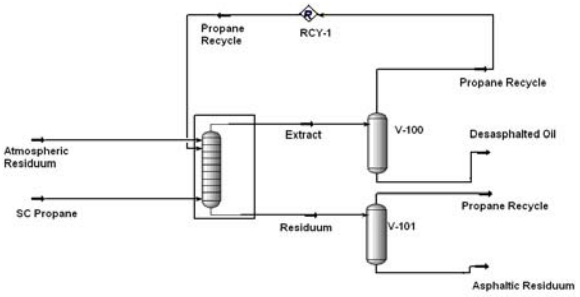
2.2. Solvent Selection

Due to their reasonable costs and favorable values of the critical and transport properties, CO_2, methane, ethane and propane are the most used solvents in the process that use supercritical technology. The selection of CO_2 is linked to some of their characteristics (low toxicity, non-flammable, low critical pressure and temperature and low cost), and mainly to the environmental issues, since in this process, CO_2 generated in the refinery will find a suitable destination and no longer represent a further source of emissions of greenhouse effect gases. In addition, the supercritical technology fits the label of "clean technology" since it leads to the minimization of energy consumption and allows the use solvents with low toxicity.

Cossolvent use can improve the extraction efficiency. Hydrocarbons with low molecular weight solubilize simultaneously the paraffinic and iso-paraffinic chains and promote the precipitation of resins and asphaltenes. As a low molecular weight hydrocarbon propane presents excellent solvency power added to a good selectivity. For these reason, the use of propane as an entrainer is analysed.

2.3. SFE Fluxogram

In the supercritical extraction a substrate is separated from the mixture through a dense pure gas or a mixture of dense gases employed as extraction agent. Downstream the extractor, the targets component are separated from the supercritical solvent by means of depressurization. This technique represents a promising alternative concerning high product quality and recovery. Moreover this technique is considered a clean technology due to the possibility of the minimization of energy consumption in the solvent recovering that is recycled to the process. The Figure 1 presents the fluxogram of the supercritical CO_2 extraction of RAT process. The supercritical solvent feed the SFE column by the bottom. Downstream the extractor, light fraction recovered from RAT is separated from the supercritical CO_2 by means of depressurization. The residuum of the process that can be used as asphaltic residuum leaves the extractor by the bottom outlet.



2.1.1. Figure 1 – Process Fluxogram

3. Results and Discussion

To evaluate the extraction efficiency the parameter representing the relative quantity of the recovered fraction of interest within the original feed was chosen as represented by equation (1). Note that the interest fraction is the sum of the three lighter pseudo-components.

$$\%Light\_Fraction\_\mathrm{Recovered} = \frac{ifm\mathrm{Prod}}{ifmRAT} x100 \quad (1)$$

where:

- *ifmProd* = sum of the mass of the lighter *pseudo-components* (NBP_470, NBP_504 and NBP_542) present in the extract stream.
- *ifmRAT* = sum of the mass of the 3 lighter *pseudo-components* (NBP_470, NBP_504 and NBP_542) present in the Atmospheric Residuum stream.

The efficiency of the separation is related to the amount of the light fraction recovered from RAT. To evaluate the individual influence of the process variables on its efficiency all variables are kept fixed while one is varied according to a scale.
The effect of the presence of small amount of propane is showed in Figure 2. According to the results the action of propane improves the efficiency of the extraction.

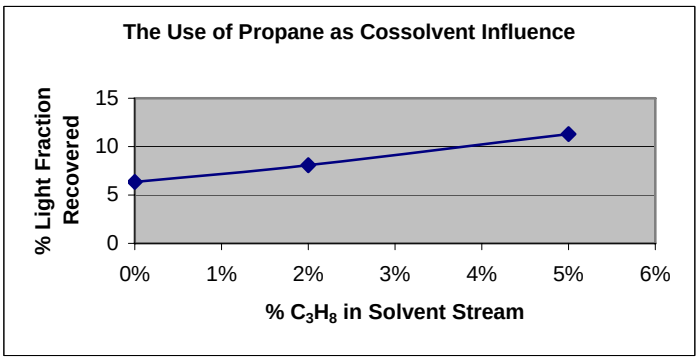


Figure 2 – The Use of Propane as Cossolvent Influence

According to the results presented in given in the Figure 3 the higher the ratio solvent to oil rates the higher the separation efficiency. The composition of solvent is 95 mol% of carbon dioxide and 5 mol% of propane. With respect to the process pressure a direct

dependence was observed since the relative quantity of the light fraction recovered increases with the pressure as shown in Figure 4.

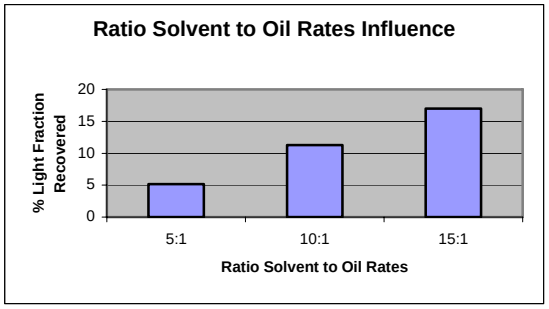


Figure 3 – Ratio Solvent to Oil Rates Influence

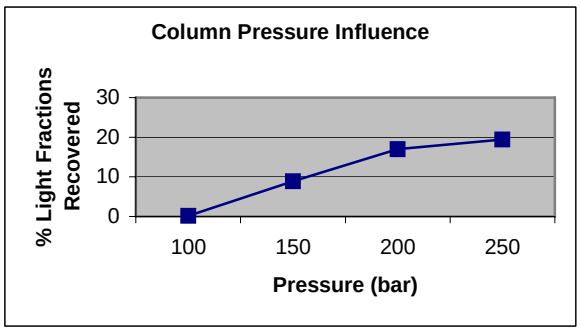


Figure 4 – Column Pressure Influence

With respect to the influence of the solvent feed temperature in the recovery results (Figure 5), the increase in temperature results in the decrease of the relative quantity of the recovered light fraction. This phenomenon can be explained as follows: In an isobaric process the increase of temperature results in the decrease of the solvent density, calculated at 250 bar and 50ºC, 100ºC and 150ºC by quadratic mixing rule of van der Waals type (table V), yielding the reduction of its solvency power and the reduction in the solubility of the solute in the solvent as well.

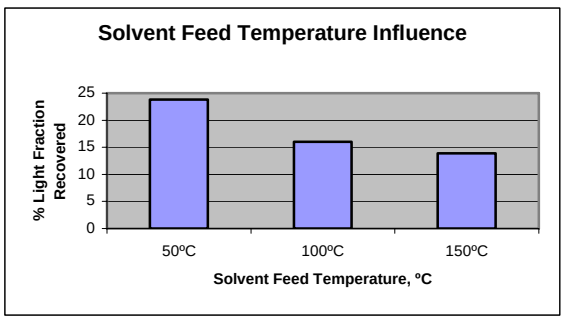


Figure 5 – Solvent Feed Temperature Influence

Table V – Solvent Density

Temperature,ºC	Density, Kg/L
50	0.8134
100	0.5696
150	0.4097

The table VI presents the results of: (i) simulation of the RAT supercritical extraction (solvent composition = 95 mol% CO_2, solvent to oil ratio = 15:1, operational column pressure=250 bar, solvent feed temperature = 50ºC) and (ii) simulation of the RAT vacuum distillation (operational pressure = 0.3 bar, temperature range= 380 to 500ºC). CO_2 in supercritical condition, has the capability to separate the light fraction present in RAT but RAT vacuum distillation showed a best efficiency to recover the light fraction present in RAT.

Table VI – Simulation Results

Stream	**RAT**	**SFC Product**	**Vacuum Dist. Prod.**
NBP-470 Mass Flow (kg/h)	75.2577	33.4999	70.6233
NBP-504 Mass Flow (kg/h)	126.5083	29.3477	36.8513
NBP-542 Mass Flow (kg/h)	111.0476	11.4463	4.8776
NBP-581 Mass Flow (kg/h)	106.7038	4.7932	1.5798
NBP-619 Mass Flow (kg/h)	106.1355	1.9877	0.6015
NBP-657 Mass Flow (kg/h)	102.1138	0.7524	0.2216
NBP-695 Mass Flow (kg/h)	95.0642	0.2673	0.0788
NBP- 736 Mass Flow (kg/h)	97.7244	0.0962	0.0288
NBP-778 Mass Flow (kg/h)	179.4447	0.0595	0.0177
Total Mass Flow (kg/h)	1000	**82.25**	**114.88**

4. Conclusion

In the analysis of the influence of process variables, through simulations has results revealed that: pressure and the ratio between solvent to oil rate presented a direct relation with process efficiency but the solvent feed temperature showed an inverse relation. Despite the fact that all the binary interaction parameters were taken, with the simulator, with the equal value. CO_2, in supercritical condition, has the capability to separate the light fraction present in RAT but the RAT vacuum distillation showed a best efficiency to recover the light fraction present in RAT. It is important to emphasize that further studies should be performed in order to investigate other mixing rules with different binary interaction parameter and solvent mixtures. In this regard, an experimental study in lab scale will be carried out by our group and the results will be compared to those obtained in this work.

References

Hwang J.; Milind D. D.; Hanson F. V.; Dynamic behavior of supercritical fluid extraction of a crude oil and its vacuum residue; Fuel. 1996, 75, 13, 1591-1595.

Zhao S.; Xu Z.; Xu C.; Chung K. H.; Wang R.; Systematic characterization of petroleum residua based on SFEF; Fuel, Vol. 84, nº 6, April 2005, 635-645.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Simultaneous flowsheet optimization and heat integration of a bioethanol processor for PEM fuel cell system

Javier A. Francesconi, Diego G. Oliva, Miguel C. Mussati and Pio A. Aguirre*

INGAR - Instituto de Desarrollo y diseño (CONICET-UTN), Avellaneda 3657, Santa Fe, CP S3002GJC, Argentina
E-mail addresses: {javierf, doliva, mmussati, paguir}@santafe-conicet.gov.ar
**Corresponding author*

Abstract

The conceptual design of a bioethanol processor for fuel cell applications with heat integration was analyzed via model-based optimization formulation. Basically, the system consists of steam reforming of bioethanol and hydrogen purification process that are coupled to a proton electrolyte membrane fuel cell (PEMFC). The system was implemented within General Algebraic Modeling System (GAMS). The model simultaneously handles the problem of optimal heat integration while performing the optimization of process flowsheet. A 50 kW power generation system is presented as case study; the objective was to obtain the operative conditions for the process units that maximize the system efficiency. The operating variables of the system considered as decision variables were: system pressure, water-ethanol molar ratio and reforming temperature, input temperatures to hydrogen purification reactors, fuel cell temperature, and fuel cell hydrogen and oxygen utilizations. The results obtained predict a global efficiency of 43%, about 5% higher than the reference case. These results demonstrate the importance of simultaneous optimal heat integration for fuel cell-based processes.

Keywords: fuel cell system, bioethanol reforming, heat integration, optimization

1. Introduction

In the last years, fuel cells have received a growing attention for power generation owing to their potential higher thermal efficiency and lower CO_2 emissions per unit of power produced. Many research and development projects have been conducted on both the fuel cell itself and the fuel processors for generating hydrogen. There exist several routes for hydrogen production from the primary fuels. Ethanol presents several advantages related to natural availability, storage and handling safety, ethanol can be produced renewably from several biomass sources. Besides the ethanol-to-hydrogen system has the significant advantage of being nearly CO_2 neutral, since the produced carbon dioxide is consumed for biomass growth, thus offering a nearly closed carbon loop (Benito et al., 2005;Francesconi et al., 2007).

Proton exchange membrane fuel cells (PEMFC) for stationary and mobile applications are highly integrated systems including fuel processing, fuel cell itself and post-combustion units (Godat and Marechal, 2003; Xu et al., 2006). Exhaustive heat integration within the system is necessary to achieve acceptable net electrical efficiency levels. Francesconi et al. (2007) examined the heat integration of the ethanol processor by means of a parametric study using simulation process software. The efficiency

reported was of 38%. The aim of this work was to analyze the simultaneous flowsheet optimization and heat integration of an ethanol fuel processor coupled to a polymeric fuel cell by means of mathematical programming techniques. In this work, the material and energy balances of the system were implemented in General Algebraic Modeling System (GAMS). GAMS is a high-level modeling system for mathematical programming problems. The nonlinear programming model of the system was optimized with CONOPT2 solver. The model simultaneously handles the problem of optimal heat integration while performing the optimization of process flow sheets. The NLP formulation proposed by Duran and Grossmann (1986) is used in order to keep maximum heat integration. This allows analyzing the system energy integration by means of the process integration method (or pinch technology), identifying the most successfully heat exchange opportunities and to define the optimal operating conditions of the ethanol processor with the aim of obtaining the best overall efficiency considering the plant balance.

2. Description of the Ethanol Processor

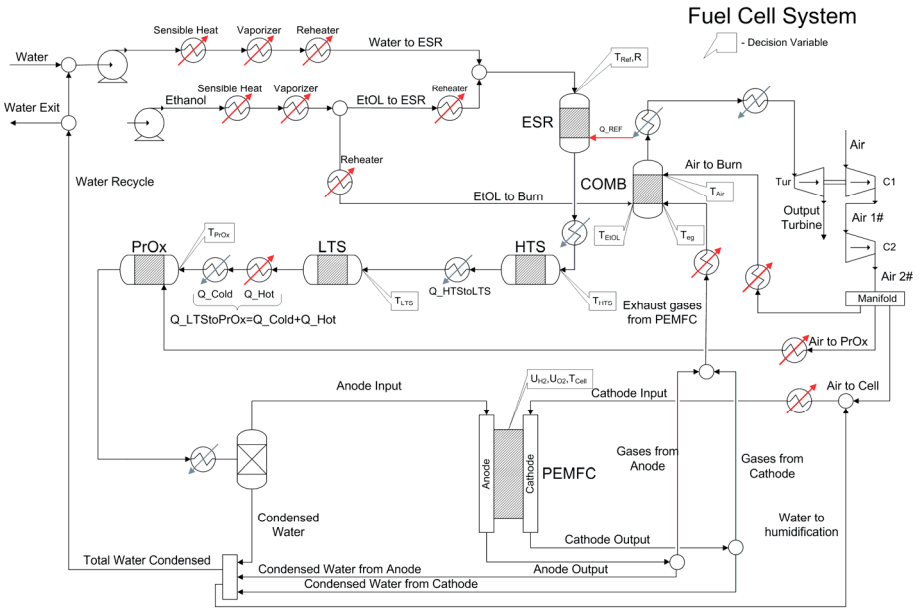


Figure 1. The fuel cell system components and their heat exchange needs.

The Fuel Cell System (FCS) includes a fuel processor which chemically converts ethanol to hydrogen, hydrogen cleanup equipment, a fuel cell stack which electrochemically converts the hydrogen energy to electric power, associated equipment for heat, air and water management, and auxiliary equipment such as pumps and blowers. The fuel processing process has been built according to Francesconi et al. (2007). Figure 1 shows the main components: ethanol steam reformer (ESR), high (HTS) and low temperature (LTS) water gas shift reactors, a preferential oxidation reactor of CO (PrOx), a proton exchange membrane fuel cell (PEMFC), a combustor, compressors and an expander. The major hypothesis regarding the modeling aspect of each unit, and the decision variables considering in the present study are described in the next subsections.

2.1. Ethanol Steam Reforming

This unit was modeled as an equilibrium reactor operating isothermally that involves ethanol steam reforming (feed is composed of fuel and steam). The reformer performs the following reactions according to the mechanism proposed by Benito et al. (2005) for a Co/ZrO_2 catalyst:

Ethanol dehydrogenation: $C_2H_5OH \rightleftarrows CH_3CHO+H_2$ ΔH^0_{298} =71 kJ/mol (1)

Acetaldehyde decomposition: $CH_3CHO \rightleftarrows CH_4+CO$ ΔH^0_{298} =-21.9kJ/mol (2)

Methane steam reforming: $CH_4+H_2O \rightleftarrows CO+3\ H_2$ ΔH^0_{298} =206.1kJ/mol (3)

Water Gas Shift (WGS): $CO+H_2O \rightleftarrows CO_2+H_2$ ΔH^0_{298} =-41.1 kJ/mol (4)

The net energy balance for this complex reaction system is endothermic, and heat has to be supplied from an external energy source in order to maintain the temperature in the reactor. In this study, the reforming temperature (T_{Ref}) and the water to ethanol molar ratio (R) have been considered as decision variables.

2.2. Hydrogen purification section

The CO produced from reforming reactions must be brought down to ppm levels because it gets adsorbed on the noble catalyst of the PEMFC and poisons it. This task is partially accomplished by water gas shift reactors. The water gas shift reaction is carried out in two adiabatic shift reactors in series with an inter-cooler in between to remove the heat of the exothermic reaction. The fist-stage reactor typically operates at 350-500ºC and is called the high-temperature shift (HTS) reactor. The HTS reactor uses a chromiun-promoted iron oxide catalyst. The second stage is a low-temperature shift (LTS) reactor which operates at 150-250ºC, using a copper-zinc catalyst supported on alumina. Performance of these units depends on input concentration of CO and H_2O/CO ratio. Besides, input temperatures affect performance because they modify the output temperature varying the CO conversion.

Chemical equilibrium limits the conversion achieved in the WGS reactor, thus, final CO cleanup occurs in a preferential oxidation (PrOx) unit in which the desired reaction is the oxidation of carbon monoxide (reaction 5). Unfortunately, the selectivity of the catalyst will not avoid the combustion of some hydrogen in the gas stream (reaction 6).

Monoxide oxidation: $CO+O_2 \rightarrow CO_2$ ΔH^0_{298} =-283 kJ/mol (5)

Hydrogen oxidation: $H_2+½O_2 \rightarrow H_2O$ ΔH^0_{298} =-242 kJ/mol (6)

In the model, the CO content at the reactor outlet is fixed at 10ppm. The air flow rate is computed as a function of the CO flow rate assuming two mole of O_2 per mole of CO. Hydrogen combustion is computed considering a complete remotion of the oxygen excess. An adiabatic operation has been considered for the PrOx reactor. The input temperatures of the HTS (T_{HTS}), LTS (T_{LTS}) and PrOx (T_{PrOx}) reactors have been considered as decision variables. Constrains were incorporated according to the operating temperature range of each unit.

2.3. Proton Exchange Membrane Fuel Cell Stack Model

The fuel cell stack was modeled according to Godat and Marechal (2003). The cell voltage (V_{Cell}), cell power (P_{Cell}) and the fuel cell cooling requirement (Q_{Cell}) were computed as follows:

$$V_{Cell} = E^{=}_{Rev}\left(T_{Cell}\right) + \frac{R_g T_{Cell}}{2F}\left[\ln\left(\frac{P^{in}_{H_2,a}+P^{out}_{H_2,a}}{2}\right)+\frac{1}{2}\ln\left(\frac{P^{in}_{O_2,c}+P^{out}_{O_2,c}}{2}\right)\right]-\eta \quad (7)$$

$$P_{Cell} = I_{Cell} V_{Cell} = 2F\left(f^{in}_{H_{2,a}} - f^{out}_{H_{2,a}}\right) V_{Cell} = 2FU_{H_2} f^{in}_{H_{2,a}} V_{Cell} \quad (8)$$

$$Q_{Cell} = \sum_{i=1}^{inlets} f_i h_i\left(T_{Cell}\right) - \sum_{o=1}^{outlets} f_o h_o\left(T_{Cell}\right) - P_{Cell} \quad (9)$$

The PEMFC was supposed to be isothermal and isobar. Here, E^0_{Cell} is the reversible voltage adjusted to the cell temperature (T_{Cell}), and the cell voltage was evaluated averaging the partial pressures of the H_2 and O_2 (arithmetic mean) between the inlet and outlet conditions. The operating voltage was chosen as the power level at which unit cell voltage drops $\eta = 0.5V$ from the reversible voltage. The current (I_{Cell}) is related to the hydrogen molar flow rate at the anode (eq. 8). The cell temperature (T_{Cell}), fuel utilization (U_{H_2}) and oxygen utilization (U_{O_2}) were considered as decision variables. The inlet oxidant to the cathode is humidified to a relative humidity of 80% and the anode inlet stream is always fed at saturated condition. In order to compute the PEMFC fuel cell cooling, an energy balance was made over the cell from the inlet to the exit conditions (eq. 9).

2.4. Post-Combustion System

The depleted fuel of the PEMFC, formed by cathode and anode outlets previous separation of condensate water, is burn in the post combustion system. The generated heat is used to balance the energy requirement of the fuel processing section. Supplementary firing of ethanol is considered if the energy content of the depleted fuel is not sufficient to satisfy the balance. Complete and stoichiometric combustion for the burner unit has been assumed. The preheating temperatures of the ethanol to burn (T_{EtOL}), exhaust gases from the fuel cell (T_{eg}), and the air supply to the combustor (T_{air}) were considered as decision variables.

2.5. Objective function

The objective is to maximize the net efficiency of the system. The FCS global efficiency (φ) is defined as the net energy output of the system obtained from the gross output by subtracting the electrical energy needed to operate FCS auxiliaries such as pumps and compressors divided by the lower heating value (LHV) of the ethanol consumed in the fuel processor for reforming and burning (f_{EtOL}).

$$\varphi = \frac{P_{Cell} - \left(P_{Pumps} + P_{Comp} - P_{Tur} + Q_C/25 + Q_H/0.95\right)}{\left(f_{EtOL} \times LHV\right)} \quad (10)$$

Q_C is the heat removal from the fuel cell system; it represents the cold utility of the process. We considered a cooling system efficiency of 25 (Larminie and Dicks, 2000). The heat utility (Q_H) is represented by an electrical heater with an efficiency of 0.95.

2.6. Optimization Problem and Heat Exchange Model

The heat exchange network is not considered directly in the model; instead, pinch technology can be used to model the integrated heat exchange system without having to impose a heat exchange network structure. The method is based on Duran and Grossman (1986) approach, it comprises including a set of constraints into the nonlinear process optimization problem so as to insure that the minimum utility target for heat recovery networks is featured. The proposed problem of this optimization with heat integration is given by the following model:

$$\begin{aligned} \min \varphi \quad &= \quad f(x, \mathrm{Q_C}, \mathrm{Q_H}) \\ s.t. \quad & \quad h(x) = 0 \\ & \quad g(x) \le 0 \end{aligned} \tag{11}$$

$$\mathrm{Q_H} + \mathrm{Q_C} + \sum_{i=1}^{n_H} \mathrm{Q_i} + \sum_{j=1}^{n_C} \mathrm{Q_j} = 0 \tag{12}$$

$$\begin{aligned} &\sum_{i=1}^{n_H} F_i Cp_i \left[\max\left\{0, T_i^{in} - T^p\right\} - \max\left\{0, T_i^{out} - T^p\right\} \right] \\ &-\sum_{j=1}^{n_C} f_j Cp_j \left[\max\left\{0, t_j^{out} - \left(T^p - \Delta T_{\min}\right)\right\} - \max\left\{0, t_j^{in} - \left(T^p - \Delta T_{\min}\right)\right\} \right] \le \mathrm{Q_H} \end{aligned} \qquad \forall p \in \Pi \tag{13}$$

where φ is the objective function, Q_H is the requirement of the hot utility, Q_C is the requirement of the cold utility, $h(x)$ is the flow sheet model, $g(x)$ represents the inequality constraints, j is the index for cold streams in the flow sheet, i is the index for hot streams, t_j is the temperature of cold stream j, T_i is the temperature of hot stream i, $F_j Cp_i$ and $f_j Cp_j$ are the heat capacity flow rates of the hot an cold streams, respectively. And the set Π defines the pinch candidate temperatures as follows:

$$T^p = \begin{cases} T_i^{in} & \text{if candidate p is hot stream i} \\ t_j^{in} + \Delta T_{\min} & \text{if candidate p is cold stream j} \end{cases} \tag{14}$$

In this representation the hot duties are considered negatives ($\mathrm{Q_H}, \mathrm{Q_i} \ge 0$), and cold duties positives ($\mathrm{Q_C}, \mathrm{Q_j} \ge 0$). In order to avoid difficulties with the NLP solver the smoothing approximation for the max operator cited in Xu et al. (2006), was utilized.
This formulation requires defining, a priori, the cold and hot streams. However, as the flows, compositions and temperatures of the streams are variables it is possible that some stream initially considered hot, could be cold in the optimized case. This behaviour can be handled considering that a specific energy duty is the sum of a cold and a hot stream. In other words, for modeling a given energy duty both possibilities are considered concurrent by means of one cold y and one hot stream, in the optimized case one of them will be null and its energetic requirement will be zero. In the system under study, conditioning the stream incoming at the PrOx reactor can present this situation, and only this stage requires this double representation (see Figure 1). With this approach, was possible to compute the influence of the decision variables so that the best system configuration can be then defined.

3. Results and discussion

The system was optimized considering a net power target of 50kW. It is intended to study the influence of the following operating variables on the efficiency of the integrated system: operating conditions of the reformer (T_{Ref} and R), effect of input temperatures of hydrogen purification reactors (T_{HTS}, T_{LTS} and T_{PrOx}), combustion preheating (T_{EtOL},T_{Air},T_{eg}), and operating conditions of the fuel cell stack (T_{Cell}, U_{H2} and U_{O2}). In addition, the pressure of the system (P) was considered as a decision variable, nevertheless pressure drops have been neglected and all streams were defined at the same operating pressure. Table 1 shows the values of the decision variables obtained with its respective upper and lower bounds. The results are compared with the data reported by Francesconi et al. (2007). The global efficiency was increased achieving a 43%. As expected, the hot requirement of the system is satisfied by hot gases post-

combustion, and the hot utility is null. The cool utility achieves the value of 60.5 kJ/h. It is worthy to note, that in the reference case the stream between LTS output and PrOx input represent a cold stream with a duty of 34.4 kJ/h, while at the optimized case this represents a hot stream with a value of -5486 kJ/h.

Table 1. Results obtained from the model-based optimization.

	Reference Case (Francesconi et al., 2007)	present work Value (Lower Bound-Upper Bound)	
Efficiency (%)	38.0	43.0	
P (atm)	3.0	2.2	(1.5-5)
R	4.0	3.0	(2-6)
T_{Ref} (ºC)	709.0	714.3	(550-900)
T_{HTS} (ºC)	350.0	365.8	(350-550)
T_{LTS} (ºC)	150.0	163.6	(150-250)
T_{PrOx} (ºC)	237.0	178.6	(150-300)
T_{Cell} (ºC)	80.0	60.0	(60-95)
U_{H2} (%)	80.0	95.0	(50-95)
U_{O2} (%)	50.0	80.2	(25-95)
T_{EtOL} (ºC)	300.0	347.5	(200-400)
T_{eg} (ºC)	500.0	800.0	(200-800)

4. Conclusions

The development of an ethanol processor model within a modeling system as GAMS allowed the simultaneous flowsheet optimization and heat integration of the process. New operating conditions were obtained increasing the global efficiency of the fuel cell system about 5% higher than the reference case. Nevertheless due to the nonconvex characteristic of the NLP model, it is not possible to guarantee a global optimum, which will be addressed in further works. The results demonstrate the importance of optimal heat integration and the mathematical programming techniques for fuel cell-based processes.

References

M. Benito, J.L. Sanz, R. Isabel, R. Padilla, R. Arjona, L. Daza, 2005, Bio-ethanol steam reforming: Insights on the mechanism for hydrogen production, J. Power Sources,151,11-17.

M. A. Duran, I.E. Grossmann, 1986,Simultaneous optimization and heat integration of chemical processes, AIChE J.,32,123-138.

J.A. Francesconi, M.C. Mussati, R.O. Mato, P.A. Aguirre, 2007, Analysis of the energy efficiency of an integrated ethanol processor for PEM cuel cell systems, J. Power Sources, 167, 15-161.

J. Godat, F. Marechal, 2003, Optimization of a fuel cell system using process integration techniques, J. Power Sources, 118, 411-423.

J. Larminie, A. Dicks, 2000, Fuel Cell Systems Explained, p247-248.

C. Xu, L.T. Biegler, M.S. Jhon, 2006, Systematic optimization of an H_2 PEM fuel cell power generation system with heat integration, AIChE J., 52, 2496-2506.

Acknowledgements

The authors want to thank for the financial support from CONICET and ANPCYT from Argentina.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Reliability Modeling of a Natural Gas Recovery Plant using *q*-Weibull Distribution

Isabel Sartori[a], Edilson M. de Assis[a], Adilton L. da Silva[b], Rosana L. F. Vieira de Melo[a], Ernesto P. Borges[b] e Silvio A. B. Vieira de Melo[a,*]

[a] *Programa de Engenharia Industrial - PEI, Escola Politécnica, Universidade Federal da Bahia – UFBA, R. Aristides Novis, n. 2, Federação, Salvador - BA, 40210-630, Brasil*
[b] *Escola Politécnica, Universidade Federal da Bahia – UFBA, R. Aristides Novis, n. 2, Federação, Salvador - BA, 40210-630, Brasil*
[*] *E-mail: sabvm@ufba.br (corresponding author)*

Abstract

We apply a recent generalization of Weibull distribution within the context of reliability engineering, inspired on mathematical developments that followed nonextensive statistical mechanics. Weibull distribution, largely used in reliability engineering, has been generalized to a q-Weibull distribution, following the lines of the q-exponential function, that appear within the context of nonextensve statistical mechanics. This new distribution is used to describe failure rate of a compression unit in a typical natural gas recovery plant currently in operation, based on time to failure data from historic process knowledge. The results show that the q-Weibull fits better the available data and this new distribution is very promising.

Keywords: reliability, natural gas, q-Weibull.

1. Introduction

A growing focus has been placed on Reliability, Availability, Maintainability and Safety (RAMS) during the design and operation of industrial systems, mainly due to the size and complexity of modern industrial plants, including oil and gas processing. Safety operation of natural gas production and recovery facilities depends on the continuous improvement of their reliability. A comprehensive assessment of operational safety requires a systemic approach based on statistical models for description of failure rates related to the equipments and their components. Development, choice or even application of a model to accurately characterize the failure rate is a non trivial task and mathematical simulation of reliability performance depends crucially on it. However, life data are still scarce and a predictive capability of the model should be often investigated. Reliability modeling, that is one of the most important steps for RAMS assessment, plays a key role to deal with incomplete life data..

There are many different statistical distributions that can be used to model life data. However, for some systems the classical life distributions are not satisfactory. Along the last two decades there has been a continuous and increasing development on what has been known as nonextensive statistical mechanics. The 1988 seminal paper by Tsallis (Tsallis, 1988) has introduced a generalization of the concept of entropy, by means of a real parameter q:

$$S_q = k \frac{\sum_i^W p_i^q - 1}{1-q} \tag{1}$$

where k is a positive constant that gives dimensional consistency to the expression (Boltzmann's constant), p_i is the probability of occupancy of the i-th microstate, and W is the total number of microstates of the system. This expression recovers the celebrated Boltzmann-Gibbs-Shannon (BGS) entropy, $S_1 = -k\sum_i^W p_i \ln p_i$ when $q \to 1$.

Nonextensive statistical mechanics has induced generalizations in other fields, e.g., in mathematics. A pair of functions naturally appears from the very beginning of the formalism, namely the q-logarithm, and its inverse, the q-exponential, defined as (Tsallis, 1994 and Borges, 1998)

$$\exp_q x = \left[1 + (1-q)x\right]_+^{1/(1-q)} \tag{2}$$

where $[x]_+ = \max\{x, 0\}$ and $q \in \mathbb{R}$. The q-exponential has many interesting properties see, for instance, Yamano (2002), as $\lim_{q\to 1} \exp_q x = e^x$, $\exp_q 0 = 1$ and

$$\int \exp_q x \, dx = \frac{1}{2-q} \exp_{1/(2-q)}\left[(2-q)x\right]. \tag{3}$$

Generalized statistical distributions developed within nonextensive statistical mechanics have been found applications in several physical, social and artificial complex systems (Gell-Mann and Tsallis, 2003). This q-type distributions (Nadarajah and Kotz, 2006, 2007) usually better fit data when compared with their equivalent "classical" versions (Picoli et al., 2003; de Souza and Tsallis, 1997), particularly if the tails are not well adjusted by Weibull distribution. Time-to-failure data has already been described by one of these generalizations, the q-Weibull distribution, by Costa et al. (2007), in the analysis of dielectric breakdown of ultra-thin oxides.

In this work we consider *q*-Weibull distribution to describe failure rate of a compression unit in a typical natural gas recovery plant currently in operation, based on time-to-failure data from historic process knowledge.

2. Reliability Modeling

Weibull distribution is largely used in reliability context. The probability density function for time to failure of the Weibull distribution in three-parameter form is defined by (Weibull, 1951)

$$f(t) = \frac{\beta(t-t_0)^{\beta-1}}{(\eta - t_0)^{\beta}} \exp\left(-\left(\frac{t-t_0}{\eta - t_0}\right)^{\beta}\right) \text{ with } t \geq t_0, \tag{4}$$

where t is time to failure or life time, β is the shape parameter, η is the characteristic life and t_0 is the location parameter.

Using the definition of the q-exponential (Eq. 2) for the generalization of Weibull distribution (Eq. 4), the q-Weibull probability density function can be written as

$$f_q(t) = \frac{\beta(2-q)}{\eta - t_0}\left(\frac{t-t_0}{\eta - t_0}\right)^{\beta-1} \exp_q\left(-\left(\frac{t-t_0}{\eta - t_0}\right)^{\beta}\right) \text{ with } t \geq t_0 . \tag{5}$$

The reliability function is $R_q(t) = \int_t^{\infty} f_q(x)dx$, i.e. (See Eq. 3),

$$R_q(t) = \exp_{\frac{1}{2-q}}\left(-(2-q)\left(\frac{t-t_0}{\eta - t_0}\right)^{\beta}\right) \text{with } t \geq t_0 . \tag{6}$$

Sample data are time-to-failure rank in ascending order and an estimate of the unreliability can be obtained using an approximation of the median ranks, also known as Benard's approximation, given by (Johnson, 1951)

$$\hat{F}_i = \frac{i-0.3}{n+0.4}, \tag{7}$$

where n is the sample size, i is the order number of failure varying from 1 to n. In this way, for every sampling time t_i we obtain

$$x_i = \ln(t_i - t_0) \tag{8}$$

$$y_i = \ln\left(-\ln_{\frac{1}{2-q}}\left(1-\hat{F}_i\right)\right). \tag{9}$$

This change of variables permits to estimate η parameter by least squares method for the linear correlation between x_i and y_i

$$\eta = \exp\left(\frac{-b}{\beta}\right)(2-q)^{\frac{1}{\beta}} + t_0 . \tag{10}$$

Then, in this way, the q-Weibull distribution parameters are calculated searching the parameters q and t_0 that return the maximum value of R^2.

3. Results

Time to failure data from a compression unit of natural gas are gathered from historic industrial plant operation.

Results of q-Weibull fit to these data are evaluated through R^2 maximization ($R^2 = 0,986$ in Fig. 1) and the following parameters are found: $t_0 = 1.23h$, $q = 0.056$, $\beta = 0,502$, $\eta = 5576h$. Fig. 2 shows the data calculated with Eqs. 8 and 9 fitted to q-Weibull

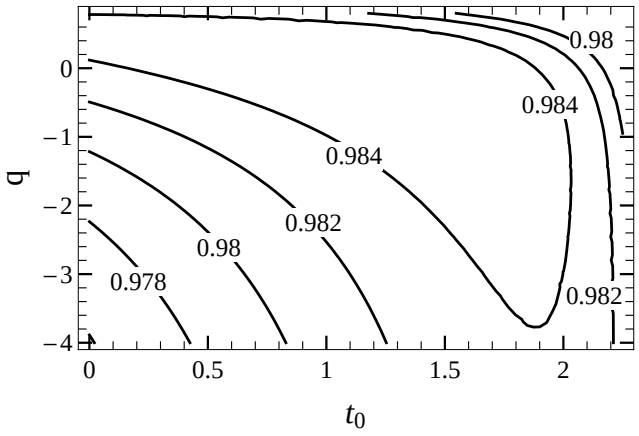


Fig.1 – Square correlation coefficient R^2 as a function of q and t_0. Lines indicate constant R^2, exhibiting the maximum in the 2-D surface.

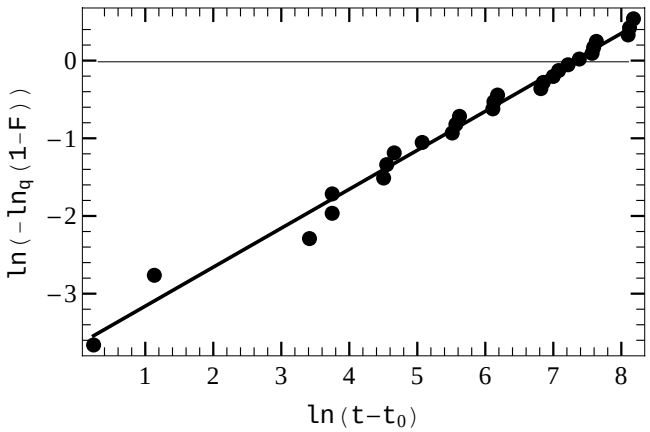


Fig. 2 – Failure probability as a function of shifted time-to-failure, properly scaled, in order to linearize the data. Square correlation coefficient is $R^2 = 0.986$.

For the classical Weibull distribution (Fig. 3) the following parameters are found: $t_0 = 0h$, $R^2 = 0,981$, $\beta = 0,609$ and $\eta = 826h$. One can observe that the values of η are quite different for both models and q-Weibull presents a bit higher value of R^2 coefficient. q-Weibull usually fits better than conventional Weibull if data are not adjusted well in the tail. This occurs because the q parameter unties the q-Weibull distribution of the exponential relation and the tail behaviour is different from conventional Weibull. Comparing the last three points in Fig.2 and 3, it can be seen that they are better fitted to the model in Fig. 2 where q-Weibull distribution is used.

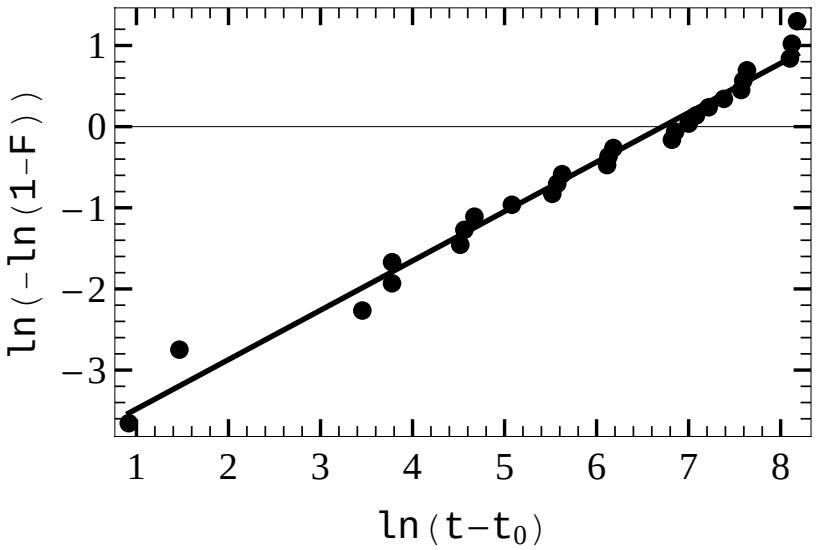


Fig. 3 – Failure probability as a function of shifted time-to-failure, properly scaled, in order to linearize the data. Square correlation coefficient is R^2 = 0.981 and parameter q =1.

4. Conclusion

In this work a comparison is done between the Weibull classical distribution and the q-Weibull generalized distributions in the context of reliability engineering. Both distributions are applied to describe life data of a compression unit in a typical natural gas recovery plant. The results show that the q-Weibull distribution fits better to the available data and is very promising for reliability modeling of industrial cases. Further research is in progress to extend this methodology to other models and applications for reliability systems engineering.

References

E. P. Borges, 1998 , On a *q*-generalization of circular and hyperbolic functions, J. Phys. A: Math. Gen. 31, 5281-5288

U.M.S. Costa, V.N. Freire, L.C. Malacarne, R.S. Mendes, S. Picoli Jr., E.A. de Vasconcelos and E.F. da Silva Jr., 2006, An improved description of the dielectric breakdown in oxides based on a generalized Weibull distribution, Physica A, 361, 209-215

M. Gell-Mann and C. Tsallis (Eds.), 2003; Nonextensive Entropy – Interdisciplinary Applications. Oxford University Press, Oxford

L.G. Johnson, 1951, The Median Ranks of Sample Values in Their Population with an Application to Certain Fatigue Studies, Industry Mathematics, 2

S. Nadarajah and S. Kotz,2006, q -exponential is a Burr distribution, Phys. Lett. A 359, 577-579

S. Nadarajah and S. Kotz, 2007, On the *q*-type distributions, Physica A, 377, 465-468

S. Picoli Jr., R.S. Mendes and L.C. Malacarne, 2003, *q*-exponential, Weibull, and *q*-Weibull distributions: an empirical analysis, Physica A, 324, 678-688

A.M.C. de Souza and C. Tsallis, 1997, Students *t*- and *r*-distributions: Unified derivation from an entropic variational principle, Physica A 236, 52-57

C. Tsallis, 1988, Possible generalization of Boltzmann-Gibbs statistics, J. Stat. Phys., 52, 479

C. Tsallis, 1994, What are the numbers that experiments provide?, Química Nova, 17, 468-471

W. Weibull, 1951, A Statistical Distribution Function of Wide Applicability, Journal of Applied Mechanics-Transactions of the ASME, 18, 293-297

T. Yamano, 2002, Some properties of *q*-logarithm and *q*-exponential functions in Tsallis statistics, Physica A 305, 486-496

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Simulation of Process Interesterification in Fluidized Bed Bioreactor for Production of Biodiesel

Jocélia S. Mendes, Jouciane S. Silva, Andrea L. O. Ferreira, Giovanilton F. Silva

Department of Chemical Engineering, UFC-Universidade Federa do Ceará
Campus do Pici, Bloco 709, Pici, 60455-760 – Ceará – Brazil

Abstract

The aim this work was investigated a new route for biodiesel production using immobilized lipase from *Candida antarctica* in continuous Fluidized Bed Bioreactor. Conventionally, Biodiesel (fatty acid methyl esters) is produced by transesterification in which, oil or fat is reacted with a monohydric alcohol in presence of a catalyst. In recent years, the use of lipases as biocatalysts for biodiesel production has become of great interest due to its environment friendly. But some alcohols such as methanol inactivated the lipases to some extent and the enzymatic stability was poor. In order to enhance the stability of the lipase, three-step methanolysis was adopted, however, glycerol, as one of the products was easy to adsorb on the surface of lipase resulting in serious negative effect on the enzymatic activity. For to solve problems was used the interesterification kinetics of triglycerides and methyl acetate for biodiesel production was modeled. A heterogeneous model describing the interesterification process in an immobilized enzyme fluidized-bed bioreactor was developed. A simplified model based on Ping Pong Bi Bi with substrate competitive inhibition mechanism was proposed to describe the reaction kinetics of the interesterification. The model without any adjustable parameters was used to predict the interesterification process. The key parameters which measured the extent of external and internal mass-transport resistances, as well as the degree of back-mixing were quantified and discussed. The fluidized-bed bioreactor considered in this investigation is composed of two phases: a fluid phase comprised mainly of the triglycerides and methyl acetate and the product (Biodiesel); and a solid phase which is the immobilized enzyme. The effects of some operating and design parameters on the performance of the fluidized-bed bioreactor were also analyzed. The model was also tested for its sensitivity to changes in hydrodynamic parameters.
Keywords: biodiesel, lipase from *Candida antarctica*, interesterification, fluidized bed bioreactor.

1. Introduction

Biodiesel is mono-alkyl esters of long chain fatty acids derived from renewable vegetable oils and animal fats. Biodiesel (fatty acid methyl esters) is produced by transesterification in which, oil or fat is reacted with a monohydric alcohol in presence of a catalyst. In the transesterification of vegetable oils, a triglyceride reacts with an alcohol in the presence of a strong acid or base, producing a mixture of fatty acids alkyl esters and glycerol. The overall process is a sequence of three consecutive and reversible reactions, in which monoglycerides are formed as intermediates. The stoichiometric reaction requires 1 mol of a triglyceride and 3 mol of the alcohol.

However, an excess of the alcohol is used to increase the yields of the alkyl esters and to allow its phase separation from the glycerol formed.
Biodiesel has become more attractive recently because of its environmental benefits and the fact that it is made from renewable resources (Schuchardt et al., 1998). The transesterification of vegetable oils with methanol or ethanol as well as the main uses of the fatty acid methyl esters are studied in this paper. The general aspects of this process and the applicability of different types of catalysts (acids, alkaline metal hydroxides, alkoxides and carbonates, enzymes and non-ionic bases, such as amines, amidines, guanidines and triamino(imino)phosphoranes) (Jordan and Gutsche, 2001; Fangrui and Hanna, 1998; Filippis et al., 1995; Antolín et al., 2002) are described, and ambient or elevated pressures and temperatures.
Conventionally, biodiesel was produced by transesterification of triglycerides and alcohols in the presence of an acid or an alkaline catalyst. In recent years, the use of lipases as biocatalysts for biodiesel production has become of great interest due to its environment friendly. However, some alcohols such as methanol inactivated the lipases to some extent and the enzymatic stability was poor. In order to enhance the stability of the lipase, three-step methanolysis was adopted, however, glycerol, as one of the products was easy to adsorb on the surface of lipase resulting in serious negative effect on the enzymatic activity Xu et al (2003), Dossat et la (1999) and Du et al (2004).
In order to solve the above-mentioned problems, we previously reported that using methyl acetate as acyl acceptor instead of methanol for biodiesel production could enhance the stability of the lipase significantly and in the process, triacetylglycerol instead of glycerol would be produced and it has been demonstrated that triacetylglycerol had no negative effect on the activity of the lipase. Moreover, triacetylglycerol was an important by-product with a higher value than glycerol and this novel route was thought to be very promising for large scale production of biodiesel.

2. Model Development

2.1. Kinetic Model

The enzymatic model used to describe the interesterification kinetics was based on the model Ping Pong Bi Bi with competitive inhibition for the substratum. The equation of that mechanism is shown:

$$V_i = \frac{V_{\max}[TG][AM]}{K_{M_{TG}}[AM]\left(1+\frac{[AM]}{K_I}\right)+K_{M_{AM}}[TG]+[TG][AM]} \tag{2.1}$$

where V_i was the initial reaction rate; [TG] and [AM] the initial molar concentrations of triglycerides and methyl acetate, respectively; K_{mTG} and K_{mA} the apparent Michaelis constants for triglycerides and methyl acetate, respectively; K_i the apparent inhibition constant of methyl acetate and Vmax the initial maximum velocity of the reaction. The kinetic constants are shown in

Table 1 The kinetic constants, Xu et al (2005).

Parameters	Value
V_{max} (mol/lmin)	1.9
K_{mTG} (mol/l)	1.0
K_{mA} (mol/l)	16
K_i (mol/l)	0.0455

2.2. Solid Phase Mass Balance

The fluidized-bed bioreactor considered in this investigation is composed of two phases: a fluid phase comprised mainly of the substrate (oil and methyl acetate) and the product (Biodiesel); and a solid phase which is the immobilized enzyme. The following assumptions are employed in the model: (1) the system is isothermal; (2) the movement of reactant within the biocatalyst can be described mathematically by Fick's law of diffusion where the effective diffusion coefficient is constant and independent of concentration; (3) the enzyme activity is uniform throughout the particle; (4) the fluid phase back-mixing can be quantified by an axial dispersion coefficient.

The general mass balance equation governing the concentration distribution in the fluid phase of the biocatalyst is given by:

$$D_L \frac{d^2 C_f}{dZ^2} - U_0 \frac{dC_f}{dZ} - (1-\varepsilon) r_V \tag{2.2}$$

and the Danckwert's boundary conditions are:

$$when\ Z = 0,\ -D_L \frac{dC_f}{dZ} + UC_f = UC_f \tag{2.3}$$

$$when\ Z = L,\ \frac{dC_{x,P,S}}{dZ} = 0 \tag{2.4}$$

This set of boundary conditions, that is equations 7 and 8, is widely used in modeling immobilized enzyme fluidized-bed bioreactors.

2.3. Solid Phase Mass Balance

The general mass balance equation governing the concentration distribution in the fluid phase of the biocatalyst is given by equation 2.5.

$$\frac{dC_S}{dZ} = k_m (C_f - C_S) \tag{2.5}$$

with the boundary conditions given by:

$$when\ Z = 0,\ C_S = 0 \tag{2.6}$$

2.4. Evaluation of the relevant parameters

The relevant parameters relating to the dispersion coefficient, external mass-transfer resistance and bed voidage are evaluated using correlations obtained from the literature and are given below.

2.4.1. Dispersion coefficient

The dispersion coefficient, D_L, in the fluidized-bed containing light, non-porous, solid particles was evaluated using the correlation of Tang and Fan (1990) given by Equation (2.7) below.

$$D_L = \frac{U_0 L}{Pe} \tag{2.7}$$

$$\frac{1}{Pe} = 4.35 \left(\frac{\rho_S}{\rho_f} \right)^{2.64} \varepsilon^{1.467} \tag{2.8}$$

The external particle-liquid mass-transfer coefficient was obtained using the correlation of Nore et al. (1992) given by:

$$k_m = 1.1 \left(\frac{U_0}{\varepsilon} \right)^{0.43} D_p^{0.24} \tag{2.9}$$

The bed porosity in the bed is determined by the mechanics of fluidization. Hence, a realistic mathematical expression for the bed fluidization is necessary. For bed of uniform spherical particles, the following relation was proposed. The Richardson and Zaki (1954) relate bed voidage to the up-flow superficial liquid velocity given as:

$$\varepsilon = \left(\frac{U_0}{U_t}\right)^{1/n} \tag{2.10}$$

The terminal velocity given by Equation (2.11) was obtained using the correlation of Khan and Richardson (1990).

$$U_t = \frac{\mu}{\rho_f D_p}\left(2.33Ga^{0.018} - 1.53Ga^{-0.016}\right)^{13.3} \tag{2.11}$$

where

$$Ga = \frac{\left(\rho_s - \rho_f\right)\rho_f D_p^3 g}{\mu^2} \tag{2.12}$$

The fractional conversion, Y, shown in Fig. 1 against the flow rate is defined by Equation (2.13).

$$Y = \frac{C_0 - C_L}{V_{\max}\, L/u} = \frac{X}{\theta} \tag{2.13}$$

where θ dimensionless residence time.

3. Results and Discussion

The modeling equations presented here constitute a system of non-linear boundary value problems which were solved by finite difference techniques. The details of solution methodology are omitted here for the sake of brevity.

Figure 1 shows the concentration of each composition triglycerides, methyl acetate and biodiesel against the height of the fluidized bed. From the Figure 2, it can be observed that, the model indicate the high concentration of biodiesel were reached in the end of the reactor. Indicating that the dispersion has influences in the interesterification of triglycerides for biodiesel production with methyl acetate.

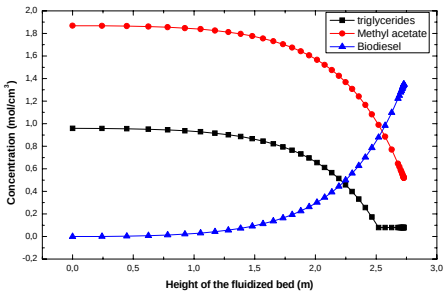


Figure 1 – The concentration of triglycerides, methyl acetate and biodiesel.

The Figure 2 show the conversions obtained for the model for fluidized bed in function of the dimensionless residence time. It can be observed that the measure that the dimensionless residence time increases the values of the conversions increased. Therefore, they are high necessary times of residence for the reactor to reach high conversions.

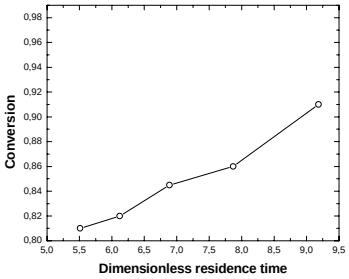


Figure 2 – Conversion in fluidized bed bioreactor

Another parameter of common interest in the design of fluidized bed reactors is the Peclet number which is the measure of the degree of back-mixing in the flow vessel. As the Peclet number approaches infinity plug-flow behavior is approached, while complete mixing is approached for a Peclet number approaching zero. It is obvious from the plot of Peclet number against the dimensionless residence time. The according Equations 2.7 and 2.8 when the value the Peclet increases the dispersion coefficient value decrease. This is also indicative of dominance of mixing over segregation. Thus, the conversion increases when the value de Peclet increases.

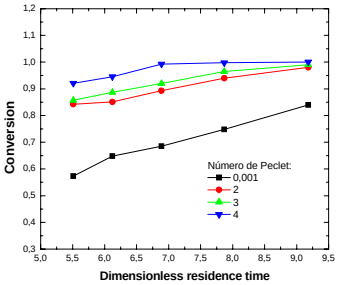


Figure 3 – Effect of number of Peclet on the conversion de the fluidized bed.

4. Conclusions

In this work a continuous fluidized bed bioreactor model with Immobilized Candida Antarctica in continuous Fluidized Bed Bioreactor was developed for simulating the steady-state performance of a bioreactor for production biodiesel. It was examined effect Peclet number and dimensionless residence time at conversion in fluidized bed. It the Peclet number decreases with the increase conversion production of biodiesel. It can be observed that conversion increased when the dimensionless residence time increased. The model can be used to simulate the biodiesel production by interesterification in fluidized bed bioreactor and the results of the simulation were satisfactory.

List of symbols

C_f	fluid phase concentration, mol/l	P_e	Peclet number, dimensionless
C_S	solid phase concentration, mol/l	u	flow rate
D_L	dispersion coefficient, m^2/s	U_0	liquid superficial velocity, m/s
D_P	particle diameter, m	U_t	terminal velocity of particle
g	gravitational acceleration, m/s	ε	porosity of the bed, dimensionless
Ga	Galileo's number, dimensionless	ρ_f	density of the fluid, kg/m^3

k_m	mass-transfer coefficient	ρ_s	mean particle density, kg/m^3
K_m	Michaelis constant	μ	viscosity, kg/m.s
L	reactor length, m	Y	Conversion
Subscripts			
P	production concentration		
S	substrate concentration		
x	biomass concentration		

References

Antolín, G., Tinaut, F. V., Briceño, Y. et alii (2002). Optimisation of biodiesel production by sunflower oil transesterification. Bioresource Technology, 83, 111-121.

Dossat, V.; Combes, D.; Marty, A. (1999) Lipase-catalysed transesterification of high oleic sunflower oil, Enz. Microb. Technol., 30 (1), 90-94.

Du, W.; Xu, Y. Y.; Liu, D. H.; Zeng, J. (2004) Comparative study on lipase-catalyzed transformation of soybean oil for biodiesel production with different acyl acceptors. Journal Molecular. Catalysis . B: Enzyme., 30 (3-4), 125-129.

Khan, A. R. & Richardson, J. F. (1990). Pressure gradient and friction factor for sedimentation and fluidization of uniform spheres in liquidg Chem. Engng. Sci., 45, 255-265.

Fangrui, M., Hanna, M. A. (1998). Biodiesel production: a review. Bioresource Tecnology, 70, 1-9.

Filippis, C. G. Scarsella, M., Sorrentino, M. (1995). Transesterification process for vegetable oils: A simple control method of methyl ester content. Journal American Oils Chemistry Society, 72, 1399-1407-1415.

Jordan, V., Gutsche, B. (2001). Development of environmentalyy benign processes for the production of fatty acid methyl esters. Chemosphere, 43, 99-108.

Nore, O., Briens, C., Margarilis, A. & Wild, G. (1992). Hydrodynamics, gas-liquid mass transfer and particle liquid heat and mass transfer in a three-phase fluidized bed for biochemical process application. Chem. Engng. Sci., 47, 3573-3580.

Richardson, J. F. & Zaki, W. N. (1954). Sedimentation and fluidization: Part I~ Trans. Inst. Chem. Engrs., 32, 35-53.

Schuchardt, U., Sercheli, R., Matheus, V. R. (1998). Transesterification of vegetable oils: a Review. Journal of the Brazilian Chemical Society, 9, 199.

Tang, W. T. & Fan, L. S. (1990). Axial liquid mixing in liquid-solid and gas-liquid fluidized-beds containing loq density particles. Chem. Engng. Sci., 45, 543-551.

Xu, Y. Y.; Du, W.; Liu, D. H.; Zeng J. (2003), A novel enzymatic route for biodiesel production from renewable oils in a solvent-free medium. Biotechnology letters, 25 (15), 1239-1241.Xu, Y. Y.; Du, W.; Liu, D. H. (2005), Study on the kinetic of enzymatic interesterification of triglycerides for biodiesel production with methyl acetate as the acceptor. Journal Molecular. Catalysis . B: Enzyme., 32, 241-245.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Kinetic Study of Biodiesel Production by Enzymatic Transesterification of Vegetable Oils

Fernando L.P. Pessoa[a], Shayane P. Magalhães[a] and Pedro W.C. Falcão[b]

[a]*DEQ - Department of Chemical Engineering, Chemistry School – Federal University of Rio de Janeiro, Av. Athos da Silveira Ramos, 149, Centro de Tecnologia, Bl. E, s. 201, Cidade Universitária, Ilha do Fundão, CEP 21941-909, Rio de Janeiro – RJ, Brazil*

[b]*PEQ/COPPE– Federal University of Rio de Janeiro, Av. Athos da Silveira Ramos, 149, Centro de Tecnologia, Bl. E, s. 201, Cidade Universitária, Ilha do Fundão, CEP 21941-909, Rio de Janeiro – RJ, Brazil*

Abstract

Although acid and alkaline catalysts are mostly used commercially for the synthesis of biodiesel, the difficulties related with the glycerol recovery, catalyst removal and the purification steps stimulate several studies of the enzymatic transesterification of oils since through this rout such problems are inexistent. Even though the high cost of the enzyme is still a drawback for the use of this process in industrial scale, the positive aspects of this technology motivate the researches in this field so that better operation conditions and competitiveness of the process are achieved. The Ping Pong Bi-Bi mechanism taking into account the competitive inhibition of the alcohol has been proposed to represent the kinetics of the enzymatic transesterification in many references in the literature, but the kinetic modeling of the enzymatic ethanolysis for the production of biodiesel has not been widely studied. In this work a kinetic modeling of the ethanolysis of sunflower oil catalyzed by lipase Novozym® 435 was carried out with kinetic parameters fitted to experimental data retrieved from the literature. The model was successful in predicting the experimental data with the reactor operating in a batch model. A semibatch reactor with continuous feed of ethanol was also investigated with respect to the influence of the ethanol flow rate on the inhibition of the enzyme.

Keywords: transesterification, lipase, ethanolysis, sunflower oil, biodiesel.

1. Introduction

The shortage of the fossil fuels and the environmental degradation due to gas emissions and/or spills have contributed for the increasing in the search for alternative fuels. Biodiesel is an alternative for the diesel derived from oil, which can be synthesized through the catalytic transesterification reaction of edible/non-edible oils or animal fats with an alcohol yielding fatty acid alkyl esters and glycerol as by product. It is environmentally friendly since it is derived from renewable and biodegradable resources and produces less gas emissions compared to the conventional fuels.

Several vegetable oils can be used as raw materials such as soybean, palm, sunflower, cotton, etc. Metanol and ethanol are the most used alcohols. In Brazil ethanol is advantageous due to its enormous availability. Acid, alkaline and biocatalysts can be employed commercially for the synthesis of biodiesel. The reaction is faster in the presence of alkaline catalysts, but when the concentration of water and fatty acids in the

oil is high the acid catalysts are preferable. As far as biocatalysts are concerned, lipases catalyze transesterification reactions of the triglycerides present in vegetable oils by acting on their ester bond. The difficulties during the glycerol recovery, catalyst removal and purification steps, which are characteristics of the reactions with alkaline and acid catalysts, are not found in transesterification routs that use immobilized lipases. Despite the high costs associated with the enzymes and their deactivation due to the presence of the alcohol, the positive aspects offered by this technology encourage the researches in this field in order to overcome the drawbacks so that the process can be used in industrial scale.

The schematic of the transesterification reaction for the production of biodiesel is shown in Figure 1. An excess of alcohol is usually required because the reaction is reversible.

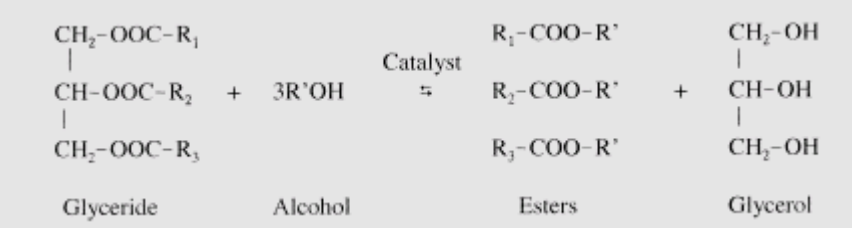


Figure 1. Scheme of triglycerides transesterification with alcohol.
R1, R2, R3, R' are the fatty acid chains

Kinetic data obtained experimentally can be modeled for scale-up purposes and also for the evaluation of the influence the different reaction parameters on the enzymatic transesterification reaction. Works that deal with modeling of the transesterification reaction for the production of biodiesel including the inhibition effect of the ethanol are seldom found in the literature. Therefore the objective of this work is to employ the Ping Pong Bi-Bi model to study the kinetic of the ethanolysis of sunflower oil with lipase Novozym® 435 as the catalyst.

2. Kinetic Modeling

The rate of an enzymatic reaction depends on factors such as substrates concentration, enzyme load, pH and temperature. Several studies in the literature report that the kinetics of the enzymatic transesterification reaction wherein raw materials derived from natural products are involved should be modeled using a model that accounts for the effects of the inhibition of the enzyme since some species present in the raw materials affect the activity of the enzyme. According to studies performed by Shimada *et al.* (2002), the deactivation of the enzyme is a consequence of the contact of the enzyme with the immiscible polar phase, containing the alcohol and glycerol, in the presence of the apolar oil phase. As reported by Dossat *et al*. (2002) e Al-Zuhair *et al*. (2006), a model widely used to describe the kinetics of this kind of reaction is based on the Ping Pong Bi-Bi mechanism which takes into consideration the influence of the alcohol and the substrates on the deactivation of the enzyme. In their model they considered the global reaction for the synthesis of biodiesel, as shown in Figure 1, instead of three consecutive reactions starting from tri-, di- and monoglyceride respectively because only small amounts of mono- and diglycerides were detected over the reaction. The rate expression given by Equation (1) was then used in the model.

$$Vi = \frac{V_m.[oil].[alc]}{Km_{oil}.[alc]\left(1+\frac{[alc]}{Ki}\right)+Km_{alc}[oil]+[oil][alc]} \quad (1)$$

Where [oil] and [alc] are the molar concentrations of oil and alcohol, respectively. Km_{oil} and Km_{alc} are the biding constants for the oil and alcohol. K_i represents the inhibition constant of the alcohol, whereas V_m denotes the maximum initial rate for the reaction.
In this work a pseudo homogeneous kinetic model with the rate per unit of enzyme load, as used by Steinigeweg *et al*. (2004), combined with the Ping Pong Bi-Bi model was applied to describe the reaction rate as given by Equation (2).

$$(-r_i) = -\frac{1}{m_{cat}}\frac{1}{V}\frac{dNi}{dt} = \frac{V_m.[oil].[alc]}{Km_{oil}.[alc]\left(1+\frac{[alc]}{Ki}\right)+Km_{alc}[oil]+[oil][alc]} \quad (2)$$

Where m_{cat} is the mass of enzyme, V is the reaction mixture volume and dNi/dt represents the number of moles of species *i* reacted per unit time.

2.1. Kinetic parameters

The experimental data taken from Hernández-Martín and Otero (2008), who studied the ethanolysis of sunflower oil catalyzed by lipase Novozym® 435, were used to fit the kinetic parameters (V_m, Km_{oil}, Km_{alc}, K_i). The reaction was carried out at ambient pressure, room temperature, initial molar ratio ethanol (pure) to oil of 20.6:1, enzyme load of 1g (50% w/w based on weight of oil) and using hexadecane as a solvent. Since the kinetic experimental data were referred as conversion of triglycerides, it was convenient to consider the sunflower oil as a pseudo component in the reaction model. Thus, the properties of the mixture of triglycerides and ethyl esters were calculated by averaging the individual properties with the corresponding mole fraction of their constituents.

The fitting of the parameters was performed with the applicative Mathcad v.14 by means of the solver *Minimize* employing the Quasi-Newton numerical method. The objective function (FO) is presented in Equation (3).

$$FO(V_m, Km_{oil}, Km_{alc}, Ki) = \sum_{j=1}^{n}\left[\left.\frac{dNi}{dt}\right|_{calc\,j} - \left.\frac{dNi}{dt}\right|_{\exp j}\right]^2 \quad (3)$$

The results for the parameters in the model were: V_m=0.00024 mol.h^{-1}.mgcat^{-1}.L^{-1}, Km_{oil}=0.000006331 mol/L, Km_{alc}=0.04 mol/L, and K_i=0.00028797 mol/L. The kinetic parameters obtained maybe considered as apparent values due to the possibility of internal and/or external diffusion limitations.

3. Results and Discussion

In order to validate the kinetic parameters a batch reactor was modeled using a set of four ordinary differential equations representing the mole balances applied to the reactants and products, which was numerically solved with the *Odesolve* routine of Mathcad v.14 with the adaptive step-size algorithm that combines the fourth-order with fifth-order Runge-Kutta methods. The initial concentrations for the reactants, products and solvent were: oil=0.4603 mol/L, ethanol=9.47 mol/L, ester=0, glycerol=0, hexadecane=0.1780 mol/L. The initial reaction volume was 4.9616 cm^3. The comparison between the predicted and experimental conversions is illustrated in Figure 2. The modeling reproduced the experimental conversion data with an absolute percent mean deviation of 4.6%. According to this plot the model predicts rather successfully the experimental data and 90% conversion was achieved in approximately 4h.

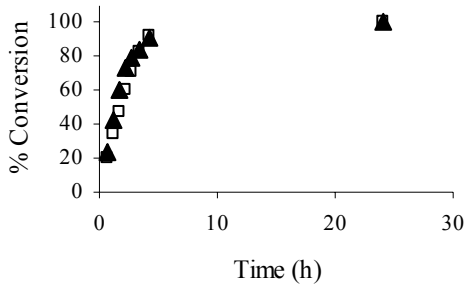


Figure 2. Sunflower oil conversion for a batch reactor. (▲) Exp., (□) Calc. Conditions: 25°C, 1g enzyme (50% w/w), ethanol:oil molar ratio=20.6:1

The model was used to calculate the molar concentrations of reactants and products during the course of the reaction. The concentration profile for oil, ester and glycerol is presented in Figure 3. This plot shows that after 6 hours no changes occured in the concentrations indicating that the reaction is finished.

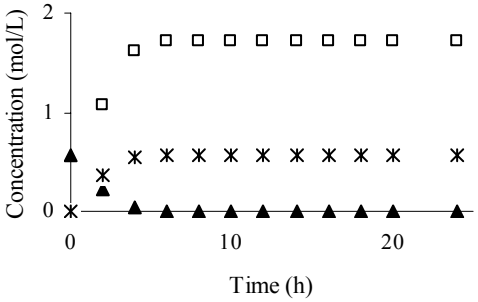


Figure 3. Concentration profiles for a batch reactor. (▲) oil, (□) ester, (*) glycerol Conditions: 25°C, 1g enzyme (50% w/w), ethanol:oil molar ratio=20.6:1

Since some enzymes suffer from deactivation due to the presence of alcohols, the stepwise reaction has been mentioned in the literature (e.g. Watanabe *et al.* (1999) and Shimada *et al.* (2002)) as an option to overcome the inhibition due to high concentration of ethanol. In such a case the transesterification reaction is carried out with sequential addition of less than stoichiometric quantities of ethanol to a fixed amount of oil. Likewise, the use of a semibatch approach would be an alternative to conduct the reaction avoiding the deactivation of the enzyme.

Based on the kinetics discussed above, the modeling of a semibatch reactor with continuous feed of ethanol was conducted in order to investigate the influence of the concentration of ethanol on the conversion. Such a model applied to a semibatch reactor comprises a system of four ordinary differential equations representing mole balances on each species and an overall mass balance. The model is described by the following set of equations:

$$\frac{dN_{oil}}{dt} = \frac{-m_{cat}\left[V_m N_{oil}(t)\right]\left[\theta_{alc} - 3\left(1 - \frac{N_{oil}(t)}{N_{oil\,0}}\right)\right]}{Km_{oil}\left(\theta_{alc} - 3\left(1 - \frac{N_{oil}(t)}{N_{oil\,0}}\right)\right)\left(1 + \frac{N_{oil\,0}\left(\theta_{alc} - 3\left(1 - \frac{N_{oil}(t)}{N_{oil\,0}}\right)\right)}{VKi}\right) + Km_{alc}\left(\frac{N_{oil}(t)}{N_{oil\,0}}\right) + N_{oil}(t)\left[\theta_{alc} - 3\left(1 - \frac{N_{oil}(t)}{N_{oil\,0}}\right)\right]} \quad (4)$$

$$\frac{dN_{oil}}{dt} = \frac{1}{3}\frac{dN_{alc}}{dt} = -\frac{1}{3}\frac{dN_{ester}}{dt} = -\frac{dN_{glyc}}{dt} \quad (5)$$

$$\frac{dN_{alc}}{dt} = 3\frac{dN_{oil}}{dt} + \frac{v_0 N_{oil0}\theta_{alc}}{V_0} \tag{6}$$

$$V = V_0 + v_0 t \tag{7}$$

Where V_0=4.9616 cm^3, N_{oil0}=0.002284 mol and θ_{alc}=20.6 represent the initial conditions for the reaction mixture volume, moles of oil and ethanol to oil molar ratio, respectively. The mol balance for oil is given in Equation (4). Equation (5) gives the stoichiometry relating the reaction rates of each species (e.g. oil, alcohol, glycerol, and esters). Equation (6) shows the mol balance for ethanol including a continuous mol flow rate, whereas Equation (7) shows the volume as a function of time. Similarly to the case of the batch reactor, the model was solved with the *Odesolve* routine of Mathcad v.14.

In order to compare the conversions obtained with batch and semibatch strategies the initial conditions of the reaction mixture volume and concentrations of the reactants have the same values in both modes. The goal was to investigate the behaviour of the conversion by varying the flow rate of ethanol added to the system continuously, v_0. The effect of ethanol flow rate on the conversion in a semibatch reactor is illustrated in Figure 4, which shows the conversions over a period of 6h for ethanol flow rates of 0.1ml/h, 1ml/h and 10 ml/h. The improvement in the conversion was basically due to the increasing in the ethanol flow rate that yields high ethanol to oil molar ratios. It appears that, under the given operation conditions, the inhibition of the alcohol is not relevant. These results are in compliance with the findings of Hernández-Martín and Otero (2008) who observed rather insignificant deactivation of Novozym® 435 by ethanol with alcohol to oil molar ratios ranging from 14.3:1 to 28:8. Additionally, Irimescu *et al.* (2002) reported a satisfactory performance of this enzyme in the ethanolysis of triacylglycerol using even high molar ratio (e.g. 77:1). The increasing in the conversion caused by the reduction in the inhibition effects might be better evaluated by starting the reaction with a concentration of ethanol lower than that used in the batch mode and consequently ethanol to oil molar ratios lower that 20.6:1. However there was no experimental data available so that the kinetic parameters could be determined.

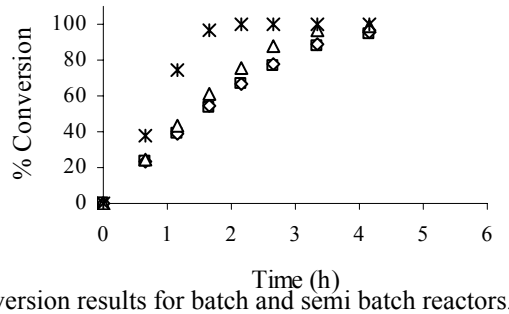


Figure 4. Conversion results for batch and semi batch reactors. (□) batch mode, (◊) semibatch with 0.1ml/h, (Δ) semibatch with 1.0 ml/h, (*) semibatch with 10 ml/h. Conditions: 25°C, 1g enzyme (50% w/w), initial molar ethanol:oil ratio=20.6:1.

4. Conclusion

A model based on the Ping Pong Bi-Bi mechanism with competitive ethanol inhibition was employed to describe the kinetics of the transesterification reaction of sunflower oil with ethanol catalyzed by lipase Novozym® 435. The kinetic constants were fit to

experimental data and the model was found to be qualitatively consistent with the proposed mechanism since it reproduced the experimental conversions rather adequately. The results of the model for a semibatch reactor with continuous addition of ethanol using the same initial conditions as in the batch mode showed improvements in the conversion basically due to the increasing in the ethanol to oil molar ratio rather than the decreasing in the inhibition effects.

5. Nomenclature

[alc]	molar concentration of alcohol, M
dNi/dt	number of moles of species *i* reacted per unit time, mol/h
$dNi/dt\|_{exp}$	experimental number of moles of species *i* reacted per unit time, mol/h
$dNi/dt\|_{calc}$	calculated number of moles of species *i* reacted per unit time, mol/h
Km_{oil}	biding constant for the oil, M
Km_{alc}	biding constant for the alcohol, M
K_i	inhibition constant of the alcohol, M
m_{cat}	mass of enzyme, mg
N_i	number of moles of species *i,* mol
N_{i0}	initial number of moles of species *i,* mol
[oil]	molar concentration of oil, M
R and R'	fatty acid chain
t	reaction time, h
V	reaction mixture volume, L
V_i	initial rate of reaction, $mol.h^{-1}.L^{-1}.mg^{-1}$
V_m	maximum initial rate, $mol.h^{-1}.L^{-1}.mg^{-1}$
V_0	initial reaction mixture volume, L
ν_0	flow rate of ethanol, $L.h^{-1}$
θ_{alc}	initial molar ratio of ethanol to oil

References

S. Al-Zuhair, K.V. Jayaraman, S. Krishnan, W. Hoong Chan, 2006, The Effect of Acid Concentration and Water Content on the Production of Biodiesel by Lipase, Biochemical Engineering Journal, 30, 212-217.

V. Dossat, D. Combes, A. Marty, 2002, Lipase-Catalysed Tranesterfication of High Oleic Sunflower Oil, Enzyme and Microbial technology, 30, 90-94.

E. Hernández-Martín, C. Otero, 2008, Different Enzyme Requirements for the Synthesis of Biodiesel: Novozym® 435 and Lipozyme® TM IM. Biores. Technology, 99, 277-286.

R. Irimescu, Y. Iwasaki, C.T. Hou, 2002, Study of TAG ethanolysis to 2-MAG by immobilized *Candida Antarctica* lipase and synthesis of symmetrically structured TAG, J. Am. Chem. Soc. , 79, 879-883.

Y. Shimada, Y. Watanabe, A. Sugihara, Y. Tominaga, 2002, Enzymatic Alcoholysis for Biodiesel Fuel Production and Application of the Reaction to Oil Processing, Journal of Molecular Catalysis B: Enzymatic, 17, 133-142.

S. Steinigeweg, J. Gmehling, 2004, Transesterification Processes by Combination of Reactive Distillation and Pervaporation, Chem. Eng. and Processing, 43, 447-456.

Y. Watanabe, Y. Shimada, A. Sugihara, Tominaga, Y., 1999, Stepwise ethanolysis of una oil using immobilized *Candida Antarctica* lipase, J. Biosci. Bioeng., 88, 622-626.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

How Modeling can Help to Discriminate Assumptions on the Influence of Nitrogen Consumption on pH during Fermentation

Huberson AKIN[a], Cédric BRANDAM[a], Xuân-Mi MEYER[a], Pierre STREHAIANO[a]

[a] *Université de Toulouse-INPT-ENSIACET, Laboratoire de Génie Chimique, 118 route de Narbonne, 31077 Toulouse cedex 4, France*

Abstract

This work deals with the use of a mathematical model to investigate the influence of the nitrogenous source on the pH during alcoholic fermentation in winemaking. The application of the model to fermentation medium whose nitrogen source was made up only of ammoniac confirms the assumption that the assimilation of one mole of ammonia releases one mole of proton in the medium. The use of the model made it possible to invalidate two assumptions concerning the impact of the assimilation of the amino acids on the pH. The most probable assumption is that the assimilation of the molecules of amino acids charged positively led to the emission of protons in the extra cellular medium. The model including this assumption was also used successfully to predict pH in the case of a fermentation realised with natural grape must.

Keywords: nitrogen consumption, modelling, pH, fermentation

1. Introduction

In a recent context of an international competition, enologists seek to better understand the wine making phenomena to better control the production quality. If the source of carbon in the process of wine making comes primarily from the glucose and fructose available in the grape, the source of nitrogen is much varied and can take forms more or less complex: ammonium ions, amino acids, peptides and proteins.

Nitrogen compounds are quantitatively the second nutriment of the yeast, after the carbon element. It is known that only ammonium ions and some amino acids can be metabolised by yeasts and that the kinetic of the fermentation and aroma production are particularly influenced by the quantity and the nature of amino acids.

Various studies have reported a pH decrease during alcoholic fermentation by *Saccharomyces cerevisiae* on medium where ammonium ions were the only nitrogenous source (Kotyk, 1989, Siegler et al., 1981). According to Won et al. (1993) and Castrillo et al. (1995), when one mole of ammonium is consumed, one mole of proton is excreted resulting in a pH decrease. The mechanisms of assimilation of amino acids by yeast seem more complex. The type of exchanges taking place when amino acids pass through cell membrane is not yet known. Several assumptions on the relation between pH values and amino-acid assimilation can be drawn from these studies but until now, no work has permitted to discriminate them.

In a previous study, Akin et al. (2008) have proposed a mathematical model to calculate the pH of a grape must during the fermentation. This model has been validated on synthetic medium and natural grape musts.

The objective of this paper is to investigate the possibility to correlate the evolution of the pH in the fermentation media to the most probable mechanism of assimilation of amino-acids by the mean of a mathematical model. Three assumptions for the effect on pH of assimilation of amino acids were tested:

- Assumption 1: assimilation of amino acid does not influence the pH
- Assumption 2: consumption of one mole of amino acid causes the excretion of one mole of H^+, like for the ammonium consumption
- Assumption 3: consumption of one mole of amino acid charged positively only causes excretion of one mole of protons to respect the electroneutrality principle.

These different assumptions drawn on the assimilation of ammonium ions and amino acids are introduced in the model and experimental and simulated values are then compared to discriminate them.

2. Materials and methods

Saccharomyces cerevisiae QA-23 commercialised by Lallemand inc. was used as it is a classical yeast strain for white winemaking.

Fermentations were carried out on synthetic media whose composition was close to white grape must: glucose (200 $g.L^{-1}$), malic acid (6 $g.L^{-1}$), citric acid (6 $g.L^{-1}$), KH_2PO_4 (0.75 $g.L^{-1}$), K_2SO_4 (0.5 $g.L^{-1}$), $MgSO_4$, 7 H_2O (0.25 $g.L^{-1}$), $CaCl_2$, $2H_2O$ (0.16 $g.L^{-1}$), NaCl (0.2 $g.L^{-1}$), a mixture of oligo elements, vitamins and anaerobia factors. The nitrogen source differed with experiments: ammonium ions for an equivalent of 420 mg $N.L^{-1}$ for the medium 1 and a mixture of amino acids for an equivalent of 384 mg $N.L^{-1}$ for the medium 2. In this mixture, only 305 $mgN.L^{-1}$ could be assimilated by yeasts since 79 $mgN.L^{-1}$ was proline which can not be metabolised by *Saccharomyces cerevisiae*. The pH of media was adjusted to 3.3 before autoclaving with a sodium hydroxide solution (8N).

Experiments were also performed on a natural grape must provided by SOPAGLY Company. The grape must contains an equivalent of 74.9 $mgN.L^{-1}$ of ammonium ions and an equivalent of 136 $mgN.L^{-1}$ of a mixture of amino acids (114.8 $mgN.L^{-1}$ could be assimilated by yeast, 21.2 $mgN.L^{-1}$ was proline). Fermentations were carried out in New Brunswick Scientific (NBS) fermentor type of 2 litres at 20°C under a 150 rpm agitation. Temperature was regulated at 20°C. The pH was not controlled during this anaerobic fermentation. For all fermentations, the initial cellular concentration was fixed to 8.10^6 $cells.mL^{-1}$.

The cell concentration was estimated using an electronic analyser Beckman Coulter. Concentrations of ammonium, malic and lactic acids were measured by enzymatic kits (MicroDom). Citric, acetic and succinic acid, ethanol and glycerol were measured by HPLC method. Mineral elements (calcium ion, sodium ion, magnesium ion, potassium ion) were measured with an emission spectrometer of Horiba Jobin Yvon. Their concentration was evaluated on the initial medium. They were considered as constant during the fermentation. Individual amino acid concentrations were determined by Biochrom 30 method. The pH was measured using an external pH-meter (PHM210 Radiometer analytical) with a precision of 0.1 pH unit.

3. Mathematical model for pH calculation

The pH calculation model used in this study was based on the thermodynamic equilibrium of electrolytic compounds in solution. The molality of hydrogen ions and so the pH, are determined by solving a non linear algebraic equation system. For more details and model validations, it is possible to refer to Akin et al. (2008)

Instead of writing classically the mass balance on each species in solution, the mass balance was written on an invariant of the reactive system. Then, it is possible to express molalities of different amino acid and acids species (m_i) as an explicite function of m_{H^+}. The equilibrium dissociation constants for organic acids are calculated as a second order polynomial function of volumic percent in ethanol of the solution (Usseglio-Tomasset and Bosia, 1978). The pKa of NH_4^+/NH_3 and of all the amino acids were taken constant with ethanol concentration.

The activity coefficient of H^+ (γ_{H^+}) necessary to calculate pH according to the equation $pH = -\log_{10}(a_{H^+}) = -\log_{10}(\gamma_{H^+} m_{H^+})$ was calculated with a Debbye-Hückel model with the MacInnes convention. The calculation of the average coefficient of activity was carried out with the law of Debye-Hückel, in the scale of the molalities

New correlations for the density and the dielectric constant were established taking into account the influence of the sugar (0 to 200 $g.L^{-1}$) and the ethanol (0 to 12 % vol) at 20°C. The formulation of the model with explicit calculations of all the species in solutions as a function of the molality in H^+ and the ionic force I, enables to solve a non linear equation system consisting of only two equations: the ionic force equation and the electro neutrality equation, with two unknowns: the molality in H^+ and the ionic force I.

$I = \frac{1}{2}\left[\sum_{i=1}^{n\_species} m_i z_i^2 \right]$ where z_i represents the charge of the electrolyte i

$$\sum_{i=1}^{n\_electrolytes} m_i(m_{H^+}, I)\, z_i = 0$$

This formulation greatly facilitates the initialisation of the Newton-Raphson iterative procedure carried out to solve the problem.

4. Results and discussions

4.1. Fermentation of the medium 1

During the first fermentation period (0 to 30 hours), there was no biologic activity. During the second period, between 30 and 110 hours, the metabolic activity of yeast increased: a consumption of sugars and ammonium and a production of biomass and ethanol can be observed. At the end of the fermentation, ethanol concentration was about 82 $g.L^{-1}$, biomass attained almost 160.10^6 $cells.mL^{-1}$ and ammonium was entirely consumed. The decrease of pH was also emphasized since its final value was 2.90.

The measurement of organic acid concentrations indicated that very small quantities were produced: 0.16 $g.L^{-1}$ of succinic acid, 0.18 $g.L^{-1}$ for acetic acid. Citric and malic acids remained constant to 6 $g.L^{-1}$ during the whole fermentation.

The pH calculated with the model (3.37) at the very beginning of the fermentation when the medium composition was perfectly known, was in agreement with the experimental value. Ethanol, sugar, organic acids and ammonium concentrations have been measured at different times for which pH values can be calculated by the model.

For this medium, the assumption of one mole of proton released for each ammonium ion (in cationic form) consumed by *Saccharomyces cerevisiae* was in agreement with the experimental data (figure 1).

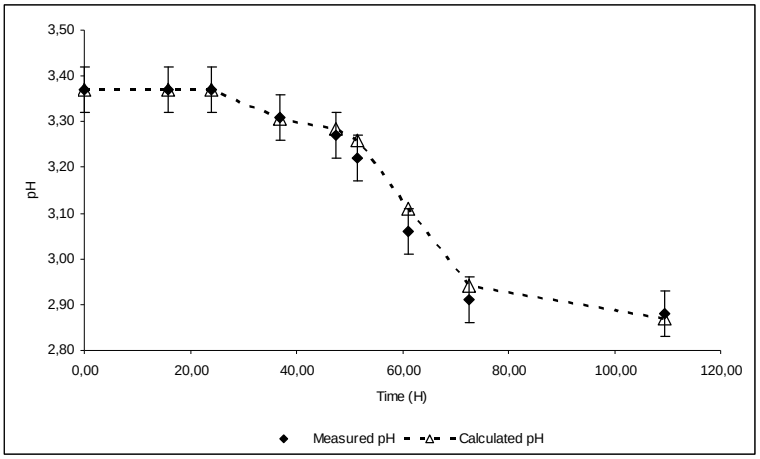


Figure 1: Comparison of experimental and simulated pH values in medium 1 (T=20°C)

4.2. Fermentation of the medium 2

During the first period of fermentation (0 to 20 hours), yeasts were in a lag phase and no changes were observed in the medium. During the second period (20 to 110 hours), ethanol production occurred to reach 94 $g.L^{-1}$ at the end of the experiment. Concentration of biomass rose to 200.10^6 $cells ml^{-1}$and amino acids were totally consumed (except the proline). 0.3 $g.L^{-1}$ of succinic acid and 0.3 $g.L^{-1}$ of acetic acid were produced. Malic and citric acids remained constant to their initial concentration at 6 $g.L^{-1}$. The pH value decreased from 3.23 to 3.13 after 50 hours of the fermentation which corresponds to the end of the assimilation of amino acids. Afterwards, pH increased regularly to reach 3.25 at the end of the experiment. This rise was previously shown to be related to the influence of ethanol on the organic acids dissociation equilibriums. The pH evolution was simulated and compared to experimental data, under the three assumptions (figure 2).

Assumption 1:
Since organic acids production was very low, no other phenomenon except amino acid consumption could explain pH decrease. Model calculation confirmed this fact, by calculating a constant pH value for this assumption.

Assumption 2:
The minimum pH value calculated was 3.03, that was lower than the experimental one (3.13). The production of proton is then overestimated by the model. This would indicate that all amino-acids can not be considered at the same level. The assimilation of some of them may lead to the excretion of one or more protons whereas other may be without any effect on pH.

Assumption 3:
To test the third assumption, the total molalities of amino acids on cationic, anionic and neutral form during the fermentation were determined. The amino acids repartition in the medium, consisted initially of 70% amino acids globally under the neutral form, 27 % with one positive charge, 1.3 % with two positive charges and 1.7 % with one negative charge. A good adequacy between the calculated and experimental pH values was found (figure 2). The maximal deviation was 0.04 point of pH that is less than the precision of the measure.
The simulation results were globally in agreement with the experimental data when the third assumption was formulated.

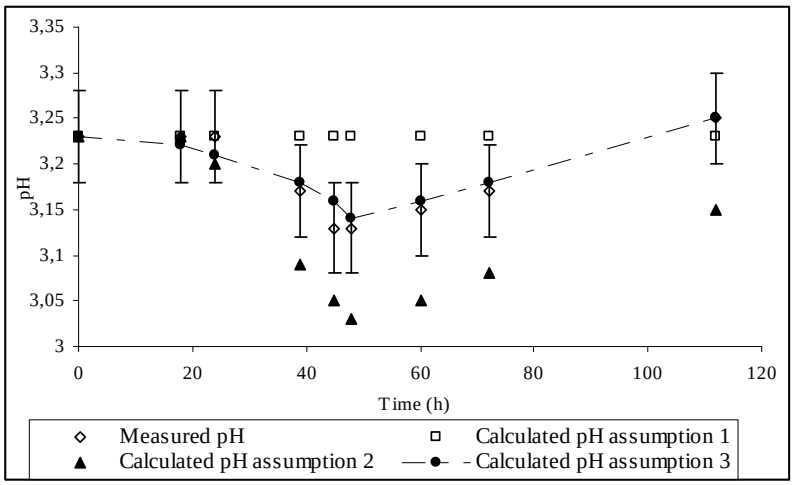


Figure 2: Comparison of experimental and calculated pH with assumption 1, 2 and 3 in MS_AA medium (Initial pH=3.23; T=20°C)

4.3. Fermentation of the natural white grape must

To confirm the assumption made for the effect on pH of assimilation of amino acids, fermentation was realised on a natural white grape must. The pH evolution was very similar to the synthetic media: a pH decrease (from 3.32 to 3.14) during the first 50 h of the fermentation corresponding to the concomitant ammonium and amino acid consumption, then a pH increase (from 3.14 to 3.3) corresponding to the ethanol production (88 $g.L^{-1}$). Organic acid production was more important than in synthetic media. There was a production of 0.7 and 0.5 $g.L^{-1}$ of succinic acid and acetic acid. Lactic acid was not produced and citric acid remained constant to 1.5 $g.L^{-1}$. Malic acid was little consumed from 6 to 5.4 $g.L^{-1}$.
The quantities of amino acid on anionic, cationic and neutral forms are calculated for the grape must. At the beginning of the fermentation, it consists of 61% of amino acids in neutral form, 36 % with one positive charge, 1.8 % with 2 positive charges and 1.2 % with one negative charge. pH values were calculated assuming the third assumption are here again in a good agreement with experimental values (figure 3).

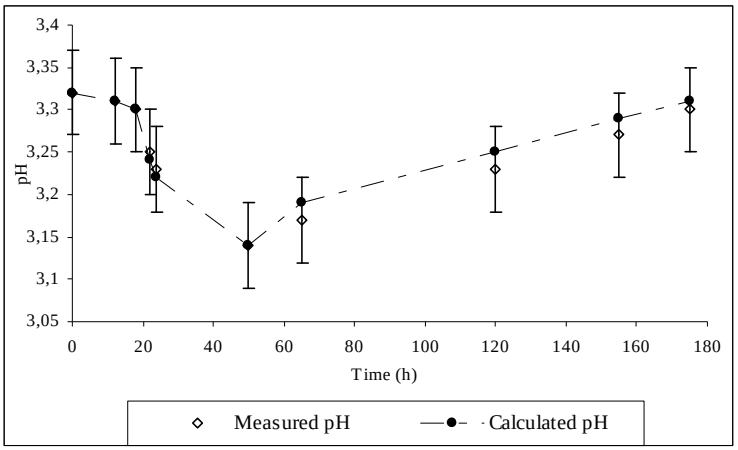


Figure 3: Comparison of experimental and calculated pH in grape must (Initial pH=3.32; T=20°C)

5. Conclusion

This study confirms that the acidification of the medium during alcoholic fermentation is directly linked to the assimilation of the nitrogen source. On a medium where nitrogen was only brought by ammonium (420 mg $N.L^{-1}$), the pH decreases down to 0.47 pH unit. With amino acids as nitrogen source (305 mg $N.L^{-1}$ of assimilable nitrogen) the decrease of the pH was lower, only 0.1 unit of pH. The results of the pH calculation model have confirmed that the mechanism of assimilation of each ammonium ions by yeast releases one proton in the medium. For amino acids assimilation, several assumptions has been tested since literature was equivocal on the effect on the pH. The model allowed us to suggest that only the cationic form of amino acids was responsible for the pH decrease. This assumption has been tested successfully on a synthetic medium but also on a natural grape must. It must be validated with additional fermentations realised on medium with different ratio of amino acids.

References

A.Kotyk, Proton extrusion in yeast, In Fleischer S. and Fleischer B. (Eds) Method in enzymology, 174 (Biomembranes, part U) Academic Press, New York (1989) 592-603

K.Sigler, A.Knotkova and A.Kotyk Factor governing substrate induced generation and extrusion of proton in yeast. Biochim. Biophys. Acta 643 (1981) 572-582

J.I Castrillo., I.De Miguel, U.O.Ugalde, Proton production and consumption pathways in yeasts metabolism. A chemostat culture analysis, Yeast 11 (1995) 1353-1365.

J.I.Won, Y.L.Yang., B.G.Kim and C.Y.Choi., Adaptative control of specific growth rate based on proton production in anaerobic fed-batch culture. Biotechnology letters 15 (5) (1993) 511-516.

H.Akin, C.Brandam, X-M. Meyer and P. Strehaiano, A model for pH determination during alcoholic fermentation of a grape must by *Saccharomyces cerevisiae*, Chem. Eng. Process 47 (2008) 1986-1993

L.Usseglio-Tomasset and P.D.Bosia**,** Determinazione delle costanti di dissociatione dei principali acidi del vino in soluzioni idroalcoliche di interesse enologico, Rivisto di viticoltura e di enologia di conegliano 31 (1978) 380-403

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Integrated procurement and operational planning of a biorefinery considering contracts and futures

Choamun Yun,[a] Young Kim,[b] Jeongho Park,[a] Sunwon Park [a, *]

[a]*Department of Chemical and Biomolecular Engineering, KAIST, 335 Gwahangno, Daejeon 305-701, Korea*
[b]*Energy Plant Research Division, KIMM, 104 Sinseongno, Daejeon 305-343, Korea*

Abstract

The profit of a biorefinery is highly affected by the supply of its raw materials and margin from the product. Taking responsive actions to the unstable supply of raw materials and the fluctuating prices are of major concern for the efficient management of biorefineries. In this study, a biorefinery complex is defined to tackle these issues by diversifying products as well as raw materials. Various raw materials are purchased based on the contracts and the spot prices of candidate raw materials while their total amount depends on product demands. The supply of required raw materials is accompanied by the futures contracts to curtail the risks involved in the procurement. In the downstream process, operational planning of fermentation and separation units is established according to the demand and margin of multiple products. The proposed model of integrated planning for a biorefinery would contribute to escalating its profitability and operational flexibility.

Keywords: Integrated biorefinery; multiproducts; procurement planning; futures

1. Introduction

Biorefinery has been appreciated for its environmentally friendly process and products. Despite of the subsidies and tax advantages from many governments, however, the industry has not been grown sufficiently for its relatively low profitability (Lakhdar and Papageorgiou, 2008). Accordingly, developing an economically sustainable as well as commercially viable process has been a crucial issue. Different approaches on alternative media, enzymes, and processes with various reaction conditions have been studied (Porro et al., 1999; Steffens et al., 2000; Vishnu et al., 2000; Willke and Vorlop, 2001; Li et al., 2006). Koutinas et al. (2004; 2007) studied the alternative upstream processes for the production of generic microbial feedstocks while downstream processing options are investigated by Joglekar et al. (2006). The idea of producing multiple products from a biorefinery as in the petrochemical industry has also been explored by several studies (Tran et al., 2004; Kamm and Kamm, 2007; Sadhukhan et al., 2008; Sammons et al., 2008). After the desirable process is selected, profit of a biorefinery still varies with the operation of the process. As in other chemical processes, the productivity is uncertain due to the changing operation parameters such as the yield of batch and the purity after separation processes (Lim et al., 2005; Lakhdar and Papageorgiou, 2008). The more significant factors for the high profit variability are the fluctuating costs for raw materials and the price change of products (Franceschin, 2008). The prices of raw materials change for a variety of reasons such as seasonal

[*] *To whom correspondence should be addressed. E-Mail: sunwon@kaist.ac.kr*

effects, states of harvest, or policy changes; and the product prices change due to the changing market conditions.
One way of curtailing the risks is to diversify the raw materials and products. By varying the source of starch and the targeting markets, a biorefinery may be actively respond to the changes. On the other hand, the risk can also be reduced with financial derivatives such as options and futures (Yun et al., 2009; Park et al., 2009). In this work, a procurement and operational planning model is proposed to decrease the profit variability of a biorefinery by flexibly operating an integrated production process for multiple products. Moreover, the risks involved in the procurement are reduced by purchasing the diversified raw materials and their futures contracts.

2. Problem description

A biorefinery process manufacturing multiple products from diversified raw materials starts from milling of purchased raw materials, e.g., corn or wheat. The processed raw materials are liquefied and saccharified in the form of starch. Subsequently, solution containing glucose is generated from the saccharification unit and it is fed into fermenters with enzymes and additional media required for the production of a desired product. The products are manufactured from the shared operating units.
The outlet streams from fermenters are separated into two main downstreams: one stream for the production of ethanol, and the other for that of biochemicals. The broth fermented for ethanol production undergoes the ethanol recovery process consisting of a beer column and a rectifier. Other biochemicals of lactic acid, itaconic acid and citric acid are assumed to be purified through an identical procedure. The broths are filtered successively in a filtration unit, an ion exchanger and a decolorization unit. The remaining liquid is condensed before the products are crystallized and dried.

3. Operational planning of an integrated process

The main assumptions of the procurement and operational planning model are as follows: (a) Processing times are constant for every batch; (b) Clean-up times are ignored; (c) Sufficient raw materials are always available; (d) The operating costs of the units are constant.
The profit to be maximized in this model is:

$$\sum_n Z_n = \sum_n \sum_k P_{k,n} \cdot X_{k,n} - \sum_n \sum_r P_{r,n} \cdot X_{r,n} - \sum_n \sum_j OC_j \delta_{j,n} - \sum_n \sum_i IC_i I_{i,n} \qquad (1)$$

Z_n is the profit of time period n, the revenue subtracted by raw material costs, operating costs, and inventory costs. $P_{k,n}$ and $X_{k,n}$ are the price and sales amount of product k; $P_{r,n}$ and $X_{r,n}$, the price and purchase amount of raw material r; OC_j, the operating costs of unit j; and IC_i, the inventory cost of storage tank i whose level is $I_{i,n}$. $\delta_{j,n}$ is a binary variable indicating whether operating unit j starts its operation at the end of nth time period.
Inventory level of a raw material is increased with its procurement and decreased when there is an inlet stream to a milling unit, M_n^{α}, i.e.,

$$I_{r,n}^{M} = \sum_r (X_{r,n}^{L} + X_{r,n}^{S}) - M_n^{\alpha} \qquad (2)$$

where $X_{r,n}^{L}$ and $X_{r,n}^{S}$ are the purchasing amount of raw material r at time period n through long term and spot contracts, respectively.

Inventory levels between milling and liquefaction unit is balanced with the outlet stream from the milling unit, $M_{r,n}^{\beta}$, and inlet to the liquefaction unit, V_l^{α}.

$$I_n = I_{n-1} + M_n^{\beta} - V_l^{\alpha}\delta_{l,n} \tag{3}$$

Similarly, intermediate storage level, $I_{i,n}$, between adjacent units j and j' are calculated as below.

$$I_{i,n} = I_{i,n-1} + V_j^{\beta} e_{j,n} - V_{j'}^{\alpha}\delta_{j',n} \qquad \text{for } \forall n \tag{4}$$

$e_{j,n}$ defines whether operating unit *j* ends its operation at the end of *n*th time period. In the downstream process from fermentor units,

$$I_{i,k,n} = I_{i,k,n-1} + V_j^{\beta} e_{j,k,n} - V_{j'}^{\alpha}\delta_{j',k,n} \qquad \text{for } \forall k,n \tag{5}$$

For final storages, the outlet stream is the sales amount of product *k*, $X_{k,n}$.

$$I_{i,n} = I_{i,n-1} + V_j^{\beta} e_{j,n} - X_{k,n} \qquad \text{for } \forall k,n \tag{6}$$

Additionally, the following time constraints are introduced to prevent the overlapped production plan in a unit.

$$e_{j,n+\tau_j} = \delta_{j,n} \tag{7}$$

$$e_{j,n} = 0 \quad \text{for} \quad n \cdot \Delta t \le \tau_j \tag{8}$$

$$\sum_{u=n}^{n+\tau_j} \delta_{j,u} \le 1 \tag{9}$$

where τ_j is the processing time at unit *j* and Δt is the length of time period *n*.

4. Risk management for procurement

The prices of the raw materials are not necessarily cheaper or the same as the expectation at the point of purchase. Accordingly, the profit of a biorefinery may alter with the changing prices to a great extent even when the plant is operated following the optimal procurement and operational plans. In this study, futures contracts are considered for hedging the risks involved in the purchase of raw materials.
The profit generated from the futures transaction is as below.

$$\sum_r (PF_{r,1} - PF_{r,0}) XF_r \tag{10}$$

where $PF_{r,0}$ and $PF_{r,1}$ are the futures prices when taking the position at time period $n = 0$ and closing out the position at time period $n = 1$, respectively. The size of futures contracts, XF_r, is decided depending on the hedging strategy. Here, the objective is to hedge against changes in the price of raw materials. As the spot and futures contracts are generally purchased independently, it is rational to add the profits generated by the both. Consequently, the profit of a biorefinery is redefined as

$$\sum_t Z_t + \sum_t \sum_r (PF_{r,1} - PF_{r,0}) XF_r \tag{11}$$

5. Scenario generation

Planning models are calculated using the scenarios generated either by decision makers or by mathematical models based on historical data. The scenarios of commodity prices are generated by assuming that the price follows Geometric Brownian Motion (Schwartz, 1997). The scenarios of futures prices are generated by assuming that their deviation from spot prices follows a normal distribution.

6. Illustration

The developed model is illustrated with a simplified example process manufacturing ethanol, lactic acid, and itaconic acid, from corn and wheat.

As a first step, price scenarios are generated for raw materials, futures of raw materials, and products on the basis of historical data from National Agricultural Statistics Service (2008), Tran et al. (2004), and Kenyon (2001). Further pertinent data are provided in Yun et al. (2009).

The results obtained for the generated scenarios are shown in Tables 1 through 3. The flexible operational planning results in the slightly higher profit with its standard deviation decreased by 6 % as indicated in Table 2.

Table 1. Some price scenarios and the corresponding optimal production plans

No. of scenarios	Price change ($/kg)				Optimal production (kg)				Profit ($/10 days)
	Ethanol	Lactic acid	Itaconic acid	Citric acid	Ethanol	Lactic acid	Itaconic acid	Citric acid	
m (σ)*	0.728 (0.253)	1.639 (0.014)	2.270 (0.475)	1.459 (0.01)	21,000	10,900	9,600	0	44,289
1	-0.435	-0.025	0.389	0.015	6,300	8,720	11,000	1,920	37,875
2	0.314	-0.003	0.351	-0.006	22,000	8,720	11,000	1,920	55,271
3	-0.154	0.005	0.039	-0.004	12,600	10,900	7,680	3,840	38,949
4	-0.438	0	0.429	0.013	12,600	8,720	9,600	3,840	39,899
⋮									
1000	0.523	0.014	0.307	0.005	22,000	6,540	11,000	3,840	56,203

*m: mean; σ: standard deviation;
data sources: http://www.the-innovation-group.com/ChemProfiles/,
http://mindbranch.com/listing/product/R154-858.html

Table 2. Profit change by flexible operational planning

Profit	Fixed operation	Flexible operation
Mean, m ($/year)	1,623,826	1,630,743
Standard deviation, σ	7,557	7,104 (-6%)

The amount of required raw materials as well as the production ratios of each product *k* is determined according to the price scenarios. The prices of raw materials are fixed in this case to evaluate the effectiveness of the proposed operational planning. Subsequently, the profit variability is further reduced using the futures contracts of raw materials. The size of futures contracts demanded is determined to minimize the profit variance or to maximize the risk-return ratio for the price scenarios of raw materials and futures contracts. Figure 1 shows that the resulting cumulative profit curve after hedging

exhibits the steeper slope, i.e., the less variance of the profit, than the one before. Table 3 indicates that the variance reduction is 25.6 % of its standard deviation where the optimal hedge ratio is suggested to be 0.24. In this case, it turns out that the σ minimizing hedge ratio also maximizes the risk-return ratio (m/σ).

Table 3. Profit change by purchasing the futures of raw materials

Profit	Without hedging	With hedging	
		Minimizing σ	Maximizing m/σ
Mean, m ($/year)	1,937,735	1,942,920	1,942,920
Standard deviation, σ	2,027	1,508 (-25.6%)	1,508 (-25.6%)
Hedge ratio*	0	0.24	0.24

*hedge ratio = (size of the futures contracts) / (size of purchasing raw materials)

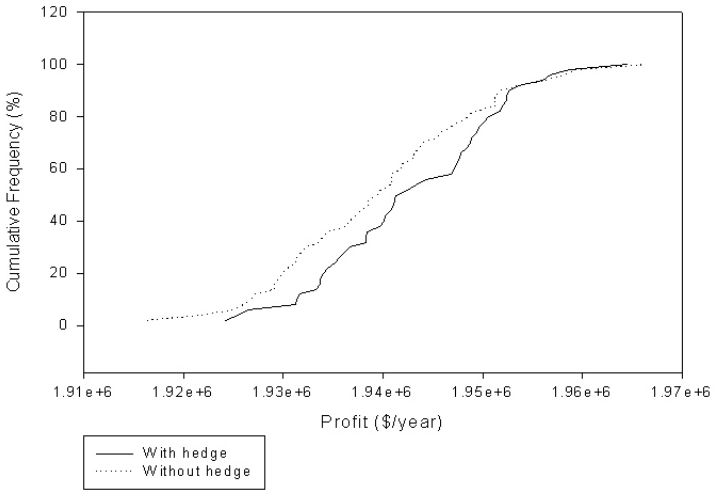


Figure 1. The cumulative profits before and after hedging using futures.

7. Conclusion

An integrated biorefinery process has been proposed to enhance the economic profitability of manufacturing and supplying the environmentally friendly products. The high profit variability of a biorefinery owing to the price fluctuations in the market could be reduced by diversifying the raw materials and products as well as by purchasing the futures contracts. The operational and financial planning model for the integrated biorefinery process has been described with an illustrative example. A biorefinery could make higher profit from the flexible production of ethanol and biochemicals. Furthermore, the profit variation was evaluated from historical data so that the procurement planning on the raw materials and futures could be established to prepare for the expected risks. The operation of the proposed integrated planning process will contribute to escalating the efficiency and flexibility of biorefineries in volatile environments.

References

G. Franceschin, Zamboni, A., Bezzo, F. and Bertucco, A., 2008, Ethanol from corn: a technical and economical assessment based on different scenarios. Chem Eng Res Des 86: 488-498.

J.C. Hull, 2005, Fundamentals of futures and options markets, 5th edition. (Prentice Hall, NJ, USA).

H.G. Joglekar, Rahman, I., Babu, S., Kulkarni, B.D. and Joshi, A., 2006, Comparative assessment of downstream processing options for lactic acid. Sep Purif Technol, 52: 1-17.

B. Kamm, and Kamm, M., 2007, Biorefineries – multi product processes. Adv Biochem Engin/Biotechnol, 105: 175-204.

D.E. Kenyon, 2001, Virginia basis tables for corn, soybeans, and wheat. Agricultural and Applied Economics, Publication Number 448-016.

A.A. Koutinas, Arifeen, N., Wang, R. and Webb, C., 2007, Cereal-based biorefinery development: integrated enzyme production for cereal flour hydrolysis. Biotechnol Bioeng, 97: 61-72.

A.A. Koutinas, Wang, R. and Webb, C., 2004, Restructuring upstream bioprocessing: technological and economical aspects for production of a generic microbial feedstock from wheat. Biotechnol Bioeng, 85: 524-538.

J.R. Kwiatkowski, McAloon, A.J., Taylor, F. and Johnston, D.B., 2006, Modeling the process and costs of fuel ethanol production by the corn dry-grind process. Ind Crops Prod, 23: 288-296.

K. Lakhdar, and Papageorgiou, 2008, An iterative mixed integer optimization approach for medium term planning of biopharmaceutical manufacture under uncertainty. Chem Eng Res Des, 86: 259-267.

Z. Li, Ding, S., Li, Z. and Tan, T., 2006, L-Lactic acid production by *Lactobacillus casei* fermentation with corn steep liquor-supplemented acid-hydrolysate of soybean meal. Biotechnol J, 1: 1453-1458.

A.C. Lim, Zhou, Y., Washbrook, J., Sinclair, A., Fish, B. et al., 2005, Application of a decision-support tool to assess pooling strategies in perfusion culture processes under uncertainty. Biotechnol Prog 21: 1231-1242.

National Agricultural Statistics Service (NASS), Agricultural Statistics Board, U.S. Department of Agriculture, (July 2008). Agricultural prices.

J. Park, S. Park, C. Yun and Y. Kim, 2009, An integrated model for financial risk management and optimal operation of a refinery, Ind Eng Chem Res, submitted.

D. Porro, Bianchi, M.M., Brambilla, L., Menghini, R., Bolzani, D. et al., 1999, Replacement of a metabolic pathway for large-scale production of lactic acid from engineered yeasts. App Env Microbiol, 65: 4211-4215.

J. Sadhukhan, Mustafa, M.A., Misailidis, N., Mateos-Salvador, F., Du, C. and Campbell, G.M., 2008, Value analysis tool for feasibility studies of biorefineries integrated with value added production. Chem Eng Sci, 63: 503-519.

N.E. Sammons Jr., Yuan, W., Eden, M.R., Aksoy, B. and Cullinan, H.T., 2008, Optimal biorefinery product allocation by combining process and economic modeling, Chem Eng Res Des, 86: 800-808.

E.S. Schwartz, 1997, The stochastic behavior of commodity prices: implications for valuation and hedging. J Finance, 52: 923-973.

M.A. Steffens, Fraga, E.S. and Bogle, I.D.L., 2000, Synthesis of bioprocesses using physical properties data. Biotechnol Bioeng, 68: 218-230.

T. Tran, Patel, T., Iland, T. and Truong, J., 2004, OU biorefining technical report for biomass production.

C. Vishnu, Seenayya, G. and Reddy, G., 2000, Direct conversion of starch to L(+) lactic acid by amylase producing *Lactobacillus amylophilus* GV6. Bioproc Eng, 23: 155-158.

Th. Willke, and Vorlop, K.-D., 2001, Biotechnological production of itaconic acid. Appl Microbiol Biotechnol, 56: 289-295.

C. Yun, Y. Kim, J. Park and S. Park, 2009, Optimal procurement and operational planning for risk management of an integrated biorefinery process, Chem Eng Res Des, accepted.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Dynamical Model for the fermentative production of fructooligosaccharides

O. Rocha[a,b], C. Nobre[a], A. Dominguez[a], D. Torres[a,b], N. Faria[b], L. Rodrigues[a,b], J.A. Teixeira[a], E.C. Ferreira[a] and I. Rocha[a,b]

[a]*IBB - Institute for Biotechnology and Bioengineering, Centre of Biological Engineering, Universidade do Minho, Campus de Gualtar, 4710-057 Braga, PORTUGAL*
[b]*Biotempo, Lda., Avepark – Zona Industrial da Gandra, Apartado 4152, 4806 – 909, Caldas das Taipas, Braga, Portugal*

Abstract

In this paper a detailed mathematical model is presented for the fermentative production of fructo-oligosaccharides with *Aspergillus* sp. The model accounts for hydrolysis and transfructolization reactions, as well as biomass formation and it contains 27 parameters that were determined from experimental data using a System Biology toolbox with the Simulated Annealing method for curve fitting. Several additional experiments were performed in bioreactors where the time variation of 7 state variables (Sucrose, Glucose, Fructose, 1-Kestose, Nystose, 1-fructosyl nystose and Biomass) was measured.
Experimental data were compared with results from simulations using the estimated parameters and it was verified that the model can predict the FOS production profile. The good agreement between simulated and experimental data was verified by calculating the relative percentage deviation modulus, which was lower than 10% for all cases except one. The derived and validated model can be used for process optimization, for example for indicating which fed-batch strategy could be used to improve the production of FOS while minimizing glucose concentration.

Keywords: Modelling, Simulation, Fructooligosaccharides.

1. Introduction

Within Industrial Biotechnology, a very promising application is the production of ingredients for functional foods, since the market for those products has been growing at very interesting rates [9]. In recent years some prebiotics have been described as beneficial food ingredients because of their proprieties of modifying the intestinal microbiota, favoring the growth of some beneficial bacteria [1;2;10;11]. Fructo-oligosaccharides (FOS) have become one of the most important prebiotic products with healthy properties, being possible to find them, usually in trace amounts, as natural components in fruits, vegetables and honey [8;11]. Although industrially these products are mainly extracted from those natural sources, they can be also produced from sucrose by the action of β- fructofuranosidase [FFase; EC 3.2.1.261] obtained from some organisms. Various fungi such as *Aureobasidum* sp., *Aureobasidum pullulans*, *Aspergillus niger*, *Aspergillus japonicas*, *Aspergillus oryzae* and *Scopulariopsis brevicaulis* [3;5;6;8;12;13;15] produce those oligosaccharides that are mainly composed of 1-Kestose, Nystose, and Fructosylfuranosyl nystose in which 1-3 fructose units are bound at the β-2,1 position of sucrose [16].

However, besides the fructosyltransferase activity, β-fructofuranosidase also exhibits hydrolytic activity [4;5;7;12], which can dominate the process depending on a combination of factors including the sucrose concentration. This fact will ultimately lead to lower production yields and to a contamination of the final product with the monosaccharides glucose and fructose. Additionally, in a fermentative process for the production of FOS, substrate consumption for biomass growth has also to be considered, increasing even further the complexity of the process and motivating the application of mathematical modelling approaches such that non-obvious operation conditions can subsequently be found that maximize the productivity of FOS and minimize the accumulation of monosaccharides.

2. Mathematical Model

The main aim of this work was to formulate a general model that characterizes the main reactions representing the fermentative FOS production process. It is based on the empirical equations of enzymatic production of fructooligosaccharides from sucrose. The model contemplates both hydrolysis and transfructosylation kinetic equations representing β- fructofuranosidase activity and growth rate equations for the microorganism. The enzymatic reactions were divided in two main categories: the hydrolysis reactions, representing FOS and sucrose degradation, and the transfructosylation reactions that describe FOS synthesis. In the formulation of the model only three different FOS were considered: 1-Kestose, Nystose, 1- fructofuranosyl nystose. A prior analysis of fermentation samples by HPLC indicated only the presence of these oligosaccharides.

2.1. Hydrolysis reactions

The hydrolysis of saccharose and FOS by β- fructofuranosidase is described by equations 2 to 5. It is considered that all the di- and oligo-saccharides can be hydrolysed by the action of the enzyme.

$$GF \xrightarrow{r1} k_1.G + k_2.F \tag{1}$$

$$GF_2 \xrightarrow{r2} k_3.GF + k_4.F \tag{2}$$

$$GF_3 \xrightarrow{r3} k_5.GF_2 + k_6.F \tag{3}$$

$$GF_4 \xrightarrow{r4} k_7.GF_3 + k_8.F \tag{4}$$

Duan and co-authors [4] proposed a Michaelis-Menten equation with substrate inhibition to represent nystose hydrolysis. In our model, this phenomenon was considered to occur also during the hydrolysis of 1-kestose and 1-Fructosylfuranosyl nystose. The FOS hydrolysis kinetic equation is then given by:

$$r_i = \frac{Vmh_{GFi} \times GF_i}{GF_i(1 + \frac{GF_i}{Kih_{GFi}}) + Kmh_{GFi}} \tag{5}$$

with i = 2, 3, 4, where r_i is the ith fructooligosaccharide hydrolysis rate (g L^{-1} h^{-1}), Vmh_{GFi} is the maximum hydrolysis rate (g L^{-1} h^{-1}), GF_i is the concentration (g L^{-1}) of nystose, 1-kestose or 1-Fructosylfuranosyl nystose, Kih_{GFi} is the substrate inhibition constant (g L^{-1}), and Kmh_{GFi} is the Michaelis-Menten constant (g L^{-1}) for GF_i.

For sucrose hydrolysis, a Michaelis-Menten equation was used, given by:

$$r_1 = \frac{Vmh_{GF} \times GF}{Kmh_{GF} + GF} \tag{6}$$

where Vmh_{GF} is the maximum hydrolysis rate (g L^{-1} h^{-1}), GF is the sucrose concentration (g L^{-1}) and Kmh_{GF} is the Michaelis-Menten constant for sucrose (g L^{-1}).

2.2. Transfructosylation reactions

The formation of oligosaccharides can occur by a transition of a fructosyl residue from one molecule to another like it was described by Duan and co-authors [4]:

$$2GF \xrightarrow{r5} k_9.GF_2 + k_{10}.G \tag{7}$$

$$2GF_2 \xrightarrow{r6} k_{11}.GF_3 + k_{12}.GF \tag{8}$$

$$2GF_3 \xrightarrow{r7} k_{13}.GF_4 + k_{14}.GF_2 \tag{9}$$

In sucrose transfructosylation, represented by equation 7, a Michaelis-Menten equation with substrate inhibition and competitive glucose inhibition was used:

$$r_5 = \frac{VmT_{GF} \times GF}{GF(1 + \frac{GF}{Ksts}) + Kmst(1 + \frac{G}{Kgst})} \tag{10}$$

where r_5 is the sucrose transfructosylation rate (g L^{-1} h^{-1}), VmT_{GF} is the maximum transfructosylation rate (g L^{-1} h^{-1}), GF is the sucrose concentration (g L^{-1}), $Ksts$ is the substrate inhibition constant (g L^{-1}) for sucrose as a substrate, $Kgst$ is the competitive inhibition constant (g L^{-1}) for glucose and $Kmst$ is the Michaelis-Menten constant (g L^{-1}) for sucrose.

Equations 8 and 9 represent the Nystose and 1-kestose transfructosylation reactions. A competitive glucose inhibition term was also included in the Michaelis-Menten equation, since Duan [4] reported this phenomenon for these two fructooligosaccharides. Equation 11 represents the fructooligosaccharides transfructosylation reaction rates:

$$r_j = \frac{VmT_{GFi} * GF_i}{GF_i + Kmt_{GFi}(1 + \frac{G}{Kit_{GFi}})} \tag{11}$$

with i =2, 3; j = 6, 7; where VmT_{GFi} is the maximum transfructosylating rate (g L^{-1} h^{-1}), GF_i is the FOS concentration (g L^{-1}), Kmt_{GFi} is the Michaelis-Menten constant (g L^{-1}) for the GF_i oligosaccharide and Kit_{GFi} is the competitive inhibition constant (g L^{-1}).

2.3. Growth reactions

The formation of biomass can either occur from glucose or fructose consumption and can be described by:

$$Y_G G \xrightarrow{r8} X \tag{12}$$

$$Y_F F \xrightarrow{r9} X \tag{13}$$

Since substrate consumption for maintenance was considered to be significantly smaller, it was neglected. The proposed Monod equations are given as follows:

$$r_j = \frac{\mu_{j,\max}.S_j.X}{S_j + KS_j} \tag{14}$$

r_j = 8, 9; where r_j is the growth rate of the microorganism (g L^{-1} h^{-1}), $\mu_{j,max}$ is the maximum specific growth rate (h^{-1}) on glucose or fructose, S_j is the glucose or fructose concentration (g L^{-1}), X is the biomass concentration (g L^{-1}) and KS_j is the affinity constant for the substrate (g L^{-1}). The Y_G and Y_F in equations 12 and 13 are the biomass yields when the glucose or fructose are used for biomass growth (g g^{-1}).

2.4. Derivation of model equations

After establishing both the reaction scheme and the kinetic equations, a general dynamical model of the process accounting for mass transfer, biomass growth and enzymatic reactions was defined. In the formulation of this model 7 state variables have been considered: sucrose (*GF*), Glucose (*G*), Fructose (*F*), 1-Kestose (GF_2), Nystose (GF_3), 1-fructofuranosyl nystose (GF_4) and Biomass (*X*).

The time derivatives of the concentration of studied components for a fed-batch bioreactor are given as:

$$\frac{dGF}{dt} = (-r_1 + k_3 \cdot r_2 - r_5 + \frac{k_{12}}{2} \cdot r_6) + \frac{F_{in}}{V} GF_{in} - D \cdot GF \tag{15}$$

$$\frac{dG}{dt} = (k_1.r_1 + \frac{k_{10}}{2}.r_5 - Y_G.r_8) - D.G \tag{16}$$

$$\frac{dF}{dt} = (k_2.r_1 + k_4.r_2 + k_6.r_3 + k_8.r_4 + Y_F.r_9) - D.F \tag{17}$$

$$\frac{dGF_2}{dt} = (-r_2 + k_5.r_3 + \frac{k_9}{2}.r_5 - r_6 + \frac{k_{14}}{2}.r_7) - D.GF_2 \tag{18}$$

$$\frac{dGF_3}{dt} = (-r_3 + k_7.r_4 + \frac{k_{11}}{2}.r_6 - r_7) - D.GF_3 \tag{19}$$

$$\frac{dGF_4}{dt} = (-r_4 + \frac{k_{13}}{2}.r_7) - D.GF_4 \tag{20}$$

$$\frac{dX}{dt} = (r_8 + r_9) - D.X \tag{21}$$

where $F_{in,s}$ is the volumetric flow rate of sucrose feeding solution (L h^{-1}); GF_{in} is the sucrose concentration on the feeding (g L); *D* is the quotient between the total feed rate ($F_{in,total}$) and the *V* is the total volume of liquid inside reactor (L).

3. Materials and Methods

Two fermentations of *Aspergillus* sp. were performed in a 5 L fermentor (B. Braun Biotech International), model Micro-DCU 200 at a pH of 5 and 30ºC, with Czapek Dox Media of OXOID and an initial sucrose concentration of 200 grams per litre. During those fermentations, the time evolutions of the 7 state variables present in model were measured. The growth was monitored by dry cell weight, where three 10 mL samples were filtered with a 0.45 micron filter and dried at 105 ºC for 20 h. The supernatant was used to determine the carbohydrates concentration. They were analyzed in a JASCO HPLC instrument with a refractive index detector using a VARIAN MetaCarb 87P column. The column was maintained at 25 ºC, and a mixture of water and acetonitrile was used as a mobile phase at 1 mL min^{-1}.

Afterwards, the collected data were used for the determination of unknown kinetic and yield coefficients with a Simulated Annealing method included in the System biology toolbox [14]. For the estimation of kinetic and yield coefficients only data from one experiment of each microorganism were used. This experiment was chosen randomly. Data from the remaining experiments were used to validate the model accuracy.
The fitting was performed by minimizing a total cost function that represents the adjustment between experimental and simulated data:

$$Total\cos t = \sum_{i=1}^{n}\left(\frac{1}{N_p}\sum_{j=1}^{p}\left(\frac{\xi_{sim,ij}-\xi_{\exp,ij}}{\overline{\xi}_{\exp,ij}}\right)^2\right) \quad (1)$$

where $\xi_{sim,ij}$ represents the simulated data and $\xi_{\exp,ij}$ is the experimental data for every point (p) for a given state variable (n) and Np is the total number of data points. The difference is divided by an average value $\overline{\xi}_{\exp,ij}$ with the purpose of attributing the same importance to all state variables.

4. Results and Discussion

The first task was to find the values for the 27 unknown parameters already described in model equations. The parameters obtained using the system biology toolbox were compared with some kinetic parameters values from literature [4], for a β-fructofuranosidase derived from an *Aspergillus japonicus*. Subsequently a simulation was carried out in System Biology toolbox using those parameters and the results were compared with the experimental data obtained from a second fermentation. The comparison between simulated and experimental data is shown in figure 1. These results show that the proposed mathematical model can predict correctly the time profiles for the state variables.

5. Conclusions

In this paper we present a detailed mathematical model for the production of fructo-oligosaccharides with *Aspergillus* sp. To the best of our knowledge, although several models representing the enzymatic reactions have been published, this is the first model that represents the fermentative process, therefore accounting for biomass formation.
Experimental data were compared with results from simulations using the estimated parameters and it was verified that the models can predict the FOS production.
Once the model is derived and validated, it can now be used for process optimization, for example for indicating which fed-batch strategy could be used to improve the production of FOS while minimizing glucose concentration.

References

1 Angeliki Kapiki, Christos Costalos, Christina Oikonomidou, Antigoni Triantafyllidou, Erini Loukatou and Vassiliki Pertrohilou. (2006) The effect of a fructo-oligosaccharide supplemented formula on gut flora of preterm infants. Early Human Development.
2 Bielecka,M., Biedrzycka,E., Majkowska,A., Juskiewicz,J. and Wroblewska,M. (2002) Effect of non-digestible oligosaccharides on gut microecosystem in rats. Food Research International, 35, 139-144.
3 Cruz,R., Cruz,V.D., Belini,M.Z., Belote,J.G. and Vieira,C.R. (1998) Production of fructooligosaccharides by the mycelia of Aspergillus japonicus immobilized in calcium alginate. Bioresource Technology, 65, 139-143.
4 Duan,K.J., Chen,J.S. and Sheu,D.C. (1994) Kinetic studies and mathematical model for enzymatic production of fructooligosaccharides from sucrose. Enzyme and Microbial Technology, 16, 334-339.

5 Fernandez,R.C., Maresma,B.G., Juarez,A. and Martinez,J. (2004) Production of fructooligosaccharides by beta-fructofuranosidase from Aspergillus sp 27H. Journal of Chemical Technology and Biotechnology, 79, 268-272.
6 Hatakeyama,Y., Takeda,H., Ooi,T. and Kinoshita,S. (1996) Kinetic parameters of beta-fructofuranosidase from Scopulariopsis brevicaulis. Journal of Fermentation and Bioengineering, 81, 518-523.
7 Kim,M.H., In,M.J., Cha,H.J. and Yoo,Y.J. (1996) An empirical rate equation for the fructooligosaccharide-producing reaction catalyzed by beta-fructofuranosidase. Journal of Fermentation and Bioengineering, 82, 458-463.
8 Lee,W.C., Chiang,C.J. and Tsai,P.Y. (1999) Kinetic modeling of fructo-oligosaccharide production catalyzed by immobilized beta-fructofuranosidase. Industrial & Engineering Chemistry Research, 38, 2564-2570.
9 Menrad,K. (2003) Market and marketing of functional food in Europe. Journal of Food Engineering, 56, 181-188.
10 Michael,B. (2002) Relationship of prebiotics and food to intestinal microflora. European Journal of Nutrition, V41.
11 Mikkelsen,L.L., Jakobsen,M. and Jensen,B.B. (2003) Effects of dietary oligosaccharides on microbial diversity and fructo-oligosaccharide degrading bacteria in faeces of piglets post-weaning. Animal Feed Science and Technology, 109, 133-150.
12 Sangeetha,P.T., Ramesh,M.N. and Prapulla,S.G. (2004) Production of fructo-oligosaccharides by fructosyl transferase from Aspergillus oryzae CFR 202 and Aureobasidium pullulans CFR 77. Process Biochemistry, 39, 755-760.
13 Sangeetha,P.T., Ramesh,M.N. and Prapulla,S.G. (2005) Recent trends in the microbial production, analysis and application of Fructooligosaccharides. Trends in Food Science & Technology, 16, 442-457.
14 Schmidt,H. and Jirstrand,M. (2006) Systems Biology Toolbox for MATLAB: a computational platform for research in systems biology. Bioinformatics, 22, 514-515.
15 Sheu,D.C., Duan,K.J., Cheng,C.Y., Bi,J.L. and Chen,J.Y. (2002) Continuous production of high-content fructooligosaccharides by a complex cell system. Biotechnology Progress, 18, 1282-1286.
16 Shin,H.T., Baig,S.Y., Lee,S.W., Suh,D.S., Kwon,S.T., Lim,Y.B. and Lee,J.H. (2004) Production of fructo-oligosaccharides from molasses by Aureobasidium pullulans cells. Bioresource Technology, 93, 59-62.

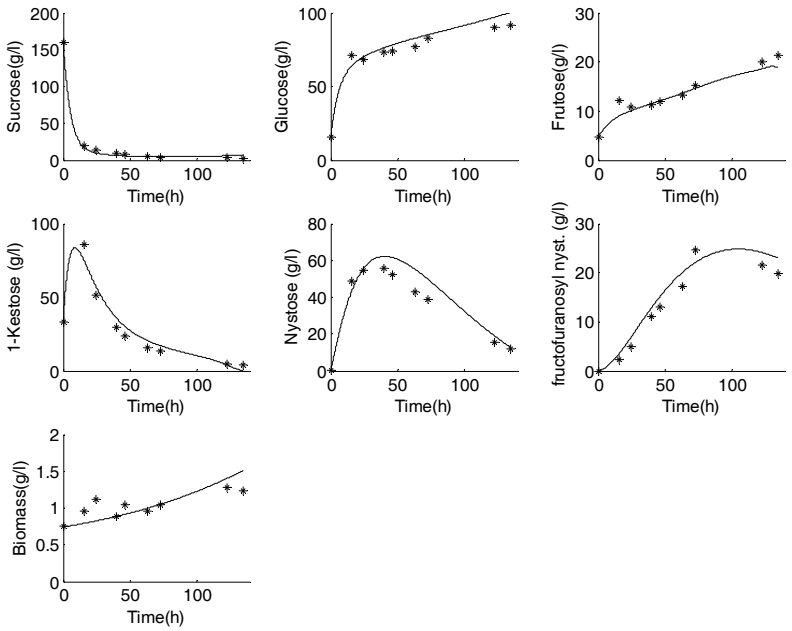


Figure 1: Comparison of *Aspergillus sp.* simulated and the experimental data from Ferm1. The simulation was performed with the parameters calculated from Ferm2.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Energy consumption minimization in bioethanol dehydration with supercritical fluids

Cecilia I. Paulo, [a] M. Soledad Diaz, [a] Esteban A. Brignole[a]

[a] *Planta Piloto de Ingeniería Química (PLAPIQUI), Universidad Nacional del Sur-CONICET. Camino de La Carrindanga Km 7, Bahía Blanca 8000, Argentina.*

Abstract

In this work, we propose a rigorous model for bioethanol dehydration process with supercritical propane to minimize energy consumption. Thermodynamic predictions are performed with an upgraded Group Contribution with Association Equation of State, GCA-EOS. As compared to the basic scheme for dehydration with supercritical fluids, vapor recompression, as well as feed preconcentration could be highly energy efficient. We further consider alternative integration schemes between process streams, associated to different nonlinear programming problems. Special attention has been devoted to a new scheme that integrates the vapor recompression scheme to the preconcentration step, which provides additional reduction in total energy consumption. We demonstrate that bioethanol dehydration can be a sustainable alternative that is energetically competitive with molecular sieves in the production of this biofuel.

Keywords: Bioethanol, dehyextraction, near critical fluid, NLP.

1. Introduction

Alternatives to fossil fuels are being investigated to reduce the world's dependence on non-renewable resources. Biofuels are currently considered as relevant sustainable technologies due to energy security reasons, environmental concerns, foreign exchange savings, and socioeconomic issues related to the rural sector. The reduction of greenhouse gases pollution is the main advantage of utilizing biomass energy. The most common renewable fuel is ethanol derived from corn grain (starch) and sugar cane (sucrose). Wood, straw and even household wastes may also be economically converted to bioethanol. However, there is need for decreasing energy consumption in the entire bioethanol supply chain to make it economically competitive with fossil fuels. Much research is being pursued on the use of lignocellulosic biomass as an attractive feedstock for future supplies of ethanol. On the other hand, downstream processes of bioethanol separation and dehydration are being studied. Karuppiah et al. (2008) have proposed different design alternatives for the transformation of corn kernels to fuel ethanol, using distillation together with molecular sieves and adsorption units with corn grits to achieve fuel-grade bioethanol. The use of pervaporation membranes has been also analyzed as an alternative to extractive distillation and molecular sieves (Hoch & Espinosa, 2008). The use of light hydrocarbons as supercritical solvents for bioethanol dehydration has been proposed as a low energy consumption technology (Brignole et al., 1987, Horizoe et al., 1993). The basic process consists of two steps, the extraction of bioethanol from the aqueous solution with a near critical solvent, and a final separation of ethanol from the solvent in a distillation train. The light hydrocarbon solvent has good selectivity for ethanol and water-solvent relative volatility becomes greater than one at the solvent recovery column (water entrainment effect). The inclusion of

different alternatives to the basic scheme has been formulated as a mixed integer nonlinear programming (MINLP) problem whose solution has provided improved energy consumption options (Gros et al., 1998, Diaz et al., 2000).
In this work, we further consider alternative integration schemes between process streams within different nonlinear programming (NLP) problems. In particular, we analyze a new scheme that integrates the vapor recompression scheme to the pre-concentration step, which can provide additional reduction in total energy consumption. Thermodynamic predictions are performed with an upgraded Group Contribution with Association Equation of State, GCA-EOS (Jorgensen, 1988; Gros et al., 1996, Ferreira et al, 2004). Numerical results show that process economics are comparable to the use molecular sieves.

2. Process Description

2.1. Basic dehydration with supercritical fluids
In a basic bioethanol dehydration process with supercritical light hydrocarbons (Diaz et al., 2000), the main units are the high pressure extractor and solvent recovery columns, as shown in Fig. 1. The supercritical fluid solvent stream enters the extraction column at the bottom, while the ethanol-water mixture is fed to the top of the column. The column is operated at conditions near the critical temperature of the solvent and at pressure above the critical. The extract (mainly ethanol and solvent) contains a small amount of water. The raffinate mainly consists of water and a very small amount of bioethanol and almost no solvent. The extract is reduced in pressure through a valve and fed to a distillation column to recover the solvent. In this column, the solvent is recovered and the complete dehydration of bioethanol is obtained, through entrainment of water by the near critical solvent. The distillate is returned as the supercritical solvent to the extractor and almost absolute bioethanol is obtained as bottoms product in the distillation unit. In this process, numerical results strongly depend on the accuracy of the thermodynamic model predictions of key phase equilibrium properties.

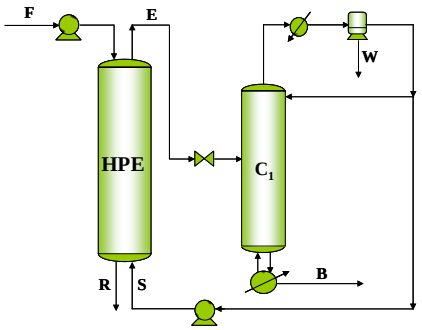


Figure 1. Basic extraction-dehydration scheme. HPE: high-pressure extractor, C_1: dehydration column, F: aqueous feed, R: raffinate, B: dehydrated bioethanol, S: solvent.

2.2. Alternative process schemes
Figure 2 shows the extraction-dehydration process superstructure that embeds process schemes analyzed in this work. The recovery of ethanol from dilute aqueous solutions can be achieved with low energy consumption by ordinary distillation, if the separation goal is the complete alcohol removal from the solution (and not dehydrated ethanol). In this case, the high alcohol-water relative volatility for dilute aqueous mixtures, allows

the increase of ethanol concentration using a simple stripping column. Therefore, *preconcentration* (PC) of the process feed reduces the flowrate of the aqueous solution to the extractor and, consequently, supercritical solvent requirements. The use of a two solvent recovery columns scheme (C1+C2) makes possible energy integration between the feed preconcentration and first solvent recovery column, such as matching the top vapor from the preconcentrator (C) to the reboiler of the first solvent recovery column (HE1). However, the stream matches depend on the preconcentrator operating pressure. There is an alternative vapor *recompression* scheme (RC), in which energy consumption is mainly determined by the compression work provided to the overhead vapor of the first distillation column (C1). The required energy in the column reboiler/condenser (HE3) is supplied by the condensation of the recompressed vapor (V_{C1}). An alternative scheme analyzed in this work is to use a turbine (TC) as driver for the compressor (RC) and integrate the exhaust steam stream to the preconcentrator reboiler (HE1). The turbine operates with middle pressure steam. No external heating services are thus required. The use of vapor recompression is justified from the low temperature difference between top and bottom at this column, achieved when a rather high concentration of solvent is kept in the bottom product. This gives low energy consumption for vapor recompression at the expense of some additional energy consumption by conventional heating in the second separating column.

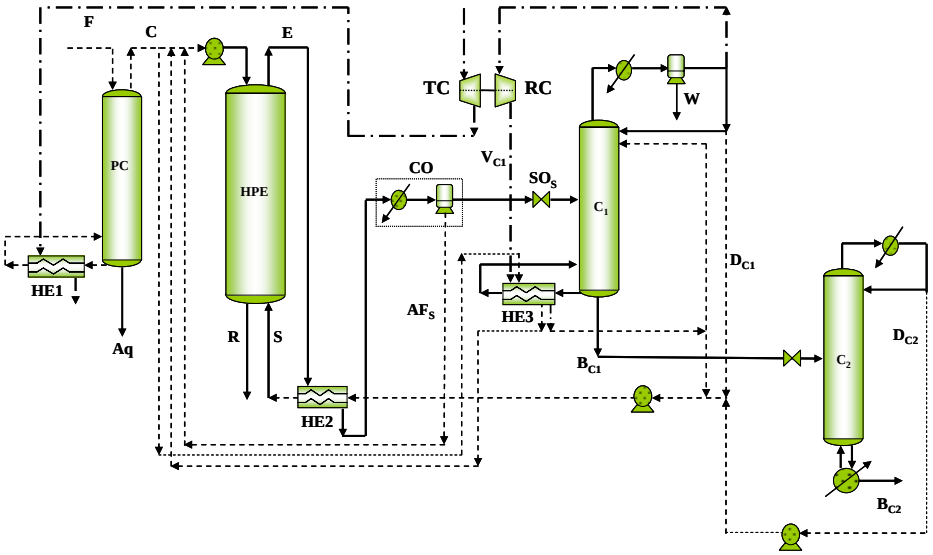


Figure 2. Superstructure for extraction-dehydration process. HPE: high-pressure extractor, C1: dehydration column, C2: second solvent recovery column, F: aqueous feed, R: raffinate, B: dehydrated bioethanol, S: solvent, CO: cooling unit, SOs: solvent organic stream, AFs: aqueous feedback stream, PC: preconcentrator, RC: recompressor, HE1: preconcentrator reboiler; HE2: solvent heater, HE3: column C1 reboiler. Full line: basic flowsheet; dashed line: alternative schemes; dashed point line: optimal scheme.

3. Mathematical model for extraction-dehydration with supercritical fluids

Different process alternatives have been formulated as a series of NLP problems. Main design variables are extraction temperature (TE) and pressure (PE), solvent flow rate (S) and reflux ratio at dehydration first column (RC1). Their bounds are shown in Table 1.

Table 1. Design variable bounds

Variable	Lower bound	Upper bound
Extractor Temperature, T_E (K)	325.00	420.00
Extractor Pressure, P_E (bar)	40.00	100.00
Reflux ratio, R_{C1}	0.30	2.50
Solvent, S (kmol/h)	45.00	1500.00

The process mathematical model (Diaz et al., 2000) includes first principles rigorous models for high-pressure multistage extractors (Kehat & Ghitis, 1981), low and high-pressure distillation columns (Naphtali & Sandholm, 1971) and a multiphase flash (Michelsen, 1982). The key thermodynamic properties of the extraction-dehydration process are based on the Group Contribution Equation of State with Association, GCA-EOS model (Skjold-Jorgensen, 1988, Gros et al, 1996) that provides reliable phase equilibrium predictions at high pressure in mixtures with association. Operating bounds and process specifications have been included as inequality constraints, as shown in Table 2.

Table 2. Inequality constraints

Unit	Constraint	Description	Bound
Extractor	r1	Ethanol recovery (%)	≥ 98.50
Solvent recovery column C_1	r2	Water composition in top vapor phase of C_1	$\leq Y_{H2O(sat)}$
Solvent recovery column C_1	r3	Ethanol recovery (%)	≥ 98.00
Solvent recovery column C_1	r4	Ethanol (solvent free basic) in bottom of C_1	≥ 99.00
Preconcentrator	r5	Ethanol recovery (% molar)	≥ 99.50
Preconcentrator	r6	Energy available from preconcentrator vapor	$\geq$ QreboilerC_1
Preconcentrator	r7	Preconcentrator vapor-reboiler C_1 temperature difference	≥ 15.00

The goal of this work is to minimize the dehydration process energy consumption to provide an economically attractive clean technology. The objective function is composed of several terms corresponding to pumping energy for liquids (solvent and aqueous feed) and heating requirement in distillation columns (PC, C1, C2), as well as their integration in the different proposed flowsheet schemes. Mechanical energy (kJ/kg) has been affected by a factor of 3.0 (Streich & Bolkart, 1982) to convert it to an amount of thermal energy of equivalent cost, so as evaluate the different alternatives on a similar cost basic. Nonlinear programming problems have been solved with a Successive Quadratic Programming algorithm (Biegler and Cuthrell, 1985).

4. Discussion of results

In this work, we consider the dehydration process downstream the fermentation step in a bioethanol plant. A typical feed to this sector is 10,000 kg/h of an aqueous solution with

10 wt% ethanol concentration. Propane is the near critical solvent (Tc=369,8, Pc=41,9 bar). Units specifications are: *extractor,* 10 stages; *first solvent recovery column*, 35 stages, pressure: 25 bar; *second solvent recovery column*: pressure: 12 bar; reflux ratio, 0,70. We have formulated four nonlinear programming (NLP) problems corresponding to the most attractive process schemes determined in previous work (Diaz et al., 2000) and a new integration scheme between turbo-compressor and preconcentrator reboiler.
Table 3 shows optimal conditions and minimum energy consumption for each scheme. As compared to the basic supercritical extraction-dehydration scheme, vapor recompression or feed preconcentration can be highly energy efficient (columns 2 and 3 in Table 3). The last column in Table 3 shows that the proposed scheme integrating the exhaust steam stream from the turbine to the preconcentrator reboiler gives the optimal structure for the high pressure bioethanol dehydration process. This option improves the integrated energy consumption in about: 29% respect to the preconcentration option, 65% respect to the recompression option and 88% respect to the basic scheme. The optimal scheme also shows lower operating temperature and pressure in the extractor, and an important reduction in solvent requirement, as compared to the remaining schemes.

Table 3 Optimal operating conditions and energy consumptions for each alternative process.

	Basic	Vapor Recompression	Preconcentration	Preconcentration + Recompression
T_E (K)	397	393.06	362.81	361.28
S (kmol h^{-1})	760	820.71	194.50	200.00
R_{C1}	1.50	1.50	0.98	0.90
P_E (bar)	100	78.93	55.00	55.00
Ethanol purity (%)	99.00	99.99	99.00	99.00
Ethanol recovery in C_2	98.00	98.00	98.87	98.65
Preconcentrator (kJ kg^{-1})	-	-	3184.98	3188.84
Column C_1 (kJ kg^{-1})	22200.00	22284.46	3735.46	3697.64
Column C_2 (kJ kg^{-1})	500.00	261.33	254.28	255.45
Pumping (kJ kg^{-1})	3500.00	2511.72	291.67	299.18
Recompression (kJ kg^{-1})	-	5.995.32	-	325.73
Integrated energy consumption (kJ kg^{-1})	26200.00	8768.36	4346.79	3082.93

5. Conclusions

We have formulated nonlinear programming problems based on first principles rigorous models with reliable thermodynamic predictions by a Group Contribution Equation of State with association (GCA-EOS), to analyze different designs for bioethanol dehydration plants, using propane as the supercritical or near critical solvent. The GCA-EOS model gives reliable properties predictions of highly nonideal azeotropic mixtures at low and high pressure conditions. Four process schemes have been optimized (basic supercritical extraction-dehydration, vapor recompression, feed preconcentration and preconcentration + vapor recompression). Numerical results show that the

preconcentrator+recompression scheme provides a significant reduction in energy consumption, making the high pressure process comparable to the use of the well-known dehydration technology with molecular sieves. We are currently including both capital and operating costs in our analysis to perform a detailed cost comparison. Nevertheless, optimization results indicate that the supercritical extraction-dehydration of bioethanol can be energy efficient, helping to make bioethanol a sustainable and economically competitive alternative to non-renewable fossil fuels.

References

L. T. Biegler, J. E. Cuthrell, 1985, Improved Infeasible Path Optimization for Sequential Modular Simulators II. The Optimization Algorithm, Computers and Chemical. Enginering 9, 257-265.

E. A. Brignole, P. Andersen, Aa. Fredenslund, 1987, Supercritical fluid extraction of alcohol frim water, Ind. Eng. Chem. Res. 26, 254.

O. Ferreira, E.A. Brignole, E.A. Macedo, 2004, Modelling of phase equilibria for associating mixtures using an equation of state, J. Chem. Thermodynamics 36, 1105-1117.

S. Diaz, H. Gros, E. A. Brignole, 2000, Thermodynamic Modeling, Synthesis and Optimization of Extraction-Dehydration Processes, Computers & Chemical Engineering, 24, 9, 2069-2080.

H. P. Gros, S. Bottini, E. A. Brignole, 1996, A group contribution equation of state for associating mixtures, Fluid Phase Equilibria, 116, 537.

H. P. Gros, M.S. Diaz, E. A. Brignole, 1998, Near Critical Separation of Aqueous Azeotropic Mixtures: Process Synthesis and Optimization, J. of Supercritical Fluids, 12, 69.

M. P. Hoch, J. Espinosa, 2008, Conceptual Design and Simulation Tools Applied to the Evolutionary Optimization of a Bioethanol Purification Plant, Ind.Eng.Chem.Res., 47, 7381–7389.

H. Horizoe, T. Tanimoto, I. Yamamoto, Y. Kano, 1993, Phase equilibrium study for the separation of ethanol-water solution using subcritical and supercritical hydrocarbon solvent extraction, Fluid Phase Equilibria, 84, 297.

R. Karuppiah, A. Peschel, I. E. Grossmann, M. Martín; W. Martinson; L. Zullo, 2008, Energy Optimization for the Design of Corn-Based Ethanol Plants, AIChE J., 54 (6), 1499–1525.

E. Kehat, B. Ghitis, 1981, Simulation of an extraction column, Computers & Chemical Engineering, 5, 171.

M. L. Michelsen, 1982, The isotermal flash problem: part II: phase split calculations, Fluis Fhase Equilibria, 9, 21.

L. M. Naphtali, D. P. Sandholm, 1971, Multicomponent separation calculations by linearization, American Institute of Chemical Engineering Journal, 17, 148.

S. Skjold-Jorgensen, 1988, Group contribution equation of state (GC-EOS): a predictive method for phase equilibrium computations over wide ranges of temperatures and pressures up to 30 Mpa., Ind. Eng. Chem. Res., 27, 110.

M. Streich, A. Bolkart, 1982, Heat pumps and ORCs can effectively compete in waste-heat utilization projects, Oil & Gas Journal, 80, 186.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Corrosion Control Document Database System in Refinery Industry

Junghwan Kim[a], Sang-Rok Park[b] and Il Moon[a]

[a]*Dept. of Chemical and Biomolecular Engineering, Yonsei University, 262 Seongsanno, Seodaemun-gu, Seoul 120-749, Korea*
[b]*GS Catex Corporation, 1056, Wollae-dong, Yeosu-si, Jeollanam-do 555-260, Korea*

Abstract

This paper focuses on techniques of improving refinery reliability, availability and profitability. Our team developed a corrosion control document (CCD) database system for refinery process. This system includes a method for defining and measuring the component risk and it also provides a powerful tool for managing many of the important elements of a refinery plant. The simulation has been applied for this system to find the critical component without direct measurement and experiment. It is possible to expect the risk parts of the refinery process.
Recent study shows the loss due to corrosion in US is around $276 billion. It's a big concern for both managers and engineers of refinery industry. Therefore it is essential to make the profitable and convenient system for refinery industry to improve reliability.
CCD database system consists of numerous parts namely damage mechanism (DM), design data, critical reliability variable (CRV), guidelines, etc. CCD database system has been developed on the basis of the corrosion control in refinery industry. It also improves the safety of refinery process and reduces the cost of corrosion greatly

Keywords: Corrosion Control Document (CCD), refinery process, damage mechanisms (DM), critical reliability variables (CRV)

1. Introduction

As an oil refinery has become bigger and accumulated recently, an accident of corrosion are more likely to be extended to serious accident. In fact, refinery process accidents have been happened continuously in the inside and outside of country. These accidents bring about many casualties and economic loss. Recently, an American report shows that corrosion cost was around $276 billion in 1998, and it was large sum accorded to 3.1 percent of American GDP of that year.
The principal case of corrosion accident in the refinery process was storage tank explosion in June 17, 2001 at Delaware. The storage tank exploded due to high-pressure gas release caused by the corrosion of connection pipe in tank. In consequence, one person who belong to repair works contractor died and eight people were wounded in this accident.
In case of Korea, the accident which took place in Oct 20, 2003 was caused by high temperature sulfidation in HOU unicracking process. Corrosion made the tube thin. With thinned tube break caused by internal pressure, distilled crude oil in tube flew out, and then the fire broke out due to heater's high fever.
Like this, corrosion accidents occur frequently in the world, and corresponding men and money loss has been serious. Therefore, systematic system is required so as to reduce

accidents caused by corrosion and to minimize corrosion cost. First of all, CCD (corrosion control document) development is necessary for development of systematic system. However, while most of internal oil companies state a few reasons such as lack of work force and time, they postpone the development of CCD.
In this study, our team developed CCD in oil refinery process and systemized development process. As following these, oil refining process as well as corrosion condition will be improved. Furthermore, on the basis of data from CCD DB, if this and currently used RBI (risk based inspection) apply to actual case together, economic effect can be achieved.

2. CCD Development

2.1. Background

In many cases, test and check of corrosion accident, in recent refined oil service as well as in petroleum chemistry equipment, was approximately recorded or it was not recorded at all. Also, information about process and arrangement which are apt to be corroded has been not well managed. Even though loss by corrosion is very small or negligible, it should be described exactly or managed well. As well, several things have to be determined, that is, which part is protected from record of loss caused by corrosion and which thing is prevented. Therefore, CCD development in refined oil and petrochemical process is needed firstly.
In this study, we developed CCD with an emphasis on CDU process in an oil refinery. Generally, Corrosion Control Document includes the following contents.

- *Inclusive document about equipment efficiency decrease*
- *Unit process & Process State*
- *Shutdown & startup condition that is related to steady-state & corrosion*
- *Process Flow Diagram, materials, corrosion diagram*
- *Corrosion control loops (Total system, Slurry system, Reflux System)*
- *Potential case of efficiency decrease in unit process and the spot where the accident took place*
- *Efficiency decrease mechanism by corrosion*
- *Corrosion control procedures, injections, inhibitors inspection, corrosion monitoring, process alteration and advice of material replacement*
- *Integrity Operation Windows (IOW)*

CCD is a document that includes information about from principle design of equipment to corrosion reason, corrosion mechanism, past record, guide line on corrosion prevention. Hence, CCD development is requisite for equipments in an oil refinery as well as petrochemical plant. Also, it is fundamental document to improve reliability of equipment and to reduce loss by corrosion.

2.2. Process

The general order of CCD development is constituted of 4 steps like following Fig.1.
The first step is to classify the quality of material according to equipment in PFD process. After that, describe DM (damage mechanism) which gives effect on equipment with each process. On the basis of that, select CRV (critical reliability variables) which

influence on inspection interval and equipment integrity. The last step is to make operation guidelines which propose proper extent of operation.
The system is needed, which can feedback the result to related team after monitoring selected CRV from online or offline, and make a report. This is called IOW (integrity operating window) system. The concept of IOW system is representing certain limit to operation guidelines of measurable operation variables (Fig. 2). With IOW system, perception of process change can assure equipment reliability by getting feedback and report earlier than condition change is recognized.

Fig. 1 Steps of Development of Corrosion Control Document (CCD)

Fig. 2 Concept of IOW system ("Establishing Integrity Operating Windows (IOW's)", John Reynolds

If it runs out of fixed value, it can make problem like devolution or corrosion of material. We developed CCD as 4 steps. Important content of each step is following.

2.2.1. Step 1: Building a MSD (Material Selection Diagram)
Considering corrosion and economic efficiency, each equipment of oil refinery and petrochemical plant is made from different material. For instance, carbon is used in upper portion of A-column in CDU process, but lower portion is made from Hastelloy that hardly ever corrode. Compared to upper portion, such corrosion materials as asphalt and coke in lower portion are comparatively less pure, and these flow more through lower portion than through upper portion. Therefore, using existing materials, carbon, may cause several problems.
At first, we drew the Process Flow Diagram of each process. Generally, one PFD includes following information.

- *Process Piping*
- *Major equipment items*
- *Control valves and other major valves*
- *Connection with other systems*
- *Major bypass and recirculation streams*
- *Operating data (temperature, pressure, mass flow rate, density, etc.)*
- *Process stream names*

On completed PFD, We make MSD which reflect all the past histories of material use. Based on design data of corresponding process and the past history, we indicate such the qualities of material as Cr, Duplex, 400SS, Hastelloy, and 300SS on PFD. While

making MSD, we marked the spots where the probes have already been installed and where we need to install other probes.
Following Fig. 3 shows MSD of A-column in CDU process.

2.2.2. Step 1: Building a MSD (material selection diagram)
After building a MSD, there's a need to select a Damage Mechanism. The corrosion materials and the mechanism should be found during this step.
First of all, the design and the operation data should be collected, and then check up the temperature, the pressure and heat treatment conditions from collected data. Once finishing collection data, it's necessary to collect the corrosion material data of each steam. What should be collected is sulfur, chloride ion, hydrogen sulfide, ammonia and so on. The followed step is selecting a DM related to an equipment or each part of equipment refer to API(American Petroleum Institute) 571.
We selected 9 DM out of 19 DM on API regarding to possibility of occurrence at current processes, Creep/Stress rupture, Fuel ash corrosion, Oxidation, High temperature sulfidation, Naphthenic acid corrosion, Hydrochloric acid(HCL) corrosion, Ammonium chloride(salt) corrosion, Wet H2S corrosion and Ammonia stress corrosion cracking.
There are the specific explanation of selected 9 DM, affected materials, main factors, the methods of prevention & alleviation and the ways of inspect & monitoring.

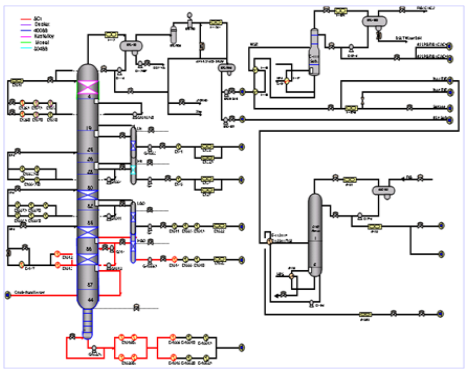


Fig. 3 MSD of CDU process

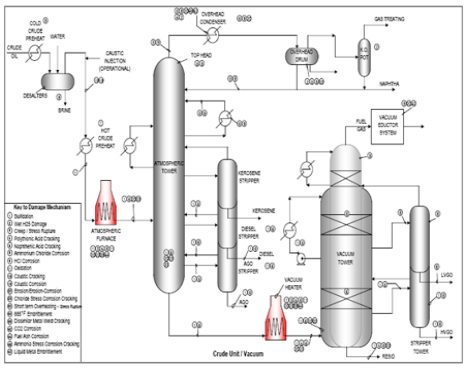


Fig. 4 DM of Equipment (API 571)

We describe the DM according to equipments and more focuses on parts of equipment. For example, the upper part of A-column has HCL corrosion and Salt corrosion. So are that the middle part has Salt corrosion as well as the lower part has Naphthenic acid corrosion and High temperature sulfidation.

2.2.3. Step 3: CRV (Critical Reliability variables) Selection
After selection the DM, the next step is choosing key parameters that influence DM severity as Critical Reliability Variables.
Main variables are PH, Chloride ion, Hydrogen sulfide, ammonia, TAN and etc refer to Best Practice, API 571 and API 580 & 581. As it's already mentioned, the upper part of A-column has HCL corrosion and Salt corrosion. The influencing factors on HCL corrosion are Water PH, Water CI-, Water Fe, Water H2S, Water NH3 and

Temperature. In addition, measuring temperature has to be always done, whereas the others could be inspected 3times a week.

2.2.4. Step4: Operating guideline proposal

The last step is to introduce a CRV operation guideline which has no influence on the reliability of the equipment (references: API 571 and API 580 & 581, Chemical Vendor Guideline).

Fig.5 illustrates the final developed CCD. It is essential to set the boundary of main variables which affects each equipment so as not to damage reliability.

For instance, PH level should be 6~7 at CRV of HCL corrosion, Cl^- must be lower than 30ppm, so is that Fe is lower than 1ppm. This guideline is the basic data of an IOW (Integrity Operating Window) system. The IOW system automatically makes a report every time each variable goes beyond their upper and lower bound. To keep the reliability of the equipment, it's required to pay special attention to the variables which goes beyond their limits too many times.

Fig. 5 Corrosion Control Document

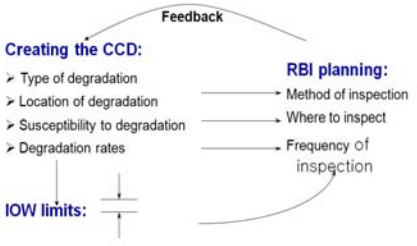


Fig. 6 Relationship among CCD, IOW, RBI

3. Result and Application

3.1. Application to RBI

CCD development becomes a basic data for RBI(Risk Based Inspection) which deals with the priority orders, methods of equipment examination as well as cycle by conducting quantitative analysis about equipment troubles and accidents caused by corrosion.

Unless there's enough data of corrosion factor, it is not possible to make full use of RBI which aims control obsolete equipment. Therefore, there's a need to apply CCD DB to RBI (Fig.6).

That would be able to figure out quantitative risk due to corrosion. Moreover, it could not only improve profit where companies have the methods of risk inspection to reduce the risk through minimum costs and endeavor, but also enhance the safety of process equipment through developing decision-making guideline based on RBI system which has different inspecting cycle and methods refer to the priority order of equipment risk.

3.2. Risk Monitoring and Process Simulation

We could achieve some side effects such as prevention against fatal accident which causes casualties and poverty damage by development of risk monitoring and analysis

algorithm. If we hope to development this system, we should define the important factors which bring about a lowering efficiency such as the continuous flames on pumps, valves, generators, ventilators or heat exchangers.
It's possible to control equipments with significant effects through process simulation. For example, the development of a corrosion scenario and simulating make an accurate estimate of risk linked with probability of accident. Furthermore, it makes reducing unexpected interruption of plant. On top of that, as we've pointed out, the optimal operation guideline with estimation of high risk brings a number of advantages such as reducing unexpected shutdown and downtime, boost output, improvement in quality, executing the optimum equipment and innovation in safety management system.

4. Conclusion

This study focuses on the development of CCD to enhance reliability of the whole process equipment as well as to operate smoothly in chemical corrosion. CCD is developed by building a MSD, selecting DM, CRV and operation guideline proposal. We sincerely believe that CCD could improve the reliability and the safety of the whole refinery industry, and so is sharing information with each department of company. The further step is IOW system could be coupled with RBI system. Besides RBI, we could reduce the risk through simulating process and risk monitoring.
Especially, building CCD is terribly important as the first step for the reliability of equipment and the safety of a plant. And IOW, RBI and Risk Monitoring become the fundamental data of subsequent research.
Finally, we could improve the reliability of the whole refinery, petroleum industry by applying developed CCD to FCC (Fluid Cracking Catalysis), HOU (Heavy Oil Upgrade), LPG and aromatics process due to CCD based on CDU process.

Acknowledgments

This work was supported by MKE (Ministry of Knowledge Economy) under the program of ETI (Energy Technology Innovation). This paper is a result of "Research Group of Energy Safety for Next Generation". (Project No 2007-M-CC23-P-08-1-000)

References

G.H. Koch, M.P.H. Brongers, N.G. Thompson, Y.P. Virmani, J.H. Payer, "Corrosion Costs and Preventive Strategies in the United States", FSP & SFPE, 2003.
M.S. Ray, M.G. Sneesby, "Chemical Engineering Design Project: A Case Study Approach", 2nd Edition, Gordan and Breach Science Publishers, ISBN 9056991361, 1998.
API(American Petroleum Institute), "API RP 571 Damage Mechanisms Affecting Fixed Equipment in the Refining Industry", 2003.
J. T. Reynolds, "Risk-Based Inspection - Where are we today?", CORROSION 2000, NACE International, Paper No. 00690, 2000.
J. T. Reynolds, "The Application of Risk-Based Inspection Methodology in the Petroleum and Petrochemical Industry", ASME, PVP-Vol. 336, p.125, 1996.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Methodology of Pipe and Equipment Layout for On-Shore Oil & Gas Industry

Luiz G. Persson[1], Flávio B. Santos[1], Cesar A. C. Tavares[1], and Arnaldo E. de Andrade[1]

[1]Chemtech – Departament of Pipe design, Rua da Quitanda, nº:50, CEP 20011-030, Rio de janeiro-RJ, Brazil.

Abstract

Traditional process engineering methods in the petroleum industry are based on the experience of engineers' and the well-succeed known cases. The maintenance of the facilities and design methods broaden the scope of traditional methods, but focus primarily on short time and simple construction ways while ignoring the best economic pipe and equipment layout. This paper presents a work method developed during the project of a new refinery located in the northeast of Brazil called RNEST licensed to Chemtech Company. Matching the study of the basic project with the petroleum chemical process it is possible to achieve a new approach which considers inputs from both maintenance and construction design attending the normative security rules. This approach was developed from the study of a small facility of RNEST, the caustic treatment unit, where the reduced number of equipments and pipes made the method possible to be done in a fast an understandable way. The preliminary study involves a review of the general refinery plant and the chemical processes of the facility, that way it is able to rearrange the equipment layout in order to optimize structures and pipe layout. Combining different disciplines and analyzing every sub product inside the unit, the pipe modeling work becomes more efficient and it improves the performance of the system. Together, pipe and equipment analysis can provide insight into the economic field, physical space and sustainability of the industrial process. The proposed method is broadly applicable to assist decision making in petroleum and chemical engineering tasks.

Keywords: refinery, equipment layout, pipe arrange.

1. Introduction

Plant layout design plays an important part in the design and engineering phases of any industrial facility. This paper discusses the role and responsibilities of the plant layout designer, provides advice on how to use project data and mainly develop a methodology of equipment arrangement design and piping layout design.

2. Methodology

In almost any engineering project gathering the basic information is crucial for a successful work. However, there are some unique features in each project, especially how this information is supplied. In a refinery, the pipe layout design requires a good number of personal resources for its full achievement, therefore, a good study at the beginning of the project can minimize redundant work and save precious hours in some cases, as if any equipment has to move from its original position because of bad

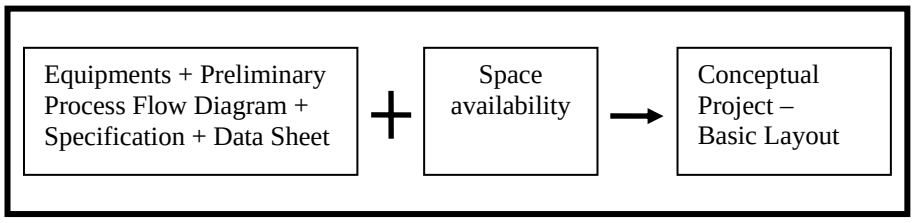
equipments location study or save space through a compact layout of all units and their interconnections.

For the conceptual project the first step is analyze the plant production capacity and also parameters there are directly connected with it. Some o those factors as electrical power available in the location, natural resources, water supply and qualified workers, rarely are taken into account, but for a good venture they are extremely important. There is also the local logistical capacity for import and export goods and the frequency of local predominant wind that certainly will influence in the final result.

In the case of pipe arrangement, the process plants (on-sites units) are the most important variables to be analyzed. In those plants are the pipes with high complexity of connections within a short area for construction, this fact requires a detailed analyses of equipments displacement in order to make a good pipe arrangement. In the off-site are the main distribution pipes, they are much bigger than the pipes placed on site, but these pipes have less complexity and their only function are connect the plants of on-shore oil and gas industry. In most of the projects the great challenge of the piping and equipment layout team is to make a layout preventing all piping difficulties such accesses to operation, predict maintenance area and plan access to build the pipes. Commonly those topics are brought into project in the pipe design phase but they could be studied along the equipment design phase in order to perform a better pipe layout and optimize the facility area.

The beginning of the main project is the study of the process plants stats with the descriptive memorials (MD's), supplied by the process team. It is also necessary the analyses of the equipments data sheet, together, these documents describe preliminary equipment sizes and inlet and outlet products and flows. The MD's describe the whole unit process, the function of all equipments and the main fluids pressure and temperature at each process stage. The equipment layout interacts directly with these documents; the MD's shows the fluid flow work and show the order that the fluid must pass through the equipments making easier the layout task. Many systems are described by this document and their interactions with others systems such blow-down, pup-out, compressor, water distributions and others.

As explained, the MD's gives a macro view of operation functionality of the refinery, a deeper analyzes should be made using the Process Diagram. With these documents information adding the reserved area for the plant, the preliminary layout of equipments can be done. A well made work at this level means a compact layout reducing the needed area with the most important pipes previously planed. This means space and pipe material economy.



Besides the technical demand, an engineering project should also respect the security and environmental regulation rules. A good equipment and pipe layout should provide fewer costs within maintenance and constructions and also ensure less cost with insurance policies.

Frequently the main company of the venture has its owns regulations and standards for an initial approach and support for the project, some regulations and standards may belong to the company but there is others who was created by a regulation institute, mainly they are set on the contract and must be followed by the project company.
Most of the times the norms are preventatives and their main purpose is orientate the designer, in some cases the designer should choose different types of solution in some appropriate occasions.

2.1. Equipment study

The equipments are displaced at certain logic order to minimize piping materials and the area needed to build the plant. However, its elevation must be considered too. The information needed to study and calculate the equipments elevation are found at their data sheet, pipe and instrumentation diagram and pipe list, which is a document that lists all pipes and their characteristics such as fluid temperature and pressure.
This study has different objective from the pipe rack study. Here, the objective is to indentify the pipes that require certain elevation to work with the equipments and by this work one can determinate the elevation necessary to all pipes to work well. Since it's impossible to think at the same time in all the structures, valves, buildings and equipments that maters in the elevation of a single equipment a lateral view pipe sketch of the equipments and the most important pipes that connects it is a good way to study its elevation.
This kind of pipe sketch works in a very similar way to the pipe sketch presented previously. It's important that this sketch incorporates information about equipment nozzle elevations, building near the equipments and equipments that works together with the studied equipment, with that ready then the important pipes must be drown. The designer should also consider that some equipment has a minimal elevation defined by the process team, and since all equipments are connected they need to respect a certain elevation between then. For example, one pipe that drains fluid at low pressure from one vase or tower and head to the pump must have the drain point higher than the pumping point; in other words, the fluid must flow by gravity.
As said before the basic project or the process team determines a minimum elevation for the equipment, and this study is responsible for determining if this minimum is enough or not. All structures and buildings in a plant are always associated with a certain equipment, defining its elevation is important, so the project of these structures can be started by other teams in the project.

2.2. Pipe Sketch

The complexity and the high number of pipes that are required into a petrochemical plant make too complex analyzing one system or determinate the structures and buildings dimensions. These analyzes must include all or almost all the pipes that pass at the studied area and depending on the quantity and the complexity of these pipes the designer imagination may not be sufficient to consider all variables. To make these analyzes we suggests a methodology presented in this paper as the pipe sketch, which consist to draw the pipes, on the layout arrangement drawing, in 2-D representing them by a line. This drawing can be simple and not precise with their length.

2.3. Pipe-rack Dimensions

In a refinery the number of pipes is too high and the pipe rack is constructed in order to support almost all pipes and it's also a way to organize the pipes grouping them all. To determinate the necessary pipe-rack dimensions and avoids area and material loss the pipes passing trough the pipe-rack must be studied. This study must consider the critical section of the pipe-rack, which means the section that has more pipes to be supported; to determinate this section the designer must make a pipe sketch of all pipes that he intends to pass trough the pipe-rack. After that the designer must locate all pipes in the section and calculate the space necessary .On our case study we added 30% of free space for further plant improvements or unanticipated pipes. When finished the designer must decide which way he can build the pipe rack, high and thin or thick and low, this choice depends on the available area.

This is for sure an arduous work, but avoiding this study make the chances of failure grow. An overestimated pipe rack means material and area loss and an underestimated one means a probable project failure.

2.4. Maintenance and building analyzes

With the basic layout done, some important points must have special attention at this moment to improve the quality of the project. To complement the equipment layout the point showed below have to be studied:

- Access to valve and instrumentation operation
- Free area to equipment maintenance
- Platforms to access and operate all needed equipments and instruments
- Emergency exits
- Area and accesses for building hoist and other building equipments

These topics has to be thought together with the all others project topics, but its essential to have an specific time just to check if these topics are correctly applied in the project at each system, each equipment and each valves control station.

The first system to be analyzed should be the pressure relief system; most of PSV's (pressure safe valves) have to be above the flair line (the header of all relief pipes). Since the flare may be at the last level of the pipe-rack it's convenient to place a platform along all pipe rack top to operate and place the PSV's control stations. These platforms can be connected with other buildings and equipment platforms, in order to make easier the accesses to the higher level of many buildings. Depending on the available area, constructing platforms at the side of the pipe rack at medium level can be a good solution since the designer can place control stations, instruments and valves that don't have space to be placed at the floor on these platforms.

Security is very important to any industry, in oil and gas industry isn't any different; respect the security rules are many times very expensive, but not attending these rules can cost much more, so at any project one must attend all security rules that apply in each case.

Access in all plant frontiers can be useful for building, maintenance and secure the plant. These accesses are important to be large enough to support the transportation of the big equipments such as reactors, towers, compressors, pumps and vases.

In the pipe design of the plant it is important to think if the pipe project is possible to be build verifying the equipments already located and buildings near the pipe. Some pipe project has welds that are impossible to be done, so the project must have these cases

reduced to the minimal possible avoiding remaking the pipe project or difficulties in the building field.

3. Case Study

As it could been seen on the Abstract, the methodology above was developed through a refinery project in the northeast of Brazil. Among all problems solved in the Project, the main focus was use a reduce area and make a compact pipe layout. There are many ways to make a compact layout but many of them may not be fully functional, resulting in difficulties for building and operate the plant.

The refinery project is ambitious and its main objective is to take advantage of the market available at the Brazilian northeast region and supply the Brazilian demand (that currently import diesel at higher prices, mostly because of the transportation and refining cost which increase the final product cost), maintain a strategic position at European market, which have a high diesel prices and a big demand, and finally to attend international quality requests such as low sulfur and other pollutants.

In this case study, we will present two plants of the refinery, were we have change the basic layout in order to have a performance improvement. The main alterations from the basic project done were: pipe rack dimensions, new access to the plant and equipment location.

The case study starts with the MD's and process diagram study and with the basic project on hand. Then the layout was modified to a more compact form respecting the norms and regulations required by the owner. The frame below (Figure 1, Figure 2, Figure 3, Figure 4) shows the after and before layout for the plants studied.

Together with the layout comes the pipe-rack dimension study from both plants. In unit I within the pipe sketch we determinate that the pipe-rack was under provisioned for the amount of pipes it should support. The total length was 24 meters divided in 4 levels. Taking the most critical session of the pipe rack from the pipe sketch we could set the dimensions to 50 meters at least without consider possible expansion or unanticipated alterations that could require more space.

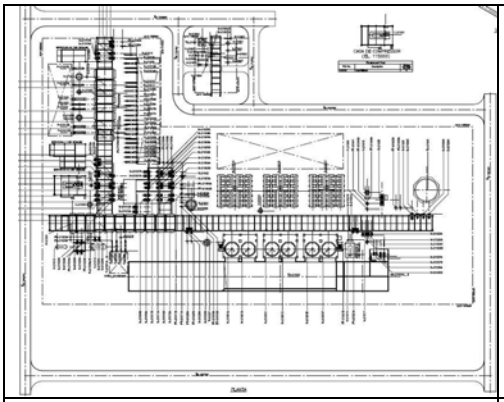


Figure 1 –Basic equipment layout Plant – Unit I

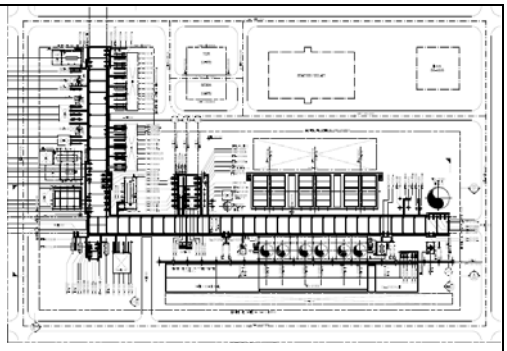


Figure 2 – Final equipment layout Plant – Unit I

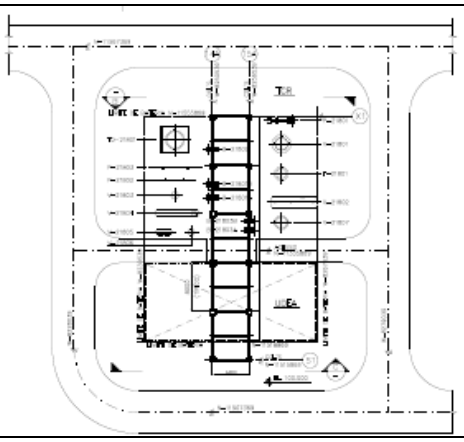


Figure 3 – Basic equipment layout Plant-Unit II

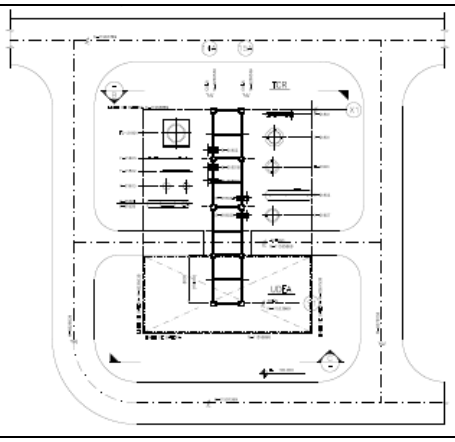


Figure 4 – Final equipment layout Plant-Unit II

The operation platforms for pressure safety valves are located in the top o the pipe-rack. This way the pipes that contain valves don't need to reach the floor and find space. With this platform pre-projected one can design a better pipe arrangement wasting less pipe components. A similar study was done for each equipment in order to develop a compact pipe layout with easy access for maintenance and operation.

4. Conclusion

It could be seen that plant layout design plays an important part in the design and engineering phases of any industrial facility. The development of equipment arrangements and piping layouts for process off-shore industries offers an opportunity to demonstrate technical ability along with a creative talent and common-sense approach to problem solving. Process facilities must be designed and engineered within extremely short schedules while adhering to maintenance, safety, and quality standards; moreover, the design must take constructability, economics, and operations into account. Although the tools to achieve these goals are changing from pencil and paper to computer graphics terminals, the responsibilities of the plant layout design remain the same.

Each plant layout develops an individual layout philosophy. Although conditions (client specifications, schedule constraints, and availability of information) may change significantly among projects, if the design style philosophy remains the same, the probability of a successful project is very high.

The main objective it is to keep in mind all possible variables to optimize the layout design project.

References

Z. R. Hunt, 1993, the Basics of Plant Layout Design, Process Plant Layout and Piping Design, PTR,

Y. P.C.S.Telles, Instalationd and Construction Industrial Dsiposal, Industrial Pipe Design, LTC, 10 Edition

X. ASME, 2006, Standards for Piping Components , Process piping, Code for Pressure Piping B 31.3, HIS.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Chemtech's Energy Efficiency Methodology

Renata Machado[a], Andreas Hahn[a],Rafael Teixeira[a], Flávio Waltz[a], Valter Souza[a]

[a]*Chemtech – A Siemens Company, Rua da Quitanda, n.50, Rio de Janeiro-RJ, CEP 20011-030, Brazil*

Abstract

There would be no lack of energy if all the existent technologies were applied for a more efficient energy use, said Greenpeace – Brazil. Aiming at an energetic sustainability of his clients, Chemtech developed a methodology that primes for the conscientious use of energy, contributing to the reduction of industries' energy consumption. This methodology is based on the traditional energy audit, consisting of an initial assessment, data survey and measurements, resulting in actions to reduce overall energy consumption, allied to a differentiated implementation. The audit embodies a tariff analysis, correction of power factor, pre feasibility study of mechanical and electrical measures and process analysis. A conceptual project is elaborated on selected measures (i.e. automation, production, electrical installations and utilities) after a detailed feasibility study. Differing from a traditional audit, a client's energy management team is established aiming the development of an internal control of energy, with the implementation of an energy efficiency awareness program for all employees and controlling the energy consumption through established energy indicators. Another differential of Chemtech's work is the approach used to develop the audit. Each milestone of the audit consists of several steps, each step used to augment the consistency of gathered data and refine obtained results through recurring meetings with the client covering specific topics. The next step is only initiated after the client's acceptance of data and results pertaining to the previous one. This approach raises the consistence and acceptance of the developed work, like a helix, guaranteeing the overall development efficiency, so as the consistency of the project.

Keywords: Energy efficiency, energy audit, energy management.

1. Introduction

Brazil have been making efforts on energy conservation since 1985 with the creation of the Programa Nacional de Conservação de Energia – Procel (National Energy Conservation Program), focused on electricity. This search for energy efficiency culminated with the creation of the Programa Nacional da Racionalização do Uso dos Derivados do Petróleo e do Gás Natural - Conpet (National Program for the Rationalization of the Use of Oil and Natural Gas Derivatives) in the year of 1991. Nowadays, with increasing concerns about the future of energy availability, energy supply security has been spotlighted in the political debate around Brazil. Not counting the global climate change, caused by the increase of green house gases on earths atmosphere, in detriment of ecosystems, causing stronger and more frequent storms, melting of glaciers, heat waves and droughts. Questions like economical and populational growth, beside energy intensity are some of the key variables responsible for the increase of energy demand.

In this context the energy efficiency activities become usual in industrial plants, buildings and even homes. Nowadays most of brazilian energy conservation companies (ESCOs) attained plenty knowledge about how to develop energy audits and implement selected measures on their clients' plants. These are the main activities developed on the brazilian energy efficiency market so far.
In spite of the high initial impact of pure technical energy efficiency measures, these lose part of their effectiveness along the time if no constant monitoring of their performance is done. For this reason and that the human factor influences profoundly the energy performance of a company, Chemtech opens a second action front through the energy management, obtaining synergy with the proposed technical measures.

2. Objective

This work intends to present the energy efficiency methodology develop by Chemtech, focused on two major fronts: one aimed on the technical and process approach. The other one aimed on the development of a continuous client side energy management. Both fronts are associated with an integration methodology widely used by Chemtech, directed to the client's feedback.

3. Energy Audit

3.1. Workshop
The first step for the development of an energy audit consists of a workshop, involving client's employees of different interest areas (automation, maintenance, operation, among others). The objective of this first workshop is to gather initial information for a preliminary energy report. The survey is mainly focused on:

- assessment of the production vs. energy flow on the site;
- identification of the main energy consumers on the site;
- identification of the main energy loses;
- identification of bottlenecks;
- determination of the energy audit's foci.

Most of the time the client's employee's opinions, due to their large experience on field activities, provide information regarding energy consumption reduction, being of great value during the planning of the next steps.
During this first contact a communication plan between the companies is established, providing an efficient link able to supply the communication needs during the project development.
At the end of this stage the team should be able to tell the main bottlenecks of the production processes and some savings opportunities.

3.2. Preliminary Energy Report
During the development of the preliminary energy report three activities are developed: on field assessment, assessment of maintenance needs and assessment of operational needs.
During on field assessment the main energy consumers are verified. Internal and external lighting systems, compressors, pumps, ventilators, motors and other equipments with energy consumption reduction potential are analyzed. Last but not least, the utility consumptions are focused: compressed air, HVAC and water. An overall analysis of the electrical structure of the site is also initiated.
Most of the time, the lack of maintenance of some equipment, due to the lack of maintenance control, turn these obsolete, inefficient or inadequate for use, causing a

higher energy consumption than necessary. In this case the use of good maintenance management practices leads to the reduction of energy and utilities consumption, sometimes with unexpected high savings.
The assessment of operational data is an essential activity for the further development of the energy audit. At this stage the energy costs, production and consumption reports and process documents are evaluated.

3.3. Energy Audit
During the development of the energy audit the on field assessment is terminated and an evaluation of the main potential energy savings is made. For each potential measure the monetary savings are calculated and a technical and economical feasibility study is developed. Then a second workshop is necessary to define, together with the client, which measures should be detailed for implementation based on the results of the feasibility study. At this workshop the proposed measures are classified in short term and long term projects, last ones leading to conceptual projects.
At the end of this stage the client has a definition with all selected actions to be implemented associated to their energy savings potential.

3.4. Short Term and Conceptual Measures
The short term measures are immediately implemented by the client with support of Chemtech's energy efficiency team, to assure the expected savings.
For the long term measures conceptual projects are elaborated aiming the development of future basic and detailed projects for implementation.

4. Energy Management

In spite of the great initial impact of pure technical measures, at long term, this approach loses his effectiveness if no further follow-up is undertaken. Since the human factor is essential for the continuity of the implemented measures, Chemtech developed an energy efficiency management methodology to guarantee the endurance of implemented technical measures and the development of new ones, and also to develop a sound client-side energy management structure. This methodology involves the assessment of the clients' administrative and communications structure, involvement of different sectors of the companies, application of specific questionnaires, training of the members from the newly created or already established energy management staff to improve their energy efficiency and management skills, including monitoring, targeting & reporting, and the creation of an energy efficiency awareness program for all employee levels. Key performance energy indexes to manage results are customized attending the client's needs, based on traditional indicators and on the cumulative sum of differences methodology, usually used in the statistical process control, but now applied for monitoring the overall energy consumption performance between well defined limits.

4.1. Methodology
Chemtech's energy management methodology is a blend of different methodologies published mainly by Natural Resources Canada and ETSU (UK) and years of experience of its energy efficiency professionals. The core of the methodology relies on an awareness program implemented together with monitoring & targeting of key indices defined together with the client.
Experience has shown that several steps are necessary for the implementation of a successful energy management program, each of them discussed below.

4.1.1. Joining the Players

This step joins the players necessary to maintain a sound energy management program. First of all a formal commitment from the senior management is necessary. No formal commitment usually leads solely into an implementation attempt, with no expressive returns. Also an energy management sponsor has to be defined together with an energy management team. An energy management leader is assigned with powers and authority to develop company's energy policies.

4.1.2. Identification of Opportunities

Together with technical opportunities, pointed out through an energy audit, human resources have to be evaluated, mainly through interviews and questionnaires. The results obtained from the human resources' feedback will indicate needs for awareness or specific training. Human resources are essential for the success of any energy management program; special care has to be taken during this step assuring that most of the employees needs for improvements are listed.

4.1.3. Establishing Targets

Targets are set not only for energy consumption levels, but also for the awareness and necessary skills of the employees. Human resource needs have been unveiled during the identification of opportunities and at this step a new level of knowledge and awareness is defined to be reached in a certain amount of time.

Energy consumption indices are determined together with the client and these will be monitored to be maintained between well defined limits. One of the main tools used during this step is the cumulative sum of differences methodology, derived from the statistical process control. Energy consumption and production data are correlated and evaluated; the best historical performance is selected and set as desired target. Since this energy consumption performance already had been reached once, no objections can be made about it. Later better energy performances will be achieved and set as new targets during the program evaluation.

4.1.4. Communications Plan

Mostly not given value to by technical personnel, communication is an essential part of energy management, since most results proceeding from it are achieved from the company's employees.

At this step existing communication lines are assessed, target audiences are defined, messages developed and communication tools identified or created. Sometimes energy efficiency events are planned at this point. A great amount of time is invested in this step, since a communication failure usually leads to poor energy management results.

Client's communication department is to be closely involved, working together with the energy management team, designing strategies for the following months.

4.1.5. Training Plan

In parallel with a communication plan a training plan is developed based on the results of the identifying of human resources opportunities and needs. The plan is designed to achieve the targets set for employee's skill and awareness levels and the training can be delivered in house or by third party. Usually the energy management leader, together with the energy team, defines the training needs, but not necessarily delivers the courses themselves. The program evaluation indicates if the desired skill and awareness levels have been achieved.

4.1.6. Energy Management Implementation

The energy management program is formally implemented only after the steps above had been fully developed. The implementation is preceded by a communication

campaign raising employee's awareness and preparing the company for the monitoring and targeting to be initiated. At this point usually a formal commitment from senior management is published.

4.1.7. Program Evaluation

After implementation the energy management program has to be constantly evaluated, adjustments have to be made to targets and to the overall management plan. New interviews and questionnaires help to evaluate the success of delivered training courses. Increasing awareness leads to more feedbacks, which, together with the necessary adjustments, creates a continuous improvement loop.

Energy performance reports are send at defined intervals to decision makers and to energy customers allowing them to follow the consumption trend and make necessary adjustments to achieve the desired consumption levels. Deviations from desired consumption levels are investigated and adjustments made to achieve targets or new target levels are set.

4.1.8. Sustainability

To guarantee the sustainability of the energy management program some activities are necessary to maintain a high level of response.

New employees have to attend training sessions about energy efficiency and the company's energy management program basics to level up with the rest of the client's employees. Recurring communication messages have to be developed to reinforce the awareness, taking care not to saturate the audience. Success has to be shared recognizing success and pointing out exceptional achievements.

Last, but not least, the whole program has to be evaluated from time to time, being adapted to be up-to-date with the company's needs.

5. Chemtech's Integration Methodology

During the whole development time of preliminary energy reports, energy audits and even the implementation of an energy management, aiming the best roll out of the project management, Chemtech applies its integration methodology.

This methodology consists of recurring meetings with the client during the development phases discussing obtained data and results obtained so far. At each meeting existing data and results are reevaluated and necessary corrections are made, like a spiral, each beginning "coil" of the spiral representing a meeting and the "height" of the spiral representing the growing knowledge and accuracy.

Chemtech's integration methodology keeps the client informed about the development at different phase levels, obtaining his necessary confidence, participation and support, managing client's expectations and avoiding unnecessary rework at the end of the project. This approach has been largely appreciated and praised by Chemtech's clients.

6. Conclusion

The traditional energy efficiency approach through a purely technical energy audit with further implementation of energy saving measures, although with high initial impact, loses effectiveness over time if no control is undertaken to assure continuity of the implemented actions.

Although mostly neglected by technical personnel, communication is a fundamental part of the success of any energy audit or energy management program. Chemtech's integration methodology is designed to improve the communication with its clients during development and a sound communications plan assures best results through an implemented energy management program.

Allied to an energy management program the results obtained through technical measures can be maintained and even increased due the interaction with company's human resources. The employee's skills and awareness can be increased leading to a sustainable energy management.

References

Saving money through energy efficiency – A guide to implementing an energy efficiency awareness Program, Natural Resources Canada, 2004

Guide 112 – Monitoring and Targeting in large companies, ETSU, 1998.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

New Model to Determine Fracture Gradient in Ultradeep Water

Clovis D. Ferreira,[a] Wilson da Mata,[a,b]

aPós-Graduação em Ciência e Engenharia de Petróleo, Universidade Federal do Rio Grande do Norte, Campus Universitário, Natal-RN 59078970, Brasil
bDepartamento de Engenharia Elétrica, Universidade Federal do Rio Grande do Norte, Campus Universitário, Natal-RN 59078970, Brasil

Abstract

The development of the offshore drilling in ultra deep water requires additional care in function of some critical points. In the beginning of drilling, just below the bottom sea level and the transition zones where pore-pressures tend to the formation fracture pressures. With the increase of the depth of water in relation to depth of sediments below the bottom of the sea, it happens the reduction of the operational window that it is the available space to work with the density of the drilling fluid, hindering the progress of the drilling. To minimize the problem, it is proposed in this work a model that uses the superior limit of the operational window (fracture gradient) equal to the value of the leakoff test available of the area in the depth considered to use as maximum value for the gradient of mud. The calibration adjusts the fracture gradient made calculations to the leakoff test. In this situation, it is been in favor of the safety, being avoided fractures of the formation with loss of circulation or kick. The proposed model is a geometric correlation that uses values of pore pressure gradient and overburden gradient of another indirect model and executes the calculations using a relationship in which the involved parameters vary in the same proportion.
The application of the proposed model implies in costs reduction in the definition of the casing shoes depths, in the casing sizing reduction, in a better cementing design and also in a more adequate mud weight for drilling all the well phases. The economic return is more significant in scenario of ultra deep water.

Keywords: fracture gradient, overburden, pore pressure

1. Introduction

The progress in the search for news offshore ultra-deep waters oil reservoirs requests from petroleum industry high investments in new methodologies and new technologies. For that, there is a growing need in the development of special tools to drill safety oil wells.
One of the challenges to drill these specific oil wells is the project of a drilling fluid which must have a hydrostatic pressure to balance the low fracture gradient normally founded in these oil reservoirs. The critical points are the beginning of drilling, just below the bottom sea level, and the transition zones where pore-pressures tend to the formation fracture pressures. In both cases, there will have a narrow margin pore-fracture of working densities for drilling fluids. Bigger is the sea water depth, less is the margin of pore-fracture pressure, limiting continuous drilling operations until reach the oil zones. By consequence, normally more complex well designs are needed (many casings and many casing shoes).

The well knowledge of pore-pressures and fracture-pressures defines operations risks to avoid column prisons by differential pressure, formation damage due to drilling fluid invasions or fluid filtrates, formation fracture and bad stability of well walls.
Several models exist for prediction of fracture gradient and normally they can be used in offshore oil well design. The fracture gradient can be obtained through direct or indirect methods.
The direct methods are accomplished through measurements directly made in front of the interest intervals (well tests). For this reason, they are usually expensive because they use rig times, in spite of supplying real and trust values. Normally they serve as a data base for fracture gradient pressure calibration curves generated by indirect methods.
The indirect methods use techniques of calculations based on theoretical models or empiric models supplying an estimation of the fracture gradient along all the well depth.
To avoid all these problems during an offshore oil well drilling, all the well phases are planned, taking in account estimating values of fracturing gradient pressures for all the well depth, normally using indirect methods based on approximation of leakoff tests (direct reference tests). For this reason, the leakoff test is assumed to be the superior limit of fracture gradient pressure, permitting more economical and safer wells.

2. Proposed Model

2.1. Observation

The proposed mathematical model is an indirect method based on geometric observations obtained from geopressure logs, which gives the pressure as a function of the well depth.
The data base used for this assumption was the one found in the literature (books, articles, field data, etc).
From geopressure logs, which represent graphics of indirect methods, can be obtained proportional relationships between the pore-pressure and the fracture-pressure, both in the same Cartesian plan profile.
The proportionality relationship is obtained when it is considered the differential pressure, for the same depth, to be represented by co-linear, consecutive and adjacent segments where the endpoints of each segment are pressure gradients.

2.2. Base of the proposal

Through an applied geometric relationship in Barison (2008) based on the principle of proportional segments, the equation (1) can be obtained where F/D (fracture gradient) represents the endpoint of segment. In equation (1), F/D is the fracture gradient, S/D is the overburden gradient and P/D is the pore-pressure gradient

$$\frac{F}{D} = 0{,}38197\frac{S}{D} + 0{,}61803\frac{P}{D} \tag{1}$$

Rearranging the equation (1), the equation (2) can be obtained inserting a Calibration factor. This equation represents a linear relationship in the geopressure profile from which can be estimated a fracture gradient approximately equal to the value of leakoff test.

$$\frac{F}{D} = Calibration \cdot \left(0{,}38197\frac{S}{D} + 0{,}61803\frac{P}{D} \right) \tag{2}$$

2.3. Calibration

The calibration is obtained in two steps according to the following conditions:
1st step - This step corresponds calibration to the initial part of the well (where do not exist information) from the bottom sea level to the depth of leakoff test.
For this, it is used the equation (3) in which the value F/D is obtained from equation 1 and X one of the following parameters, depending on the sea water depth:
- Wells located in sea water depth between 300 meters and 800 meters, X =1,000.
- Wells located in sea water depth superior to 800 meters, X =10,000.

$$Calibration = \frac{Leakoff\ \ Test}{Leakoff\ \ Test - \left(\frac{Depth}{X} \cdot \frac{Leakoff\ \ Test}{\left(F/D \right)_{proportional\ \ segments}} \right)} \tag{3}$$

where:
Depth - measure since the rotary kelly bushing until the calibration depth before the leakoff test.
2nd step - Calibration from the depth of the leakoff test to the final dept of the well
For this, it is used the equation (4) in which the value F/D is obtained from proportional segments in the same depth of the test. The calibration of this step is a constant value that adjusts the F/D curve, not depending on the sea water depth.

$$Calibration = \frac{Leakoff\ \ Test}{\left(F/D \right)_{proportional\ \ segments}} \tag{4}$$

The proposed model is a geometric correlation that uses the values of pore pressure gradient and overburden gradient from another indirect model and executes the calculation on a simplified form using for this a direct adjust with values of available leakoff test of the area. This procedure is repeated for each depth throughout all the well intervals.

2.4. Application and results

Figure 1 observed in Otatti et al (2008) presents a profile of an offshore directional well 1-GM-1-USA drilled in a seawater depth of 1,493 meters with 5,200 meters final depth.
The available data for this well is three leakoff tests and two tolerances to kick tests, which are used to be compared with the literature models and the one proposed in this work.
It can be also verified that the calibration was made using only the depth of 3,100 meters and the curve generated for the proposed model was valid to the others two values of the leakoff test for the depths of 4,350 meters and 4,800 meters.
This evidenced that the values of the pressure gradients obey a proportionality correlation showing that the fracture gradient is a linear function of the proportional segments.

It is observed in Figure 1 that the curve of fracture gradient calculated throughout all the well for the methodology proposal, presented approach values with the prompt values of leakoff test that the gotten ones with the existing methodology.

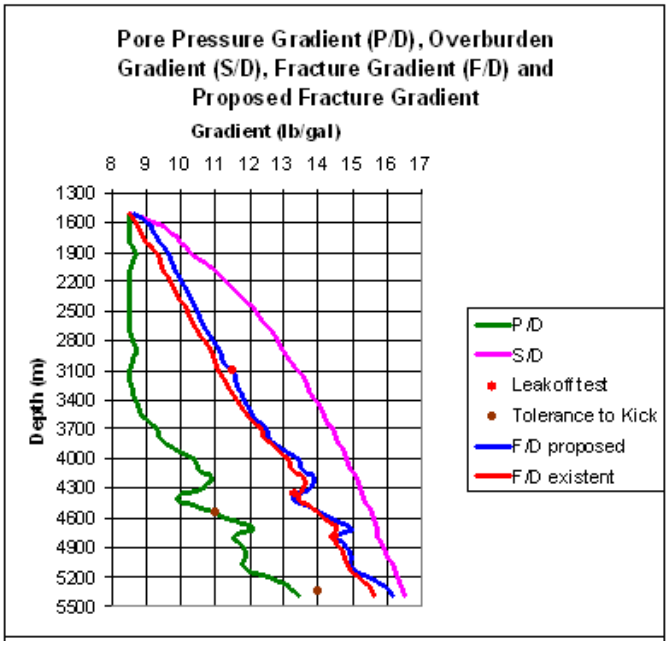


Figure 1 – Different between profiles of fracture gradient proposed and method existent observed in Otatti et al (2008).

The available data for this well is four leakoff tests shown in the Table 1, which are used to be compared with the model proposed in this work.

Table 1 - Presents the differential pressure between the values of calculated fracture gradient and the proposed model, based on leakoff tests of Figure 1.

Depth (m)	Leakoff Test (lb/gal)	Existent Model		Proposed Model	
		F/D (lb/gal)	Error %	F/D (lb/gal)	Error %
3,100	11.5	11.1	3.48	11.5	0.00
4,350	13.3	13.5	1.50	13.4	0.75
4,800	14.5	14.3	1.38	14.5	0.00

Other example from the literature of Rocha et al (2004) in wellbore stability analysis for a deepwater well drilled in Marlim Sul Oilfield, offshore Rio de Janeiro of Brazil.
The Figure 2 is shown the profile of geo-pressures of the referred well.
It is a profile of an offshore directional well drilled in a seawater depth of 1,205 meters with displacement of 3,210 meters, total vertical depth of 2,903 meters and total measured depth of 5,211meters.

It is observed in Figure 2 that the values calculated for the existing methodology for fracture gradients throughout the entire well, had been very above of the value of the leakoff test that is the referential considered superior limit.
As the sizing of the mud is based on the values of fracture gradient, it means that, problems with losses of circulation had probably occurred caused by the weight of the mud raised in great interval of the well.
With exception only of the depth of 2200 m that the value of the calculated fracture gradient almost equaled with the value of the leak off test.
It is verified that the curve of the fracture gradient obtained in Rocha et al (2004) it crossed the value of the leakoff test in the depth of 1700 m in 0,9 lb/gal and also in the depth of 2870 m where the difference crosses 1,4 lb/gal above the value of the leakoff test, meaning that that value is not compatible to establish the maximum density for the drilling fluid, because certainly it can result in loss of circulation for fracture of the formation.

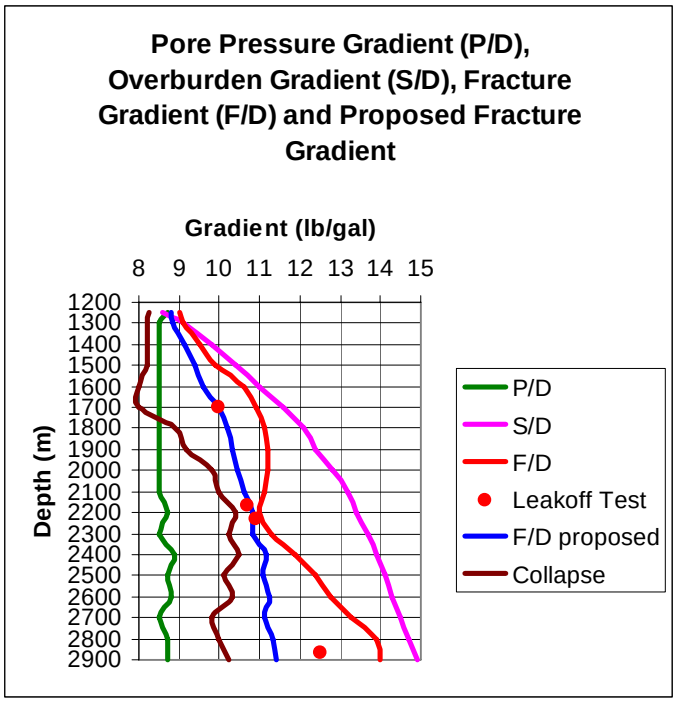


Figure 2 - Different between profiles of fracture gradient proposed and method existent observed in Rocha et al (2004).

The available data for this well is four leakoff tests shown in the Table 2, which are used to be compared with the model proposed in this work.

Table 2 - Presents the differential pressure between the values of calculated fracture gradient and the proposed model based on leakoff tests of Figure 2.

Depth (m)	Leakoff Test (lb/gal)	Existent Model		Proposed Model	
		F/D (lb/gal)	Error %	F/D (lb/gal)	Error %
1,700	10	10.9	9.00	10	0,00
2,170	10.7	11.1	3.74	10.6	0.93
2,230	10.9	11.2	2.75	10.8	0.91
2,870	12.5	13.9	11.2	11.3	9.6

2.5. Conclusions

The model proposed in this work was developed for offshore wells scenarios, since it exists, for the specific area, pore pressure gradient, overburden gradient and leakoff test. It is a geometric correlation that can be used as an alternative form substituting an indirect method to estimate the fracture gradient not depending on the Poisson Ratio of the interval perforated sediments.

The proposed model in this work showed, for the available literature data, that the fracture gradient (F/D) is a linear function of the proportional geometric segments for each depth. For this, a linear geometric relationship allows to establish an estimative value for the fracture gradient with better approach to leakoff tests than those compared with the literature models.

The simplicity of the model proposed in this work, besides can give easier and better results to estimate fracture gradient (F/D), can produce greater efficiency in the calibration with the leakoff test making possible a better exploitation of the available narrow range of pore-fracture. In consequence, a significant cost reduction can be obtained with better well designs, especially offshore ones where the costs are normally very important.

References

L. Rocha, R Andrade. H. Freire, 2004, Important aspects related to the influence of water depth on ERW. IADC/SPE 87218, 15 pages.

H. Otatti, R Lepeleire, V. Ferreira, 2008, Trabalho Final de Curso de Engenharia de Poços da Pontifícia Universidade Católica do Rio de Janeiro (PUC-Rio) On Line in: 26.08.2008. Available in: http://www.gtep.civ.puc-rio.br/petro/movie/pocos/swf/projeto.swf.

B. Barison, 2008, Geométrica: Desenho, Geometria e Arquitetura On Line in 08.08.2008. Available in: http://www.mat.uel.br/geometrica/php/pdf/dg_prop_%C3%A1urea.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimization of Steam and Solvent Injection as an Improved Oil Recovery Method for Heavy Oil Reservoirs

Edney R. V. P. Galvão,[a] Marcos A. F. Rodrigues,[a] Jennys L. M. Barillas,[a] Tarcilio V. Dutra Jr.,[a,b] Wilson da Mata[a,c]

[a]*Programa de Pós-Graduação em Ciência e Engenharia de Petróleo, Universidade Federal do Rio Grande do Norte - Campus Universitário, Natal-RN 59078-970, Brasil*
[b]*Departamento de Engenharia Química, Universidade Federal do Rio Grande do Norte - Campus Universitário, Natal-RN 59078-970, Brasil*
[c]*Departamento de Engenharia Elétrica, Universidade Federal do Rio Grande do Norte - Campus Universitário, Natal-RN 59078-970, Brasil*

Abstract

Currently a resource more and more used by the petroleum industry to increase the efficiency of steam flood mechanism is the addition of solvents. The process can be understood as a combination of a thermal method (steam injection) with a miscible method (solvent injection), promoting, thus, reduction of oil viscosity and interfacial tensions between injected fluid and oil. Solvents are hydrocarbons well known for reducing these tensions and facilitating the production of heavy oil. The use of solvent alone tends to be limited because of its high cost. When co-injected with steam, the vaporized solvent condenses in the cooler regions of the reservoir and mixes with the oil, creating a zone of low viscosity between steam and heavy oil. Mobility of the displaced fluid is then improved, resulting in an increase of oil recovery. To better understand this improved oil recovery method, a numerical study of the process was done contemplating the effects of some operational parameters (distance between wells, steam injection rate, solvent type and injected solvent volume) on cumulative oil production and oil rates. A semi synthetic model was used. Some reservoir data were obtained similar to those found in Brazilian Potiguar Basin and others ones were obtained from literature. Simulations were performed in STARS (CMG, 2007.11). It was found that injected solvent volumes increased oil recovery and oil rates. Further, the majority of the injected solvent was produced and can be recycled. High initial productions achieved by models that use solvent have normally a significant impact on the operation economics, because earlier productions suggest that fluids injection (steam and solvent) can be interrupted earlier. On environmental point of view, it would have a reduction of energy and water consumptions for steam generation, having diminished Green House Gases (GHG) emissions. Also it is important to emphasize that the high oil rates presented by these models can generate an earlier financial return, and this would be decisive for the economic viability of the project.

Keywords: solvent, steam flood, heavy oil, reservoirs modeling.

1. Introduction

Steam flood is an improved oil recovery method applied generally in viscous oil reservoirs. This method consists of injecting heat to reduce viscosity, increasing oil

mobility and facilitating its production. To increase the efficiency of this mechanism, a resource more and more used by the petroleum industry is the addition of solvents.
Diverse hybrid steam-solvent processes are being developed, mainly by Alberta Research Council (ARC), in Canada (Nasr & Ayodele, 2006). Experimental results, numerical studies and field tests suggest that great benefits can be achieved with the addition of solvents to the injected steam, as an increase of oil rates and oil-steam ratios, reduction of energy and also a reduction in water consumptions for steam generation, having diminished Green House Gases (GHG) emissions. Thus, it is indispensable to study the viability of this new technology in heavy oil fields from Brazil, with right adjusts to particularities of each reservoir.

2. Process Modeling

In this work, steam and solvent injection was analyzed through a vertical wells system. A semi synthetic model was used. Some reservoir data were obtained similar to those found in Brazilian Potiguar Basin and others ones were obtained from literature. Simulations were performed in STARS (Steam, Thermal and Advanced Reservoir Simulator) 2007.11 from CMG (Computer Modelling Group).

2.1. Reservoir Modeling

The physical model corresponds to an oil reservoir of 100 m x 100 m x 29 m, on a Cartesian coordinates system (x, y and z directions), as shown in **Fig. 1**. To reduce simulation time and considering system symmetry, the chosen flooding pattern was one quarter of an inverted five-spot, represented by an injector and a producer wells. Some of used data, like rock-fluid properties and heat loss parameters, are represented in **Table 1**. In the bottom hole direction, injector well was completed until the 8th layer, while producer well was completed until the 11th one.

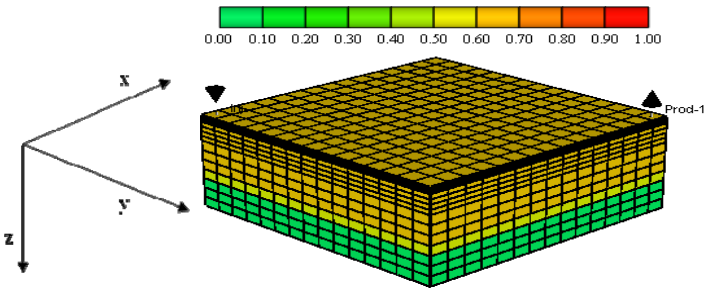


Fig. 1. Reservoir model. Initial oil saturation.

Table 1. Reservoir simulation input parameters.

Item	Value
Reservoir thickness (m)	29
Oil zone thickness (m)	20
Water zone thickness (m)	9
Initial reservoir temperature (ºC)	37.8
Reservoir depth (m)	200
Porosity (%)	30
Vertical/horizontal permeabilities ratio (%)	10

Table 1. Reservoir simulation input parameters *(continuation).*

Item	Value
Oil viscosity*	1000 cP@37.8°C
Initial oil saturation (%)	70
Irreducible water saturation (%)	30
Steam temperature (ºC)	288
Steam quality (%)	55
Maximum pressure in the injector well (kPa)	7,200
Minimum pressure in the producer well (kPa)	197
Thermal conductivity of rock and surrounding formation overburden and underburden (J/m-s-K)	1.7
Thermal conductivity of water (J/m-s-K)	0.6
Thermal conductivity of oil (J/m-s-K)	0.13
Thermal conductivity of gas (J/m-s-K)	0.04
Volumetric heat capacity of surrounding formation overburden and underburden (J/m^3-K)	66,465
Production time (yr)	16

*Viscosity *versus* temperature curve was obtained from *Barillas, 2005*.

Equations from mathematical modeling are embedded in STARS, which is a commercial software. Thus, it is not necessary to insert them in the input data file.
Description of the process is done using specific keys from STARS library. Input parameters depend on available rock-fluid properties, as ones showed in **Table 1**. Because ideal models were used, they were not compared with real models.

2.2. Analyzed Operational Parameters

The minima (-1), intermediates (0) and maxima (+1) levels as well as nomenclature of analyzed operational parameters are shown in **Table 2**. A factorial planning of three levels was realized to study the effects of parameters interactions, resulting in 81 simulations.

Table 2. Range of analyzed operational parameters.

Parameter	Minimum (-1)	Intermediate (0)	Maximum (+1)
Distance between injector/producer wells (m)	65	100	135
Steam injection rate (m^3/day)	20	35	50
Solvent type	Pentane	Hexane	Heptane
Injected solvent volume/Injected steam volume (%)	5	10	15

Used maximum production rates correspond to the double of respective values of fluids (steam + solvent) injection rates. Oil recoveries (OR) were calculated dividing cumulative oil (CO) by original oil *in place* (OOIP) associated to each grid configuration (OR = CO/OOIP). Being a function of rock total volume, porous volume also varied, while all the others reservoir properties remained constant.

3. Results

Through the analysis of parameters interactions and taking oil recovery as objective function, an optimized model was achieved. **Table 3** shows the parameters from this model and others models used to emphasize the effect from injected solvent volume and to make a comparison with the case in which steam is injected without solvent.

Table 3. Operational parameters values for models 0%, 5%, 10% and 15% of injected solvent.

Parameter	0% Solvent	5% Solvent	10% Solvent	15% Solvent (Optimized Model)
Distance between injector/producer wells (m)	100	100	100	100
Steam injection rate (m^3/day)	20	20	20	20
Solvent type	_	Heptane	Heptane	Heptane
Injected solvent volume/Injected steam volume (%)	0	5	10	15

3.1. Comparison among Primary Recovery and 0%, 5%, 10% and 15% Solvent Models

In **Fig. 2**, cumulative oil (without solvent) *versus* time are showed for primary recovery and models listed in **Table 3**. From second year on, curves with solvent models registered cumulative oil (without solvent) higher than the ones obtained by model 0% Solvent. This suggests that presence of solvent, for the analyzed amounts, accelerated the arrival of warm oil as far as producer well, promoting, thus, an earlier production with regard to model 0% Solvent. Moreover, it can be evidenced that an increase of injected solvent volume improves oil recovery (70,3% for model 5% Solvent, 71,1% for model 10% Solvent and 72,4% for model 15% Solvent).

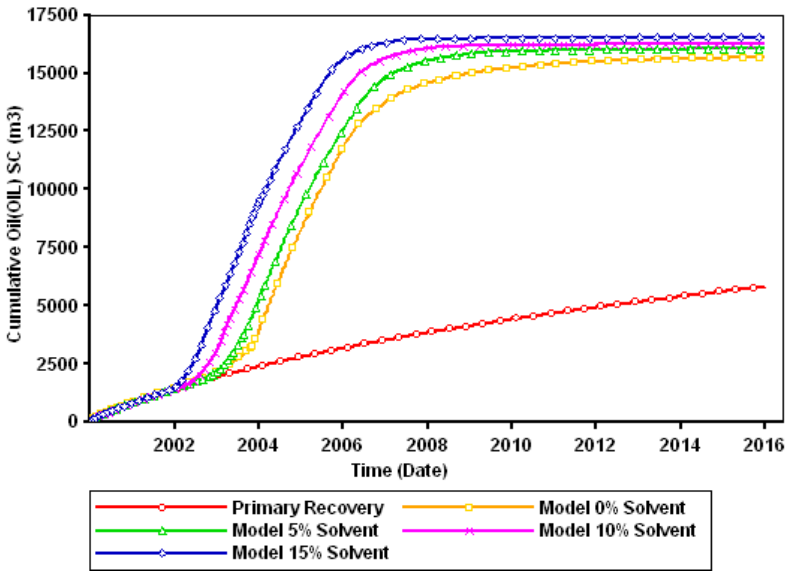


Fig. 2. Cumulative oil (without solvent) *versus* time. Primary recovery and models 0%, 5%, 10% and 15% Solvent.

Fig. 3 shows oil component (without solvent) rates *versus* time for these models. It can be observed that the higher injected solvent volume, the earlier production rates. This effect occurred mainly between second and fourth year, in which Optimized Model (15% Solvent) achieved 15.0 m^3/day, while model 0% Solvent achieved only 5.0 m^3/day at the same time. High rates obtained initially by Optimized Model contributed for a faster reservoir depletion, justifying the lower rates achieved by one in the following period.

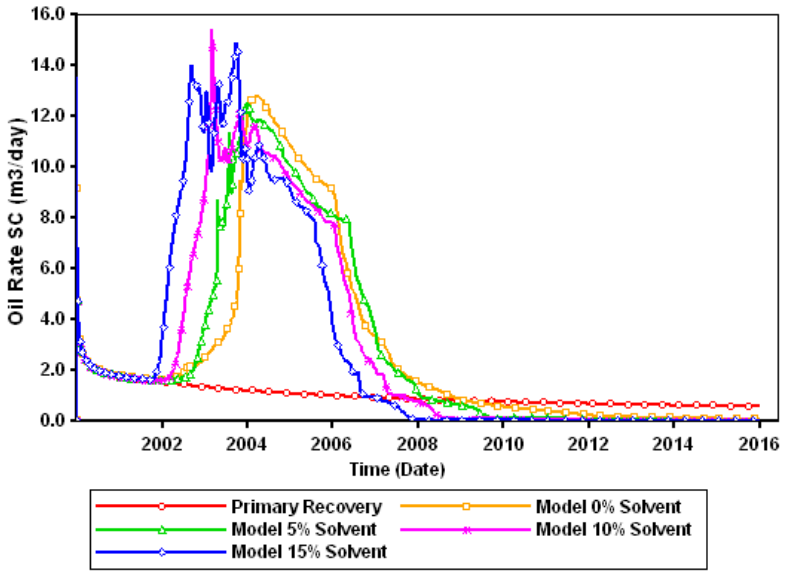


Fig. 3. Oil component (without solvent) rates *versus* time. Primary recovery and models 0%, 5%, 10% and 15% Solvent.

High initial productions achieved by models that use solvent have normally a significant impact on the operation economics, because early production suggests that fluids injection (steam and solvent) can be interrupted earlier. On environmental point of view, solvent injection can provide a reduction of energy and also a reduction in water consumptions for steam generation, having diminished Green House Gases (GHG) emissions. Also it is important to emphasize that the higher oil rates presented by these models can generate an earlier financial return and, by consequence, a project with a good economical viability (Galvão, 2008).

Injected and produced volumes of solvent *versus* time for models are registered in **Fig. 4**. Practically all injected solvent was produced together with oil from reservoir. To quantify the reuse of produced solvent is not in the scope of this work, however it is expected that produced fluid is lighter (high ºAPI) and has higher economic value (Galvão, 2008). Also it can be evidenced that the higher injected solvent volume, the earlier its recovery began.

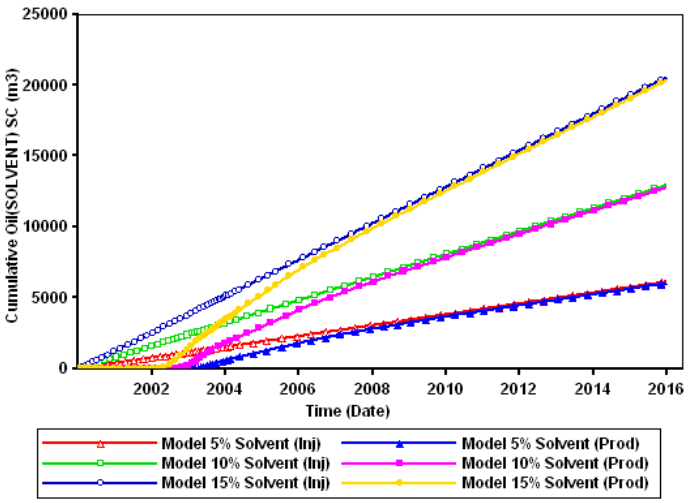


Fig. 4. Injected and produced volumes of solvent *versus* time. Models 5%, 10% and 15% Solvent.

4. Conclusions

• An increase of injected solvent volume improved the cumulative oil production with regard to model without solvent;

• Presence of solvent, for the analyzed amounts, accelerated the arrival of warm oil as far as producer well, promoting, thus, an earlier production with regard to model without solvent. This suggests that fluids injection (steam and solvent) can be interrupted earlier, reducing energy and water consumptions for steam generation, Green House Gases (GHG) emissions and generating an earlier financial return;

• Practically all injected solvent was produced together with oil from reservoir. It is expected that produced fluid is lighter (high ºAPI) and has higher economic value. Also it could be evidenced that the higher injected solvent volume, the earlier its recovery began.

Acknowledgements

The authors want to thank CAPES and Laboratório de Estudos Avançados em Petróleo (LEAP-UFRN) for the support received in the execution of this work.

References

NASR, T. N.; AYODELE, O. R. New hybrid steam-solvent processes for the recovery of heavy oil and bitumen. SPE 101717. Nov, 2006.

GALVÃO, E. R. V. P. Injeção de vapor e solvente como um método de recuperação avançada em reservatórios de óleo pesado. 2008. 106p. Dissertation (Master Degree in Petroleum Engineering) – Centro de Tecnologia, Programa de Pós-Graduação em Ciência e Engenharia de Petróleo, Universidade Federal do Rio Grande do Norte, Natal.

CMG, Computer Modelling Group Ltda. Guía para el usuario. Steam, Thermal and Advanced Reservoir Simulator - STARS. Versão 2007.11, Calgary-Alberta-Canada.

BARILLAS, J. L. M. Estudo do processo de drenagem gravitacional de óleo com injeção contínua de vapor em poços horizontais. 2005. 163p. Dissertation (Master Degree in Chemistry Engineering) – Centro de Tecnologia, Departamento de Engenharia Química, Programa de Pós-Graduação em Engenharia Química, Universidade Federal do Rio Grande do Norte, Natal.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Multi objective optimization using Life Cycle Environmental Impact and Cost in the Operation of Utility plants

Pablo Martínez, Ana María Eliceche

Departamento de Ingeniería Química,Universidad Nacional del Sur, PLAPIQUI-CONICET, Camino La Carrindanga km 7, 8000 Bahía Blanca, Argentina.

Abstract

Environmental and economic objective functions are used simultaneously to select the operating conditions of a steam and power plant. Different methodologies to solve multi objective optimization problems were implemented successfully. The life cycle potential environmental impact and the operating cost of the power plant are minimized simultaneously. A methodology is presented to estimate the potential environmental impacts during the most important life cycle stages associated with imported fuel and electricity in the utility plant. Mixed Integer Non Linear multi objectives problems are formulated and different strategies are implemented and successfully solved in GAMS.

Keywords: multi objective optimization, life cycle environmental impact, utility plant.

1. Introduction

The reduction of environmental impact of process operations is one of the imperative challenges to achieve sustainable development in process industry for the current century; also economical and social objectives are imperative issues for sustainability. Economic objectives have been used in process system engineering. However, environmental and economical objectives do not follow comparable behaviours. While the economic implication of a process is minimized, the associated environmental impact rises and an environmental friendly process is often cost intensive. This fact evidences the need of solving problems with more than one objective, that is, to find solutions that satisfy environmental and economic objectives at the same time. This solutions could be found formulating Multi objective Optimization Problems in the decision making process. The multi objective optimization problem could, in theory, be solved using similar methods as those employed in single objective optimization problems, converting the multiple objectives into a single objective. Multi-objective optimization applied to environmental and economic objectives has been treated by authors like Ciric and Huchette (1993) minimizing the amount of waste and the net profit of an ethylene glycol production plant. Dantus and High (1999) proposed a method to convert a bi objective optimization problem into a single objective optimization problem; the method proposed is a variation of the utopia point distance minimization, including discrete variables to select the type of reactor to be used in the methyl chloride superstructure plant design. Pistikopoulos and Hugo (2005) have treated multi objective optimization with environmental and economic objectives applied to supply chain network design and planning using the e-constraint method.

In the present work the operating conditions of a steam and power plant are selected to minimize life cycle environmental impact and operating cost simultaneously solving a multi objective optimization problem. The environmental objective is the life cycle environmental impact associated with solid wastes, gaseous and liquid emissions of a steam and power plant. In the life cycle context, the battery limits of the steam and power plant need to be extended in order to include emissions of imported natural gas and electricity generated by nuclear, hydroelectric and thermoelectric plants. The operating cost includes costs of imported electricity, natural gas feed, makeup water and water treatment. A Mixed integer non linear multi objective optimization problem is formulated and solved in GAMS (Brooke et al, 2003).

2. Environmental and Economic Objective Functions

2.1. Potential Environmental Impact Evaluation

The Potential Environmental Impact (PEI) function considered is a multi objective function itself, since nine environmental impact categories are considered: global warming, acidification, ozone depletion, photo oxidant formation, eutrophication, fresh water ecotoxicity, human toxicity, source depletion and the impact due to ionizing radiation. The Potential Environmental Impact is calculated using the Guinée et al. (2002) methodology. The contribution of the emission of a pollutant **k** to a given environmental impact category **j** is evaluated multiplying the pollutant **k** flow rate $\mathbf{F_k}$ emitted into the environment by a characterization factor $\boldsymbol{\gamma_{kj}}$ published by Guinée et al. (2002). This characterization factor represents the effect that chemical **k** has on the environmental impact category **j**. Hence, the Potential Environmental Impact, PEI, is calculated as follow:

$$\mathbf{PEI} = \sum_{\mathbf{j}} \sum_{\mathbf{k}} \boldsymbol{\alpha}_{\mathbf{j}} \times \mathbf{F}_{\mathbf{k}} \times \boldsymbol{\gamma}_{\mathbf{k,j}} \qquad (1)$$

Where $\boldsymbol{\alpha_j}$ represent the weighting factors for each environmental impact category **j**. More information can be found in Eliceche et al. (2007). Eq.1 transforms the pollutants emissions flow rates into potential environmental impacts.

2.1.1. Utility Plant Environmental Impact

The emissions of the steam and power plant are evaluated from the modelling of the main processes formulated in GAMS. The emissions come mainly from the combustion in the boilers of a mixture of natural gas, $\mathbf{F_{ng}}$ and residual gas, $\mathbf{F_{rg}}$. Liquid emissions of purge streams, $\mathbf{F_p}$, in the boilers and cooling system are also considered. The pollutants emissions from the utility plant (UP) are calculated as follow:

$$\mathbf{F}_{\mathbf{k}}^{\mathbf{UP}} = \mathbf{F}_{\mathbf{gn}} \times \mathbf{e}_{\mathbf{k,ng}} + \mathbf{F}_{\mathbf{rg}} \times \mathbf{e}_{\mathbf{k,rg}} + \mathbf{F}_{\mathbf{p}} \times \mathbf{e}_{\mathbf{k,p}} \qquad (2)$$

Where $\mathbf{e_{k,ng}}$ is the emission factor for pollutant **k** due to the combustion of natural gas, $\mathbf{e_{k,rg}}$ is the corresponding emission factor for residual gas combustion and $\mathbf{e_{k,p}}$ is the pollutant emission factor for liquid emissions. The emissions factors express the amount of pollutant **k** emitted by unit mass of natural gas, residual gas and liquid stream, respectively. The CO_2 emission due to natural gas and residual gas combustion are estimated stoichiometrically with the gas composition following the IPCC 2001

recommendations. Nearly 100 gaseous pollutants emissions are estimated from AP-42 report (EPA, 1998). Emissions from liquid discharges were estimated from the Electrical Power Research Institute report (2000).

2.1.2. Life Cycle Environmental Impact

Life cycle approach considers emissions during the entire life cycle of a product or service accounting by emissions from raw material extraction to waste disposal. In the case study presented in this work, the life cycle emissions are considered for the natural gas feedstock and the imported electricity needed to move some electrical motors in the superstructure of the steam and power plant. Pollutant flow rate for natural gas (NG) life cycle $\mathbf{F_k^{NG}}$ is calculated in the following equation:

$$\mathbf{F_k^{NG}} = \mathbf{F_{ng}} \times \sum_{\mathbf{l}} \mathbf{e_k^l} \qquad \mathbf{l} = 1,...,\mathbf{l_{ng}} \tag{3}$$

Where $\mathbf{e_k^l}$ is the emission factor for pollutant k in the life cycle stage **l**, $\mathbf{l_{ng}}$ is the total number of life cycle stages considered for the natural gas fuel cycle: exploration, extraction and transportation stages. As the residual gas is produced in the ethylene plant, no life cycle stage has been considered for it.
For imported electricity (IE) life cycle, emissions have been assessed through the life cycle of different electricity generation plants. The electricity generation sector in Argentina has contributions from thermoelectric, hydroelectric and nuclear plants. Thermoelectric power generation consumes coal, oil and natural gas as fuels; nuclear power generation consumes natural uranium fuel. The estimation of pollutant emissions in the electric power generation includes the following life cycle stages: extraction and processing of raw materials, transport, refining (where it is applicable) and electricity generation itself:

$$\mathbf{F_k^{IE}} = \sum_{\mathbf{q}} \sum_{\mathbf{l_q}} \mathbf{W_q} \times \mathbf{e_{k,q}^{l_q}} \qquad \mathbf{l} = 1,...,\mathbf{l_{ie}} \tag{4}$$

Where $\mathbf{W_q}$ is the electricity imported and generated with technology **q**, $\mathbf{l_q}$ superscript accounts life cycle stage **l** in electricity generated by option **q**, finally $\mathbf{e_{k,q}^{l_q}}$ is the corresponding emission factor of pollutant **k** in electricity generated with option **q**, for the life cycle stage $\mathbf{l_q}$. The life cycle stages considered are: (i) exploration, extraction, refining and transport of natural gas, oil, coal and uranium consumed in thermoelectric and nuclear plants; (ii) submerged biomass decay in hydroelectric plants (iii) waste treatment and disposal for nuclear plants and (iv) transport in the construction stage of hydroelectric and nuclear plants.
The utility plant potential environmental impact, PEI^{UP} is calculated as follows:

$$\mathbf{PEI^{UP}} = \sum_{\mathbf{j}} \sum_{\mathbf{k}} \boldsymbol{\alpha}_{\mathbf{j}} \times \mathbf{F_k^{UP}} \times \boldsymbol{\gamma}_{\mathbf{k,j}} \tag{5}$$

The component **k** life cycle emissions $\mathbf{F_k^{LC}}$ are estimated adding the component **k** emissions in the utility plant, life cycle of imported natural gas and electricity:

$$\mathbf{F_k^{LC} = F_k^{UP} + F_k^{NG} + F_k^{IE}} \qquad (6)$$

The life cycle potential environmental impact is evaluated as follows:

$$\mathbf{PEI^{LC} = \sum_j \sum_k \alpha_j \times F_k^{LC} \times \gamma_{k,j}} \qquad (7)$$

Global warming due to combustion emissions is the most relevant environmental category for steam and power plants and for fossil fuels electricity generation.

2.1.3. Economical Objective Function

The operating cost of the utility plant includes costs of imported electricity (IW), natural gas feed (NG), makeup water (MW) and water treatment (WT); where $\mathbf{c_{ng}}$, $\mathbf{c_q}$, $\mathbf{c_{MW}}$ and $\mathbf{c_{WT}}$ are the cost coefficients:

$$\mathbf{C = F_{ng} \times c_{ng} + \left(\sum_q W_q \right) \times c_W + F_{MW} \times c_{MW} + F_{WT} \times c_{WT}} \qquad (8)$$

A detailed mathematical model of the utility plant operation is presented in Eliceche et al (2007).

3. Multi objective Optimization with Environmental and Economic Objectives

The multi objective (MO) optimization is a system analysis approach to problems with conflictive objectives. A key factor of MO optimization is that rarely exist a single solution that simultaneously optimizes all the objectives. In its place, there is a set of solutions where one objective cannot be improved except at expense of another objective. This set of compromise solutions are generally referred as non-inferior or Pareto optimal solutions. A variety of strategies to solve multi objective optimization problems exist, that can be found in Alves et al (2007). The general approach consists in converting the multiple objectives into a single objective. Some of these methods are: weighted sum, utopia point distance minimization, e-constraint method and global criteria method. The general formulation of a bi objective optimization problem considering continuous and discrete variables follows:

$$\begin{aligned} &\mathbf{\min_{x,y} Z = Z[PEI(x,y), C(x,y)]} \\ &\mathbf{s.t:} \quad \mathbf{h(x)} = 0 \\ &\qquad \mathbf{g(x) + A(y)} \le 0 \\ &\qquad \mathbf{x^{LB} \le x \le x^{UB}} \\ &\qquad \mathbf{x} \in \mathbf{R^n} \\ &\qquad \mathbf{y} \in \{0,1\}^{\mathbf{m}} \end{aligned} \qquad P1$$

Where **x** and **y** are the continuous and binary optimization variables, respectively. Superscripts **U** and **L**, indicates upper and lower bounds on vector **x**, respectively. The equality constraints **h(x) = 0** are the system of non-linear algebraic equations that represent the steady state modelling of the process plant, including mass and energy balances; enthalpy and entropy prediction. The inequality constraints **g(x) + A(y) ≤ 0**

represent logical constraints, minimum and maximum equipment capacities, operating and design constraints, etc. The A matrix includes linear relations between binary variables such as logical constraints.
Different strategies to solve multi objective optimization problems have been implemented successfully. The multi objective function Z in problem *P1* for the global method presented by Dantus and High (1999) follows, with the nomenclature presented in section 2:

$$\mathbf{Z} = \boldsymbol{\omega}_1 \times \left[\frac{\mathbf{PEI} - \mathbf{PEI}^{*}}{\mathbf{PEI}^{**} - \mathbf{PEI}^{*}} \right]^{\mathbf{p}} + \boldsymbol{\omega}_2 \times \left[\frac{\mathbf{C} - \mathbf{C}^{*}}{\mathbf{C}^{**} - \mathbf{C}^{*}} \right]^{\mathbf{p}} \qquad (9)$$

Where $\boldsymbol{\omega}_1$ and $\boldsymbol{\omega}_2$ are weighting factors, these preference weights $\boldsymbol{\omega}_i$ are used to represent the relative importance of each objective. The decision-maker's preferences are also expressed in the compromise index $\mathbf{p}$ ($1 \leq \mathbf{p} \leq \infty$), which represents the decision-maker's concern with respect to the maximal deviation from the utopia point. As a result, the non-inferior solutions defined within the range $1 \leq \mathbf{p} \leq \infty$ correspond to the compromise set from which the decision maker still has to make the final choice to identify the best compromise solution (Dantus and High, 1999). The single asterisk indicates the minimum values of a given objective function solving a single objective optimization problem, while double asterisk indicates the alternative objective function value obtained. The objective functions used are life cycle potential environmental impact (PEI) given in Equation 6 and operating cost (C) given in Equation 8.

4. Numerical Results

A rigorous modelling of the utility plant is formulated in GAMS, including the power and steam demands of the ethylene plant (Eliceche et al, 2007). The continuous operating variables selected are temperature and pressure of the high, medium and low pressure steam headers and the deareator pressure. Binary operating variables are introduced to represent discrete decisions such as the selection of: (i) alternative pump drivers such as electrical motors and steam turbines and (ii) boilers which are on or off, and their auxiliary equipment such as feed pumps and air fans. Thus a multi objective Mixed Integer Nonlinear Programming problem is formulated and solved in GAMS.
Different strategies were implemented to solve the multi objective problem The solution point reported in Table 1 was obtained with the Dantus and High (1999) method and the following parameters for equation 9: $\boldsymbol{\omega}_1$=0.1, $\boldsymbol{\omega}_2$=0.9, $\mathbf{p}$=1. The following GAMS options were used: DICOPT as the outer approximation algorithm; CONOPT3 to solve the Non Linear Programming sub problem and CPLEX to solve the Mixed Integer Linear Programming sub problem.

Table 1. Multiobjective problem solution.

Objective Functions		Initial Point	Solution point	Reductions
PEI^{LC}	(PEI / h)	33627.33	29581.88	**12 %**
Cost	(U$A / h)	561.84	470.97	**16 %**

Significant reductions in the order of 12 % in the life cycle environmental impact and 16 % in the operating cost can be achieved selecting the operating conditions with the methodology proposed. Regarding the selection of pump's drivers, steam turbines are chosen rather than electrical motors, due to the fact that the environmental impact to

power generated ratio is smaller in the steam and power plant than in the generation of the imported electricity. This is due to the fact that natural gas is burned with residual gas from the demethanizer column. The residual gas is a Hydrogen rich stream, having higher combustion heat and lower combustion emissions than natural gas or any other fossil fuel. The operating cost is also cheaper with steam turbines than with electrical motors. The number of the boilers in operation is reduced from four to three, due to a proper selection of temperature and pressure of steam headers, mainly the high pressure steam header.

This is a process where improving process efficiency, environmental impact and cost are reduced simultaneously. They are not conflictive objectives. This is not the case if the environmental impact evaluation is reduced to the battery limits, where minimizing environmental impact and operating cost leads to different solutions mainly regarding the selection of alternative drivers.

5. Conclusions

A methodology has been presented to select the operating conditions minimizing simultaneously life cycle environmental impact and operating cost, solving a mixed integer nonlinear multi objective optimization problem. Imported natural gas and electricity life cycle environmental impacts have been estimated. Different strategies to solve multi objective optimization problems were implemented successfully. The ethylene steam and power plant analyzed, has a relevant contribution to combustion emissions, global warming, consumption of non renewable fossil fuels and operating cost. Thus, significant improvements in the plant operation can be achieved as shown in Table 1, with the strategy presented. This is a plant where improving process efficiency, environmental impact and cost are reduced simultaneously. It is also very important to extend the battery limits to include life cycle analysis, when environmental objectives are used to support a decision making process.

References

A. Brooke, D. Kendrick, A. Meeraus and R. Raman, 2003, GAMS development Corporation. A user guide.

A. Ciric and S. Huchette, 1993, Multiobjective Optimization Approach to Sensitivity Analysis: Waste Treatment Costs in Discrete process Synthesis and Optimization problems, Ind. Eng. Chem. Res., 32, 2636-2646.

A. Eliceche, S. Corvalan, P. Martinez, 2007, Environmental life cycle impact as a tool for process optimization of a utility plant, Comp.& Chem. Eng., 31, 648-656.

Electrical Power Research Institute, 2000, Aqueous discharges from steam electric power plants, data evaluation, Updated Report. http://www.epri.com

E. Pistikopoulos and A. Hugo, 2005, Environmentally conscious long-range planning and design of supply chain networks, Journal of Cleaner Production, 13, 1471-1491.

J. Guinée, R. Heijungs, G. Huppes , R. Kleijn, A. Koning, L. van Oers, A. Sleeswijk, S. Suh, H. Udo de Haes (eds.), 2002, Handbook on Life Cycle Assessment. Operational Guide to the ISO Standards. Kluwer Academic Publishers, Dordrecht.

IPCC. Climate Change 2001: the scientific basis. Intergovernmental Panel on Climate Change. Cambridge University.

M. Alves and J. Climaco, 2007, A review of interactive methods for multiobjective integer and mixed-integer programming, European J. of Operatinal Research, 180, 99-115.

M. Dantus, K. High, 1999, Evaluation of waste minimization alternatives under uncertainty: a multi objective optimisation approach, Comp. Chem. Eng. 23, 1493-1508.

U.S. Environmental Protection Agency,1998, AP-42. Compilation of air pollutant emission factors.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling of Flowcharts of Permeation through Membranes for Removal of CO2 of Natural Gas

Andressa Nakao[a], Ana P. F. Macedo[a], Betina M. Versiani[a], Ofélia de Q. F. Araújo[a], José L. de Medeiros[a]

[a] *Departament of Chemical Engineering, Escola de Química da Universidade Federal do Rio de Janeiro. Sala E-201, Bloco E, Centro de Tecnologia, Ilha do Fundão, CEP: 21949-900 Rio de Janeiro, Brazil.*

Abstract

In this present work, a simulator of permeation membranes is presented, the Permeation Membrane Simulator (PMS), which is based on rigorous engineering principles including sound thermodynamics, fluid mechanics, heat transfer and mass transport. The use of hollow-fiber membrane units is studied in the context of CO_2 removal from natural gas, which seems to be the preferred configuration in offshore plants. The feed stream is considered multi-component, with real chemical species, enabling reliable performance evaluations at different scenarios of pressure, temperature and composition of the gas to be treated by permeation. The resulting differential equations are solved along the characteristic contactor size: (i) steady state mass balances for the species present in each phase (L and V); (ii) momentum balances in each stationary phase; (iii) force balances in each phase. Major simplifications arise from adopting steady state and one-dimensional space over the membrane contactor.
The process arrangements are based on recent studies concerning subsalt oil finds in the offshore Santos Basin, Brazil. A process flowsheet is simulated, comprising compressors and membrane modules, corresponding to a scenario of 4MMm³/d and 12% mol CO_2. The process is composed of three permeation and compression stages, producing one stream of commercial grade natural gas and another stream rich in CO_2 for Carbon Capture & Geological Storage – CCGS - purposes. Considering the simulation results of the scenarios presented, the application of membrane technology in offshore facilities is proven technically feasible.

Keywords: Simulation, Natural Gas, CO_2, Hollow-Fiber.

1. Introduction

The recent intensification of Global Climate Change indicators highlights the need of reducing Greenhouse Gas (GHG) Emission. In this scenario, the main strategy considered by world leaders is synthesized by CCGS - Carbon Capture & Geological Storage; i.e. the capture of industrial emissions of CO_2 followed by underground sequestration into geological formations such as deep saline aquifers. The CCGS is presented as a Bridge-Technology in the sense that it is a temporary solution that should include CO_2 emissions giving subside to the World Economy of Energy, currently based on carbon fossil burning, while new sustainable energy sources become mature for widespread using. Brazil recent finds of natural gas (NG) in the giant subsalt reserves of the Santos Basin contains an expressive quantity of CO_2 (Formilgi, 2008), requiring

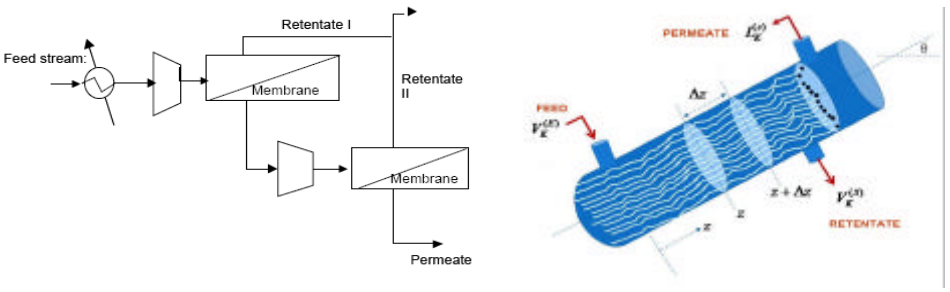
offshore treatment for CO_2 removal to: (i) minimize GHG emissions; (ii) increase the heating value of NG for use as fuel; (iii) reduce the potential for corrosion in metal pipelines to transport gas (Engelien, 2004); and (iv) to meet regulation standards for commercial NG (ANP, 2008).

Hence, NG purification is a contribution to CCGS strategies. Currently, .the main technologies used in industry are membrane modules and absorption in aqueous solutions of ethanolamines. For offshore platforms, membranes offer comparative advantages over other separation technologies – e.g. absorption and adsorption - such as: (i) simple operation requiring no heat exchange in either phase, (ii) free of chemical additions such as corrosive, toxic or flammable compounds, (iii) small foot-print due to compact, low weight units; and (iv) modular arrangements, which turns the process a flexible technology (Baker et al, 2008). These characteristics make the technology in permeation membranes more appropriate to the capture of CO_2 in offshore presalt conditions. However, it should be noticed that the main obstacle to the use of membranes is the immense process scale required in presalt NG purification. It is apparent that membrane systems lose selectivity when operating under high flow rates.

The selection of membrane configuration depends in the specifications set to the separation process. This study presents a process simulator developed as aid for performance evaluation of hollow-fiber membrane modules, the most used configuration for separation with gas permeation in membranes in offshore plants.

2. Model Formulation

The formulation of mathematical models capable of describing membrane separation processes is not trivial. The challenge is to develop models that contain the fundamental equations that describe the pertinent phenomena of this type of separation, with simple specifications, targeting potential industrial applications. The model premises are: (i) steady state; (ii) one independent variable – axial position z; (iii) rigorous 1-D compressible flow; (iv) heat transfer; (v) and permeation fluxes described via linear driving forces in terms of fugacities.
In this work, the retentate is represented by V and the permeate by L. Figure 1(a) shows the process configuration used and Figure 1(b) presents the membrane module used in the simulator, with variable labels employed in the model equations presented.



(a) **(b)**

Figure 1: Simulated Process. (a) membranes arrangement; (b) membrane module used in the simulator

The basic features of this model are: (i) independent variable: z; (ii) four dependent variables: pressure (P), temperature (θ), fluid inventory (I); and flow rates of species k (V_k, L_k); (iii) heat transfer taken into account from the distribution of external temperature, $?_E(x)$, and the distribution of heat transfer coefficient O(x); (iv) friction term via Churchill Equation for universal Darcy friction factor (*f*); (v) ideal gas properties according to Poling, Prausnitz and O'Connel (2001); (vi) viscosity of dense compressible fluid model proposed by Chung *et al.* (1988); (vii) rigorous thermodynamics via Peng-Robinson equation of state. Mass, momentum and energy balances are presented in Equations 1 through 7, derived in the hull's axial direction.

Mass Balance - Phase V:

$$\frac{\partial V_K}{\partial z} = -S\ a\ N_K \tag{1}$$

Mass Balance - Phase L:

$$\frac{\partial L_K}{\partial z} = S\ a\ N_K \tag{2}$$

Momentum Balance - Phase V:

$$\left\{1 - G_{P_V}\left(\frac{q_V}{r_V S_V}\right)^2\right\}\frac{\partial P_V}{\partial z} + \left\{-G_{T_V}\left(\frac{q_V}{r_V S_V}\right)^2\right\}\frac{\partial T_V}{\partial z} = \frac{S\ a\ q}{S_V}\left(\frac{q_V}{r_V S_V}\right) - r_V\, g\ sen(q) - \frac{Y_V(pD + S\ a)}{S_V} \tag{3}$$

Momentum Balance - Phase L:

$$\left\{1 - G_{P_L}\left(\frac{q_L}{r_L S_L}\right)^2\right\}\frac{\partial P_L}{\partial z} + \left\{-G_{T_L}\left(\frac{q_L}{r_L S_L}\right)^2\right\}\frac{\partial T_L}{\partial z} = \frac{S\ a\ q}{S_L}\left(\frac{q_L}{r_L S_L}\right) - r_L g\ sen(q) - \frac{Y_L S\ a}{S_L} \tag{4}$$

where

$$\Gamma_{P_V} = \left(\frac{\partial r_V}{\partial P}\right)_{T,\underline{N}}, \Gamma_{P_L} = \left(\frac{\partial r_L}{\partial P}\right)_{T,\underline{N}}, \Gamma_{T_V} = \left(\frac{\partial r_V}{\partial T}\right)_{P,\underline{N}}, \Gamma_{T_L} = \left(\frac{\partial r_L}{\partial T}\right)_{P,\underline{N}} \tag{5}$$

Energy Balance - Phase V:

$$\left\{1 - G_{P_V}\left(\frac{q_V}{r_V S_V}\right)^2 + \frac{T_V G_{T_V}}{r_V}\right\}\frac{\partial P_V}{\partial z} + \left\{\frac{r\overline{C}_P^V}{M_V} - G_{T_V}\left(\frac{q_V}{r_V S_V}\right)^2\right\}\frac{\partial T_V}{\partial z} = -r_V g\ sen(q) + \frac{S\ a\ q}{S_V}\left(\frac{q_V}{r_V S_V}\right) + \frac{W_E pD r_V}{q_V}(T_E - T_V) + \frac{W\ S\ a\ r_V}{q_V}(T_V - T_L) \tag{6}$$

Energy Balance - Phase L:

$$\left\{1 - G_{P_L}\left(\frac{q_L}{r_L S_L}\right)^2 + \frac{T_L G_{T_L}}{r_L}\right\}\frac{\partial P_L}{\partial z} + \left\{\frac{r\overline{C}_P^L}{M_L} - G_{T_L}\left(\frac{q_L}{r_L S_L}\right)^2\right\}\frac{\partial T_L}{\partial z} = -r_L g\ sen(q) + \frac{S\ a\ q}{S_L}\left(\frac{q_L}{r_L S_L}\right) + \frac{W\ S\ a\ r_L}{q_L}(T_V - T_L) + \frac{S\ a\ r_L}{q_L}\sum_K N_K\left(E_K^V - E_K^L\right) \tag{7}$$

3. Case Study

The flowsheet GNBRPS presented in Figure 3 is based on studies from Santo Basin, Brazil (Formigli, 2008), and defines a process fed with current GNBRPS with average content of CO_2 (12% mol), pressure of 60 bar, temperature 25°C and flow rate of 4 MMm^3/d, comprised of 3 permeation modules with intermediary compression operations. The compression operates in 3 stages with efficiency of 75% and intercoolers to 35 ºC. Process specification is to perform the permeation at low pressure in module 1 (P = 1bar) before feeding the module 2.

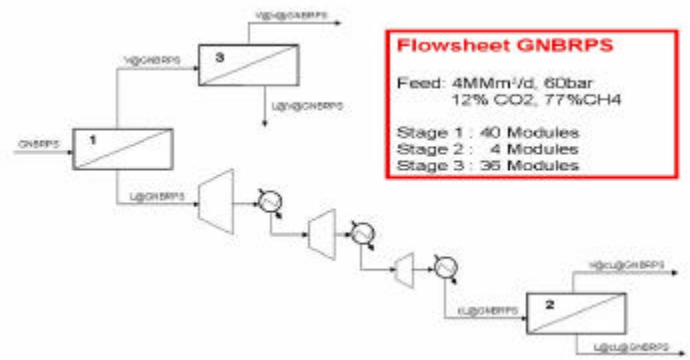
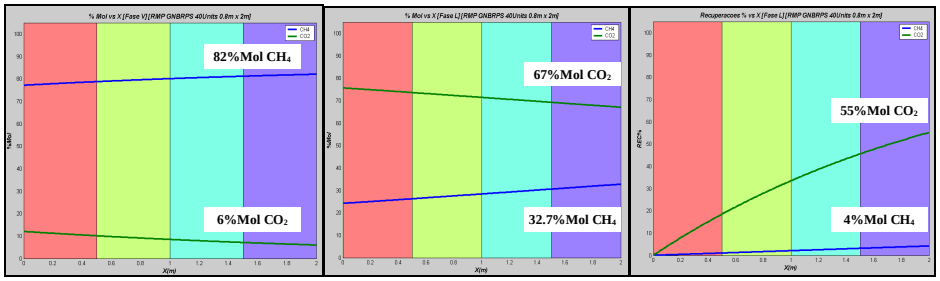


Figure 3: Flowsheet GNBRPS

Simulation Results for First Stage: Figures 4(a) and 4(b), respectively, show composition profiles (mol %) of CO_2 and CH_4 in both phases V and L of the process. Figure 4(c) illustrates the profile (%) of CO_2 recovery in phase L (permeate 1). In phase V, there is a considerable removal of CO_2, from 12% of CO_2 in feed, to 6% CO_2 content. According to Brazilian regulation agency (ANP, 2008), commercial NG is required to show a maximum of 3% mol CO_2, and phase V still does not fit within the specifications for marketing. This product, V@GNBRPS, is hence submitted to a second permeation. In phase L, CO_2 represents 67% of the permeate, and a recovery of 55% of CO_2 in phase L is achieved, requiring a further permeation module (corresponding to Stage 2) to increase the content of CO_2.

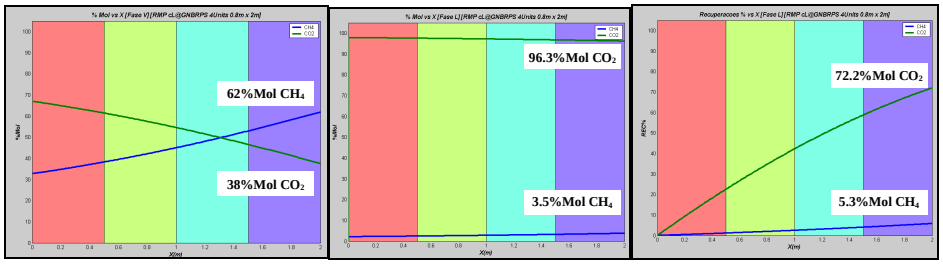


(a) (b) (c)

Figure 4: Stage 1 - (a) Profile (mol %) of phase V; (b) Profile (mol %) of phase L; (c) Profile (mol %) of CO_2 recovery in phase L

Simulation Results for Second Stage: As the content of CO_2 in phase L is not high enough to meet specification for injection into reservoirs, a second permeation stage is required. Furthermore, the pressure of permeate 1 is 1 bar, demanding recompression to 60 bar, by means of a train of compression as a single stage operating from 1 to 60 bar.
The profiles displayed in Figure 5 represent the mol% of the two gases studied (CO_2 and CH_4) in both phases of the process (Figures 5(a) and (b), respectively) and the profile of recovered CO_2 in phase L (permeate) (Figure 5(c)). In phase V, - stream V@cLGNBRPS, a good removal of CO_2 is obtained: from 67% of CO_2 in the feed to 38%. This stream unfortunately is not specified for use in any of the key objectives (commercial GN or CO_2 for injection), although it can be used as burning gas. For stage L, CO_2 reached nearly 96.3% of the composition of the permeate, a level acceptable to

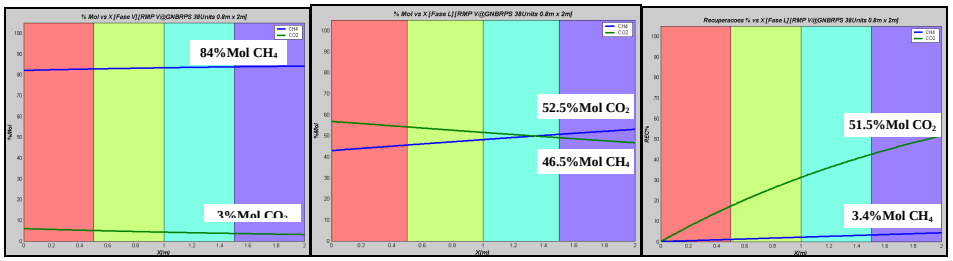
be used for injection in reservoirs. Therefore, it reached one of the objectives of this work, which would be to obtain a stream of CO_2 for injection. There was a recovery of almost 72.2% of CO_2 in phase L.



(a) (b) (c)

Figure 5: Stage 2 - (a) Profile (mol %) of phase V; (b) Profile (mol %) of phase L; (c) Profile (mol %) of CO_2 recovery in phase L

Simulation Results for Third Stage: This stage is fed with the retentate of Stage 1, which is at 59.62 bar, therefore not requiring further compression. This stream has 6% CO_2 and 82% of CH_4, equivalent to 90% of the original feed GNBRPS, justifying the use of 38 permeation modules in Stage 3. Figures 6(a) and (b) present the profiles of of CO_2 and CH_4 (mol%) in phases V and L, respectively. The profile of CO_2 recovery in phase L is shown in Figure 6(c). In phase V, stream V@V@GNBRPS shows a good removal of CO_2, based on a 6% CO_2 in the feed, slightly inferior to 3%, hence meeting the second process objective of this work - obtaining a specified stream to NG marketing. In phase L, CO_2 reached 46.5% of the composition of permeate and there was a recovery of nearly 51.5% of CO_2. This stream can still be used for burning or to turbines.



(a) (b) (c)

Figure 6: Stage 3 - (a) Profile (mol %) of phase V; (b) Profile (mol %) of phase L; (c) Profile (mol %) of CO_2 recovery in phase L

Table 1 2 summarize the results of GNBRPS simulated in PMS.

4. Conclusion

This work was presented and simulated a stream for purification by permeation membranes in a natural gas compatible with the scenario of the Pole Pre-Salt of the Santos Basin. Membrane modules of cellulose acetate in hollow-fiber were considered. The study found the technical feasibility of the use of membranes for CO_2 separation from it being a typical load of natural gas. This feasible technique is to obtain a stream

of treated natural gas suitable for marketing and a main stream of residual CO_2 suitable for geological injection.

Table 1: Stream specifications

Stream	MMm^3/d	P (bar)	T (^oC)	CO_2 (%)	CH_4 (%)
GNBRPS	4	60	25	12	77
V@GNBRPS	3.6	59.6	20	6	82
L@GNBRPS	0.4	1	19	67	32.7
cL@GNBRPS	0.4	60	35	67	32.7
V@cL@GNBRPS	0.2	59.7	4.5	38	62
L@cL@GNBRPS	0.2	1	2	96.2	3.5
V@V@GNBRPS	3.37	59.2	17.4	3	84
L@V@GNBRPS	0.23	1	17	46.5	52.5

Table 2: Simulation results

	MMm^3/d	CO_2 (%)	CH_4 (%)	RECOVERED CO_2 (%)	RECOVERED CH_4 (%)
Commercial NG	3.37	3	84	22%	91.9%
CO_2 for Injection	0.2	96.2	3.58	40%	0.2%
NG for Burning	0.2 + 0.23	38/46	62/52.5	38%	7.9%

Considering the simulation results of the presented case study, this work demonstrates the technical feasibility of applying membrane technology in offshore installations. In addition to the listed advantages of membrane technologies over traditional absorption columns, specified levels of CO_2 and CH_4 are herein proven to be achievable though the use of MPS, meeting process specifications and complying with environmental regulations.

References

Agencia Nacional de Petróleo e Biocombustível. Resolução 16, DOU 17.06.2008

R. W. Baker, K. Lokhandwala,2008. Natural gas processing with membranes: An Overview. Ind. Eng. Chem. Res., 41, p. 2109 – 2121

H. K. Engelien, 2004 Process Integrated Membrane Separation - with Application to the Removal of CO2 from Natural Gas. Department of Chemical Engineering, NTNU.

J. Formigli, 2007. Pre-Salt Reservoirs Offshore Brazil: Perspectives and Challenges

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal Economic Decision Making for Gas-to-Liquid Product Selection considering Competition in Market Dynamics

Chul-Jin Lee and Chonghun Han

School of Chemical and Biological Engineering, Seoul National University, San 56-1, Shillim-dong, Kwanak-gu, Seoul, 151-742, Korea

Abstract

For a chemical process industry, an invested capital cost is too large and when a decision making that which product should be produced is once determined, it is almost impossible to be modified or repeated. Also, since the huge uncertainty exists in the whole market such as recession today, it is so important to judge an economic decision making for selection of which target product would be the most profitable in the future. For the sake of a robust decision, we need to compare several points of view including profitability, reliability, risk, and so on. In this paper, we will try to get an optimal economic decision making for Gas-to-Liquid product selection considering market dynamics. The targets are three GTL products; FT-diesel, DME, and Methanol. At first we could design the three GTL processes using a process simulator. Then, through the constructed models, comparative economic analysis among them was conducted using an economic index. For a robust analysis we can apply the market dynamics related with demand expansion and shrinkage as well as a competition within players. Finally, we can get an optimal solution for selecting GTL product considering market variation and compare the robustness for the decision.

Keywords: decision making, process design, economic analysis, Gas-to-Liquid, Market dynamics

1. Introduction

There have been several economic analyses for GTL processes consuming natural gas as feedstock. Han et al. (2002) assessed an economic evaluation for methanol synthesis process and Zhou et al. (2007) studied the economic and environmental influences of DME process [1], [2]. In general, however, many people concentrated on individual analysis for a single GTL product, so there were few studies considering a comparative economic analysis to select the most profitable product using natural gas as raw material. Along the way Gradassi et al. (1995) compared the economic analysis of several GTL products, and Morita (2001) concentrated on three profitable GTL products (DME, FT-diesel, Methanol) used for transportation fuels. However, these studies did not cover the characteristics of process industry such as the fact that the profitability of each product is largely dependent on the raw material cost [3], [4]. Lee et al. (2009) focused on the influence of natural gas cost variation as raw material on the profitability of GTL products and found that the order of profitability of GTL products could be reversed according to the raw material cost variation [5]. However, the profitability of chemical products is largely influenced by consumer market. We need to consider the market dynamics of GTL to get an optimal economic decision for selecting GTL product.

In this paper, we modeled three GTL processes producing DME, FT-diesel, and Methanol using ASPEN PLUS™ and performed an economic assessment of each of them. Furthermore considering the market variations of each product, we analyzed the profitability of them using net present value (NPV) as an economic index. As a result, the priority of profitability was reversed according to the market variations in the future.

2. Modeling and Simulation

The GTL process is largely divided into two parts. At first we could produce the synthesis gas which is a gaseous mixture of hydrogen and carbon monoxide from natural gas as raw material, which is called to reforming process. Then, we can synthesize the target chemical product such as FT-diesel, Methanol, and DME. Both stages of reaction go through in high temperature (~1000℃) and pressure (~3000kPa), the Peng-Robinson EOS was used as a thermodynamic package model. The detailed models and results for three GTL processes were described by Lee et al. [5].

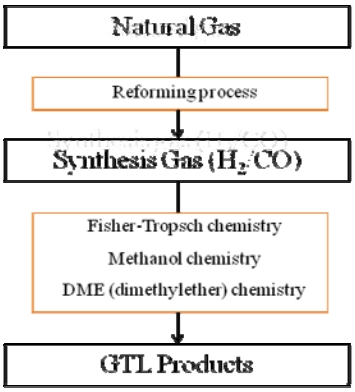


Figure 1. Schematic diagram of the Gas-to-Liquid process

3. Economic Evaluation

3.1. Previous economic analysis summary

Through the constructed process model, we can calculate a required capital and operating cost. The evaluation result of each product is shown in Table 1. The detail of each is presented by Lee et al. [5].

As the payout time means the required time to recover an invested capital, the case of the lower payout time means the more profitable one. Thus, producing FT-diesel is the most profitable among three GTL products. In this case, we assumed that all of the produced products are sold in the market. However, there can be a competition between the producers in market or the whole market size would be expanded or shrink according to the boom days or recession. To obtain an optimal decision making to select a prosperous GTL product, we need to consider a market competition and variation over time series.

Table 1. Summary of an economic analysis for three GTL products.

	unit	FT-diesel	MeOH	DME
revenue	$\$x10^6$/yr	\$1,038	\$844	\$767
capital cost	$\$x10^6$	\$2,374	\$2,254	\$1,289
utility cost	$\$x10^6$/yr	\$8	\$18	\$11
NG cost	$\$x10^6$/yr	\$606	\$606	\$606
transportation cost	$\$x10^6$/yr	\$21	\$49	\$17
payout time	yr	5.91	13.24	9.76

3.2. Net Present Value

For an economic index in this paper, we used the Net Present Value (NPV) to consider a variation of cash flow to the investment period [6]. The value of minimum expected rate of return, r, was set to 5.7%, by taking account of the yield of treasury bonds over 10 years in South Korea [7].

$$NPV=\sum_{i=1}^{N} \frac{CF_i}{(1+r)^i} - I \tag{1}$$

4. Market Competition

Considering the market dynamics, we can assume the competition in each GTL market. For instance, the stronger player in GTL market will occupy more market share annually than the weaker one. Thus, the sales quantity of each would be varied to the expansion rate of market occupation of each, which can make change the revenue of them.

4.1. Assumption

(1) There are three players(P1, P2, P3) for FT-diesel, MeOH, and DME markets individually.

(2) The whole market size is fixed for each product during investment period.

$$Q_A = Q_{A1} + Q_{A2} + Q_{A3} \tag{2}$$
$$Q_B = Q_{B1} + Q_{B2} + Q_{B3} \tag{3}$$
$$Q_C = Q_{C1} + Q_{C2} + Q_{C3} \tag{4}$$

(3) Among the three players in each product, there exists a competition which can make vary the market portion of them to 10%, 0%, and -10% annually.

$$\frac{dQ_{Aj}}{dt} = \alpha_i \text{, i = 1, 2, 3 for each player in FT-diesel market} \tag{5}$$
$$\frac{dQ_{Bj}}{dt} = \beta_i \text{, i = 1, 2, 3 for each player in MeOH market} \tag{6}$$
$$\frac{dQ_{Cj}}{dt} = \gamma_i \text{, i = 1, 2, 3 for each player in DME market} \tag{7}$$

4.2. Analysis results in considering competition among players in each market

The trends of NPV result for each GTL product are analogous to the previous economic result, which used the payout time as an economic index shown in Table 1. When the market of FT-diesel annually expands to 10%, the value of NPV of FT-diesel is slightly larger than DME case as depicted in Figure 2 and 4 (α1 and γ1 respectively). However, when the market portion shrinks to -10% every year, which means the falling on competitive power or recession of the whole market, the difference of NPV is quite increased that a negative NPV of FT-diesel is almost double comparing DME value, which indicates the possible loss of FT-diesel production would be larger than DME case.

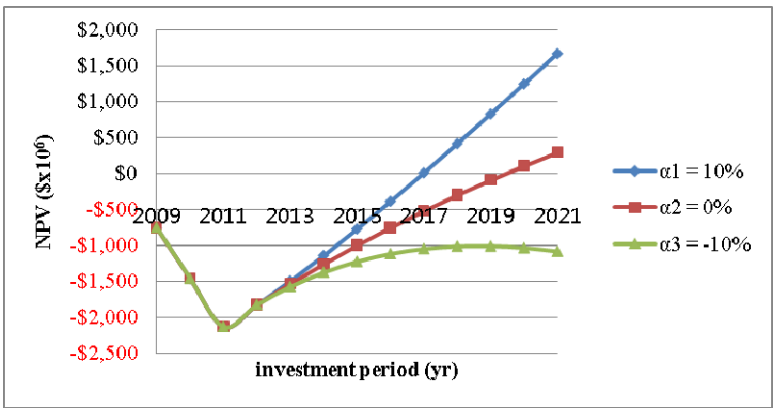


Figure 2. NPV variation of FT-diesel production over the service life for annual expansion rate of market share for FT-diesel player 1, 2, 3.

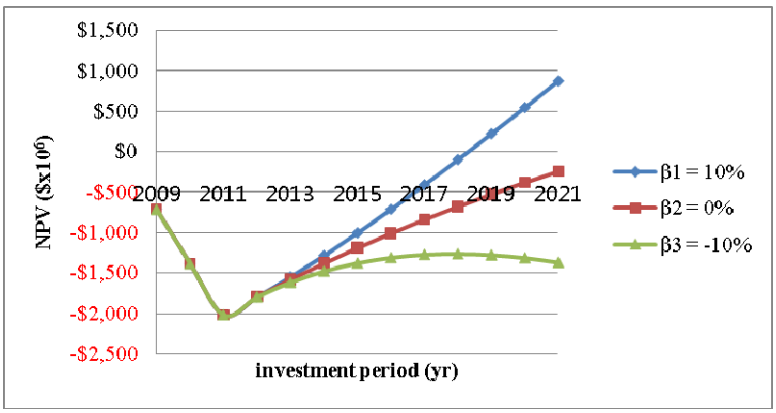


Figure 3. NPV variation of MeOH production over the service life for annual expansion rate of market share for MeOH player 1, 2, 3.

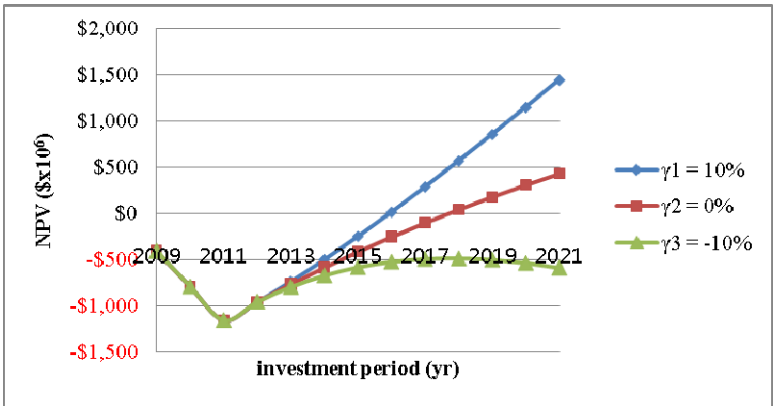


Figure 4. NPV variation of DME production over the service life for annual expansion rate of market share for DME player 1, 2, 3.

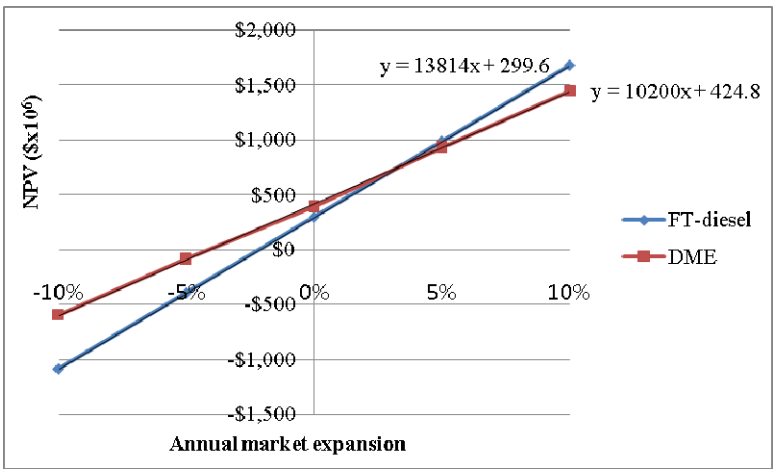


Figure 5. NPV comparison between FT-diesel and DME according to the annual market expansion.

Thus far, we could analyze the economic feasibility among three GTL products considering market competition among players. Then, we can guess how the profitability can be changed in taking account of the variations of whole market size. When there may be prosperous condition or recession in the whole market, which product would be the most robust decision? We can get the answer in Figure 5. Because the MeOH production always gets the lowest profitability, we can concisely focus on economic comparison between FT-diesel and DME to select an optimal product.

When we assume that both markets of FT-diesel and DME would be influenced by the same amount for expansion or shrinkage, we can get the correlation function between NPV and annual market expansion rate. The intersection position for both NPV graphs points out an expected market growth rate balancing the profitability of them. When the both markets grow more than 3.6% per annum, producing FT-diesel can get more profit. On the other hand, it is the better way to choose DME below 3.6% per annum, and further, it can reduce the expected loss at the lower position than the break-even point at zero value of NPV.

5. Conclusion

In this article, we investigated the optimal economic decision making for Gas-to-Liquid product selection when a competition between players in each market exists. Furthermore, when we expect the expansions or shrinkage of the whole market size, we can compare the profitability among GTL products. When the market expands over 3.6% per annum, it is better to choose the FT-diesel. On the contrary, DME production is more profitable below 3.6% per annum. Suggested methodology including market dynamics would be especially useful to decision makers when the risk and uncertainty of the market environment considerably matters.

Nomenclature

d_i = discount factors
CF_i = after taxes cash flow of period
I = the capital investment
r = minimum expected rate of return
Q_A = production quantity for FT-diesel
Q_B= production quantity for MeOH
Q_C= production quantity for DME
Q_{Ai} = production quantity of player i for FT-diesel
Q_{Bi} = production quantity of player i for MeOH
Q_{Ci} = production quantity of player i for DME
α_i = annual increase rate of market portion of player i for FT-diesel
β_i = annual increase rate of market portion of player i for MeOH
γ_i = annual increase rate of market portion of player i for DME

References

[1] Han, S. R.; Chung, Y.; Joo, O.; Jung, 2002, K. Economics of methanol synthesis process. Theor. Appl. Chem. Eng., 8, 3113-3116.

[2] Zhou, L.; Hu, S.; Li, Y.; Zhou, Q., 2008, Study on co-feed and co-production system based on coal and natural gas for producing DME and electricity. Chem. Eng. J., 136, 31-40.

[3] Gradassi, M. J.; Green, N. W., 1995, Economics of natural gas conversion processes. Fuel Process. Technol., 42, 65-83.

[4] Morita, Y., 2001, Marketability of GTL from Natural Gas; The Institute of Energy Economics: Japan.

[5] Lee, C. –J.; Lim, Y.; Kim, H. S.; Han, C., 2009, Optimal Gas-to-Liquid Product Selection from Natural Gas under Uncertain Price Scenarios. Ind. Eng. Chem. Res., vol. 48, no. 2., pp. 794-800.

[6] Miguel, Bagajewicz, 2008, On the Use of Net Present Value in Investment Capacity Planning Models. Ind. Eng. Chem. Res., vol. 47, no. 23., pp. 9413-9416.

[7] Lim, S. –R.; Park, D.; Lee, D. S.; Park, J. M., 2006. Econmic evaluation of a Water Network System through the Net Present Value Method Based on Cost and Benefit Estimation. Ind. Eng. Chem. Res., vol. 45, pp. 7710-7718.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Bioprocess Systems Engineering applied to the Production of Protein Hydrolysates in a Multipurpose Plant

Gilson A. Pinto,[a] Roberto C. Giordano[b]

[a]*Chemtech – a Siemens Company, Av. Ermano Marchetti, 1435, 05038-001, São Paulo – SP, Brasil*
[b]*Department of Chemical Engineering, Federal University of São Carlos, Rod. Washington Luis, Km. 235, São Carlos – SP, P.O.Box 676, Brasil*

Abstract – The continuous advances in process computing have provided in recent years several sophisticated tools for process analysis and simulation. The effective use of these tools demands the capability to interface with a pre-existent process system hierarchy. Within this perspective, an important issue is to provide meaningful and useful simulation and optimization applications for complex systems that require integration with data-intensive experimentation. This study proposes an integrated environment, using Internet as development platform, for simulation, monitoring, control and optimization of a cheese whey refinery, employing immobilized and stabilized enzymes as catalysts. The multipurpose process described here, the *cheese whey biorefinery*, provides, besides lactose, whey protein concentrates and hydrolysates that can be applied in food and pharmaceutical formulae. Under these circumstances, it is possible to add value to this significant by-product of the dairy industry, avoiding its disposal *in natura* (what is mostly done by small cheese manufacturers). In parallel, relevant aspects of the enzymatic reactions within the cheese whey biorefinery were investigated. The lumping of substrate molecules in pseudo-components, using a hybrid phenomenological-neural approach for description of the enzymatic de-polymerization kinetics, is the suggestion to follow the complex reaction dynamics. During the estimative of model parameters and stationary state simulation/optimization of the multipurpose plant, non-linear search algorithms were implemented and evaluated (Successive Quadratic Programming, Simulated Annealing, Differential Evolution and Particle Swarm Optimization). The work, thus, is concerned with two of the main dilemmas/challenges of the present millennium: reduction of environmental problems related to the disposal of agro-industrial residues (i.e., cheese whey) and development of processes for food production from alternative sources (whey refinery). All implemented algorithms for remote process monitoring and control were validated with data from laboratory-scale assays. The validation of the hydrolytic kinetic models followed the same procedure. According to the achieved results, is possible to conclude that the robustness of the integrated computational environment was demonstrated. The prediction capability of the approaches employed for description of the proteolytic enzymatic reactions was verified, as well.

Keywords: cheese whey biorefinery, enzymatic kinetic modeling, nonlinear optimization, Internet based application.

1. Introduction

The concept of a "biorefinery" is usually tied to the processing of lignocellulosic biomass to produce fuels and chemicals. Nevertheless, if we have in mind an oil refinery as our paradigm, any plant that has a pool of macro- and micromolecules for feedstock, providing a set of different products (usually, pools of molecules), using bioreactors (either enzymatic or biochemical – "fermenters") for the "refining" of the raw materials, any of these plants may be named a biorefinary as well. It is within this broader concept that we report in this work the application of (bio)process systems engineering tools to what we call a "cheese whey biorefinery".

A simplified block diagram of the whey biorefinery, as proposed in Pinto *et al.*, 2009, is shown in Fig. 1. In this work, we detail the integration of algorithms for plant-wide optimization, bioprocess remote monitoring and control (via internet), and hybrid models of complex de-polymerization enzymatic reactions into a user-friendly environment. The originality of the concept is putting together a static simulator (with design, optimization and cost analysis tools) and an applicative for remote monitoring and control of the bioprocess.

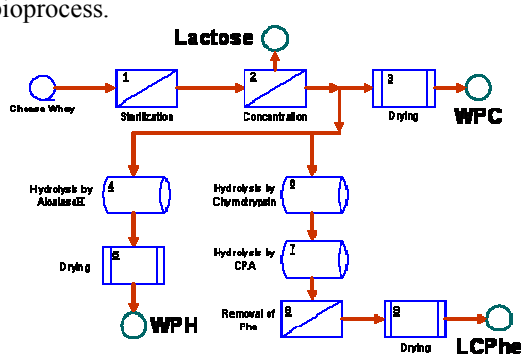


Fig. 1 – Cheese whey biorefinery block diagram.

The enzymatic customized hydrolysis of cheese whey may provide a collection of value-added products: Advances in bioprocess technologies have enhanced the use of cheese whey for commercial applications such as: edible films, whey protein concentrates, functional foods, athletes' supplements, clinical diets and pharmaceuticals (Clemente, 2000). In this way, a strong pollutant disposal of the dairy industry could be transformed into commercial products. Indeed, whey BOD is 35000 $mgO_2.L^{-1}$, approximately 10^2 times greater than domestic sewage's BOD; Siso, 1996.

In Fig. 1, raw cheese whey is micro- and ultra-filtered, and lactose is separated, following to further enzymatic hydrolysis (not shown). A whey protein concentrate is a possible first product of the biorefinery, after spray-drying (WPC). Alternatively, WPC may go through a series of controlled proteolyses with immobilized enzymes. Supplements for parenteral feeding come from the hydrolysis with alcalase®, providing a pool of small peptides with lower osmolality than a mixture of amino acids – WPH, Whey Protein Hydrolysates (Pinto et al., 2007). A different product, LCPhe (Low Contents of Phenylalanine) can be produced after two sequential enzymatic hydrolyses with chymotrypsin and carboxypeptidase A, with subsequent separation of amino acids and remaining lactose (Cabrera-Padilla et al., 2009; Galvão et al., 2009). This product can be used for feeding phenylketonuria patients.

The computational applicative previously mentioned has tools for the preliminary design, simulation, remote monitoring and control. The simulation module of the web application was developed to provide stationary mass balances, preliminary design and cost analysis of the main equipments. Economic analysis of the process is another

feature of the application, which includes an optimization module for the multipurpose plant. In the integrated Internet-based environment here presented, on-line monitoring and control modules were also put forth. The proteolytic steps of the biorefinery were investigated within this perspective, since real-time inference of the product composition is crucial for the fine-tuning the control of the enzymatic reactors. Given that mechanistic kinetic models of the enzymatic reactions would be very complex, hybrid neural-network models were employed. The prediction capability of the models and their implementation within the web application for real-time monitoring of the reactors are presented here.

2. Software Development and Implementation

2.1. Non-linear Search Algorithms

The approach used here was to implement the stochastic methods described below, in Matlab 6.5 (MathWorks, Inc.), coupled to a set of heuristic corrections for each implementation. The main idea was to use these techniques as tools for a real optimization problem, rather than concentrating a major effort in tuning the algorithms.

The initial temperature, T^0, of simulated annealing (SA) was selected according to Kirkpatrick et al., 1983, computing a temperature such that the initial acceptance ratio is approximately equal to a given value λ_0 (here, $\lambda_0 = 0.8$). First, a large value for T^0 is chosen. Then, a number of transitions using this temperature are performed. The ratio of accepted transitions is compared with λ_0. If it is less than the acceptance ratio, then T^0 is multiplied by 2. The procedure continues until the observed acceptance ratio exceeds λ_0.

The cooling scheduling for each iteration k was $T^{k+1} = 0.98T^k$. The neighborhood structure was constructed by simple random values constrained by the lower and upper bounds of parameters. The implementation of particle swarm optimization (PSO) followed the pseudo-code described by Kennedy and Eberhart, 1995. Introduced by Shi and Eberhart, 1998, the inertial weight w – not present in the original form – was also used. According to the authors, the parameter w can assure the PSO convergence provided it decreases linearly at each iteration k, according to $w^k = 0.9 - 0.5(k/it_{\max})$. The implemented strategy for Differential Evolution (DE) was DE/rand/1/binomial, as described by Storn and Price, 1997. The algorithm *fmincon* from Matlab 6.5 was used for SQP (sequential quadratic programming).

The optimization of stationary states of biorefinery – each whey flow rate defines a problem (stream *1* in Fig. 1) – is solved with the definition of two auxiliary variables, showed in the process flowsheet described in Fig. 1: p_{WPC} and p_{WPH}, respectively, the percentage of stream *7a* that will be directed to WPC production and the percentage of stream *15a* that will be used for WPH production. The net present value of the project, *NPV*, is implicitly a function of p_{WPC} and p_{WPH}, since these variables define stationary process flow rates and, consequently, the costs and yields involved in the biorefinery project (equipment dimensions, capital costs, etc.). Therefore, the nonlinear programming problem can be stated as:

$$\underset{p_{WPC}, p_{WPH}, \tau_{R-101}, \tau_{R-102}, \tau_{R-103}}{Min} \{\phi\}, \qquad \text{(Eq. 1)}$$

$$\phi = [-NPV + (F_{13} - MD_{WPC})\,\mathrm{Price}_{WPC} + (F_{20} - LM_{WPH})\,\mathrm{Price}_{WPH} + (F_{31} - LM_{LCPhe})\,\mathrm{Price}_{LCPhe}]$$

where MD_{WPC}, MD_{WPH} and MD_{LCPhe} are, respectively, market demand constraints for products WPC, WPH and LCPhe; F_{13}, F_{20} and F_{31} are the mass flow rates of the same products; and τ_{R-101}, τ_{R-102} and τ_{R-103} are the operation times of reactors present within biorefinery.

In an attempt to avoid local minima, the three implemented stochastic methods and SQP were applied in parallel for the optimization problem described in Eq. 1. The results were compared in section 3.

2.2. Computation Solution for the Web-based Application

The model architecture of the web application described here employs Java technology (Servlets and Java Server Pages; Sun Microsystems, Inc.). It is basically a Model-View-Controller (MVC) and Layers architecture. MVC has become important in user-interface programming because emphasizes a design that preserves disjoint contents and layout (Deitel and Deitel, 2007). The *Information Layer* provides data to the application (process data for example). Client requests, data processing and contents presentation are in charge of the *Application Logic Layer*. The *Client Layer* is the interface with the user. Applets were used to provide appealing graphical user interface (GUI) through the web. An applet is a program written in Java that is intended not to be run isolated, but rather to be embedded inside another application like a client browser, since it can be in included in an HTML page.

We used Tomcat version 5.5.17 as the application server. Tomcat is an open-source servlet container developed under the Apache license. Its primary purpose is to serve as a reference implementation of the Servlets and Java Server Pages specifications.

The algorithms written in Matlab® 6.5, MathWorks, Inc., are called from the server with the open-source class JMatLink (Müller and Waller, 1999). The server database was MySQL, a popular open-source SQL relational database management system developed, distributed, and supported by MySQL AB. Complete the computational solution for our case study the software Labview® (version 7.1) and the board NI PCI-6052E for data acquisition (National Instruments, Inc.), both installed on a standalone computer. The TCP/IP protocol was applied to data transfer between Labview and application server.

2.3. Hybrid Phenomenological-neural Modeling

The reaction medium within enzymatic reactors present in biorefinery characterizes a complex mixture of amino acids, proteins and peptides. In addition, the variability of its composition is high, depending not only on the upstream process conditions, but also on climate and seasonal factors that affect the original characteristics of the raw material, bovine milk. In this scenario, a mechanistic based model of these reactions would be intractable, leading to high-correlated parameters.

Combination of mass balances with the prediction capability of neural networks may supply hybrid models able to capture the behavior of enzyme kinetics (Pinto et al., 2007; Rivera et al., 2007). Multilayer perceptrons (MLP) neural networks were trained to map the states of each reactor – molar concentrations of components – into the vector of their reaction rates. Differential mass balances of each component were solved numerically, using the MLP predictions of reaction rates. The structure of hybrid approach and the methodology here put forth are described in Pinto et al., 2007.

3. Results and Discussion

Fig. 2a compares performances of the implemented algorithms during optimization of whey biorefinery. Algorithms DE and PSO achieved the best Net Present Values for stationary state simulations while SA only achieved sub-optimal solutions.

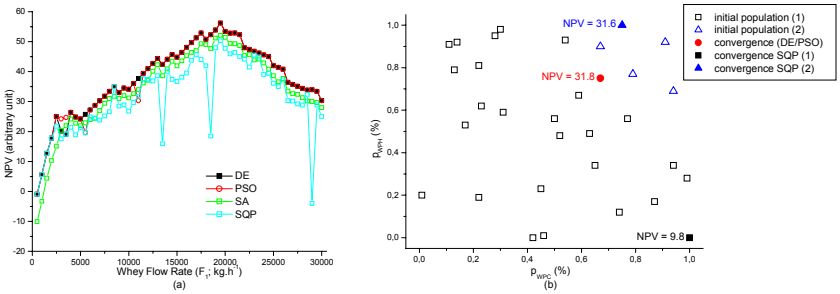


Fig. 2 – Net Present Value as a function of whey flow rate (each whey flow rate defines an optimization problem). **(a)** Performances of algorithms DE, PSO, SA and SQP during the optimization of whey biorefinery (stationary states). **(b)** Performances of DE, PSO and SQP for a whey flow rate of 7500 kg.h^{-1}. In figure, NPV is showed in arbitrary units.

As Simulating Annealing, SQP algorithm also showed poor performances during whey biorefinery optimization. The gradient-based approach only converged to the bound constraints of the variables p_{WPC}, p_{WPH}, $\tau_{R\text{-}101}$, $\tau_{R\text{-}102}$ and $\tau_{R\text{-}103}$, while best results achieved by DE and PSO were obtained inside the feasible region. Fig. 2b illustrates this behavior for one case (whey flow rate of 7500 kg.h^{-1}). Initial population of DE and PSO was employed as initial guesses for SQP. Note that while PSO and DE converged to the same result inside the feasible region (the "best" NPV – 31.8), SQP attained two different results in the contours of the same search space (31.6 and 9.8).

Figure 3a presents the web application GUI during the monitoring of a proteolysis to remove phenylalanine from the mixture of peptides, providing LCPhe. An appealing and consistent, real-time updated description of the process was attained. With the hybrid modelling approach, we believe to be possible to follow the course of the enzymatic reactions with real-time accurate predictions, avoiding lengthy off-line HPLC analyses (Pinto et al., 2007).Figure 3b illustrates the control of the pH of the enzymatic hydrolysis of WPC by alcalase®. Possible dynamic time delays caused by the Internet traffic were also overcome by the Internet-based architecture.

The simulation module of the software also allows an economic analysis of the process, including the impact of the process costs (such as enzyme prices and life times, product market demands, etc.). Subject to market demand constraints for the biorefinery products, it is also possible to simulate optimal stationary states of the process using the implement nonlinear search algorithms (DE or PSO for instance).

4. Conclusion

A web-based remote application for economic evaluation, monitoring and control of a whey biorefinery was presented in this paper. In a single integrated environment, is possible to perform real-time operations, such as control and monitoring of operational units, and to simulate stationary states of the biorefinery. The simulation module provides preliminary design of the major equipments, estimative of their fixed and operational costs, economic evaluation of the entire plant and impact of process and market indices variations on process economics. Nonlinear search algorithms (SA, DE and PSO), for process design and simulation were also developed and implemented within the application. With the simulation and optimization modules, the application can act as a decision-support tool for the analyses of different dairy market scenarios.

Hybrid phenomenological-neural models were incorporated as part of the real-time monitoring system, conferring accurate predictions of the components profiles along

reactions. The monitoring of the three enzymatic reactions of the biorefinery, crucial steps of the plant, was an important case study for the web application, which provided an appealing and reliable depiction of the enzymatic reactors.

(a) (b)

Fig. 3 – Graphical user interfaces (GUI) of the Web application. (a) On-line remote monitoring of the Phe released during hydrolysis with CPA. (b) Remote pH control of WPC production: servo problem.

Acknowledgments

The authors thank the Brazilian research funding agencies FAPESP, FAPESP/Tidia-Kyatera, CNPq and PADCT/CNPq.

References

Cabrera-Padilla, RY; Pinto, GA; Giordano, RLC, et. al. 2009, A new conception of enzymatic membrane reactor for the production of whey hydrolysates with low contents of phenylalanine. *Process Biochem*, 44, 4, 269–276.

Clemente, A, 2000, Enzymatic protein hydrolysates in human nutrition. *Trends Food Sci Tech*, 11, 7, 254–262.

Deitel, HM; Deitel, PJ. 2007, *Java How to Program*. 7 ed., Prentice-Hall, New Jersey.

Galvão, CMA; Pinto, GA; Jesus, CDF; et.al. 2009. Producing a phenylalanine-free poll of peptides ater tailored enzymatic hydrolyses of cheese whey. *J Food Eng*, 91, 1, 109–117.

Kennedy, J; Eberhart, RC. 1995, Particle swarm optimization. In: *Proceedings of IEEE International Conference on Neural Networks*. Perth, Australia, 4, 1942-1948.

Kirkpatrick, S; Gelatt, CD; Vecchi, MP. 1983, Optimization by simulated annealing. *Science*, 220, 4598, 671–680.

Müller, S; Waller, H. 1999, *Efficient integration of real-time hardware and web based services into Matlab*. Available in http://jmatlink.sourceforge.net/docs/ESS99.pdf.

Pinto, GA; Giordano, RLC; Giordano, RC. 2007, Neural network inference of molar mass distributions of peptides during Tailor-made enzymatic hydrolysis of cheese whey: effects of pH and temperature. *Appl Biochem Biotech*, 143, 2, 142–152.

Pinto, GA; Giordano. RLC; Giordano, RC. 2009, Remote engineering for a cheese whey biorefinery: an Internet-based application for process design, economic analysis, monitoring, and control of multiple plant sites. *Bioproc Biosys Eng*, 32, 1, 67-98.

Rivera, EC; Costa, AC; Andrade, RR; et. al. 2007, Development of adaptive modeling techniques to describe the temperature-dependent kinetics of biotechnological processes. *Biochem Eng J*, 36, 2, 157–166.

Shi, Y; Eberhart, RC. 1998, A modified particle swarm optimizer. In: *Proceedings of the IEEE International Conference on Evolutionary Computation*, Anchorage, Alaska, 69–73.

Siso, MIG. 1996, The biotechnological utilization of cheese whey: a review. *Bioresource Technol*, 57, 1, 1–11.

Storn, R; Price, K. 1997, Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *J Global Optim*, 11, 4, 341–359.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimization of bioethanol distillation process – evaluation of different configurations of the fermentation process

Marina O. S. Dias,[a] Tassia L. Junqueira,[a] Rubens Maciel Filho,[a] Maria R. W. Maciel,[a] Carlos E. V. Rossell,[b] Daniel I. P. Atala[c]

[a] *School of Chemical Engineering, State University of Campinas – UNICAMP, P.O. Box 6066, Campinas, SP 13083-970, Brazil*
[b] *Interdisciplinary Center for Energy Planning, State University of Campinas, UNICAMP, P.O. Box 6192, 13400-970, Campinas – SP, Brazil*
[c] *Sugar Cane Technology Center, Fazenda Santo Antônio, 13400-970, Piracicaba-SP, Brazil*

Abstract

Process simulation was used to analyze bioethanol distillation process, which requires a large amount of thermal energy. As it is shown in this study, in the ethanol production process the fermentation stage has a significant impact on energy consumption in the purification step. Thus, alternative configurations in the fermentation and distillation processes were proposed and evaluated. The results showed that vacuum extractive fermentation coupled with triple effect distillation presented the lowest energy demand among the studied configurations.

Keywords: bioethanol, fermentation, distillation, simulation, multiple effect distillation.

1. Introduction

Increase in oil prices and growing international concern about climate change and greenhouse gases emissions have motivated the use of renewable fuels, such as bioethanol, as either a gasoline additive or its complete substitute (Balat et al., 2008). Corn and sugarcane are the feedstock used in the US and in Brazil, respectively, which are the largest ethanol producers in the world. Sugarcane is so far the most efficient raw material for bioethanol production: the consumption of fossil energy during sugarcane processing is much smaller than that of corn (Macedo et al., 2008). Furthermore, optimization of bioethanol production process from sugarcane is still possible, and significant reduction of energy consumption can be achieved through the optimization of fermentation and distillation processes.

Two configurations of the fermentation process are most frequently employed in Brazil: feed-batch and continuous, both with cells recycle. Average ethanol content of the wine in several production units in São Paulo state reaches only 9 % (v/v), varying between 7 and 10 % (v/v). This relatively low ethanol content is a consequence of the inhibitory effects of both ethanol and substrate on the yeast cells, among other factors; as a result, a large consumption of energy is required during the purification step, on which a series of distillation and rectification columns are used.

Several techniques to remove ethanol from fermentation broth have been studied, for instance, Bui et al. proposed selective adsorption on silicalite, active carbon and ZSM-5 and concluded that silicate appeared to be the most suitable adsorbent for this process.

Other alternative is the vacuum extractive fermentation, which consists in the continuous removal of ethanol from the fermentation medium by integrating the fermentation reactor to a vacuum flash evaporator (Silva, 1999 and Atala, 2004). This configuration allows the use of higher substrate concentrations in the reactor, compared to conventional fermentors, and the removal of ethanol from the fermentation media keeps ethanol concentration in the reactor in low levels (around 5 % v/v), which has a low inhibitory effect. In addition, wine removed from the flash evaporator contains a higher ethanol content, what decreases energy consumption during the purification step.
In this work, simulations of both conventional and vacuum extractive fermentation processes, as well as the purification step for each case, were carried out using software Aspen Plus. Different distillation columns configurations were studied: conventional, double effect and triple effect. The alternative configurations were evaluated and compared to the conventional ones in terms of energy consumption.

2. Fermentation

2.1. Conventional fermentation

In the conventional fermentation process sugarcane juice is used as raw material, with a concentration of sugars around 200g/L. Sugars are converted into ethanol, CO_2, acetic acid and glycerol. Wine concentration is relatively low (around 8 wt% ethanol).

2.2. Vacuum extractive fermentation

The vacuum extractive fermentation process consists of a fermentation reactor coupled to a vacuum flash evaporator, which allows ethanol produced to be simultaneously removed from the reactor. Thus, ethanol concentration in the reactor remains at low levels (around 8 wt% ethanol), while the wine obtained in the flash chamber contains a concentration equal to approximately 36 wt% ethanol. Because ethanol is continuously removed from the fermentation media, sugarcane juice may have a larger concentration of sugars (around 450 g/L), which is more than two times greater than that of the conventional process. Configuration of this process is depicted in Figure 1.

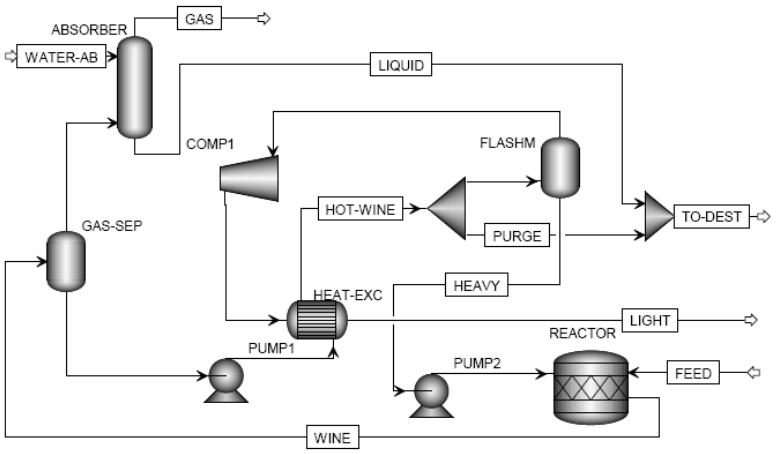


Figure 1 .Configuration of the vacuum extractive fermentation process.

2.3. Simulation data

In both cases, sucrose is almost entirely converted into glucose and fructose (conversion of 99 %); 99.5 % of the reduced sugars are converted into ethanol and CO_2, and the remaining fraction is used as source for the production of by-products (acetic acid and

glycerol), represented by equation 1 (Franceschin et al., 2008):

$$2C_6H_{12}O_6 + H_2O \rightarrow C_2H_4O_2 + 2C_3H_8O_3 + C_2H_5OH + 2CO_2 \quad (1)$$

3. Distillation

In the conventional fermentation process, only one ethanol-rich stream (wine) is produced, as opposed to the alternative process, on which two streams are obtained ("light" and "to-dest", comprised by flash vapour phase and a mixture of a reactor purge and alcoholic solution from the absorber, respectively, as shown in Figure 1). Thus, the distillation columns for each case may have different configurations, due to the characteristics of the feed streams.

3.1. Conventional distillation process

The conventional configuration considered in this work is based on a typical industrial scale process, on which a series of distillation (A, A1 and D, located above one another) and rectification columns (B and B1) are employed to produce hydrous bioethanol (around 93 wt% ethanol). In this process, the feed streams are fed to the distillation columns, on which vinasse, vapour and/or liquid phase phlegms and gases are produced: liquid phlegm is a bottom product of column D and vapour phlegm is produced near the top of column A. Phlegms are fed to the rectification columns, producing hydrous ethanol and phlegmasse. All columns work under pressures around atmospheric. This simulation was taken as base for the alternative configurations.

3.2. Double effect distillation process

The double effect process is similar to the conventional configuration, but the distillation columns operate under vacuum (19 – 25 kPa), while rectification columns operate under atmospheric pressure (101 – 135 kPa). In this way, different temperature levels are observed between column A reboiler and column B condenser (65 and 78 ºC, respectively), allowing thermal integration of these equipments and consequently reducing energy consumption on the distillation stage. Configuration of this process applied to the alternative fermentation is depicted in Figure 2.

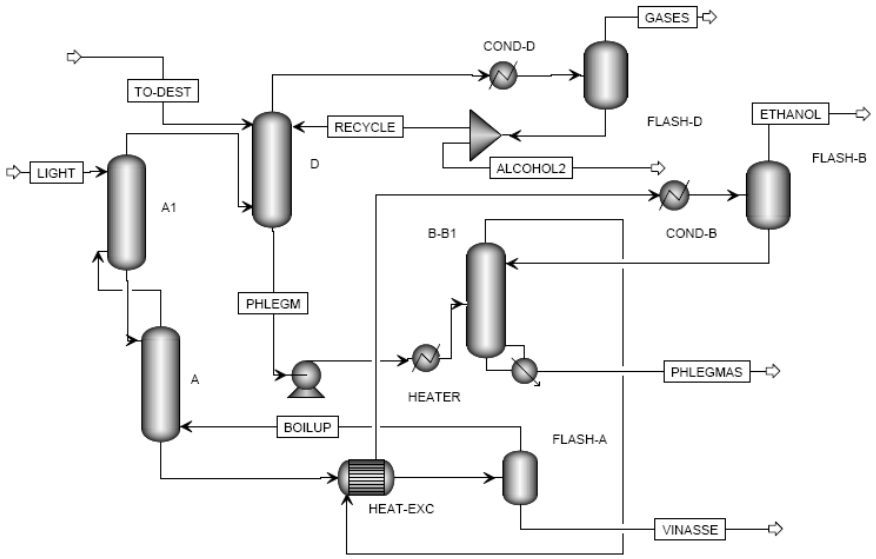


Figure 2.Configuration of the double effect distillation process.

3.3. Triple effect distillation process

In the triple effect configuration, the distillation columns operate under vacuum (19 – 25 kPa), and the liquid phlegm stream produced on column D is split in two: one of them is fed to a rectification column operating under nearly atmospheric pressure (column "B", 70 – 80 kPa) and the other is fed to a rectification column which operates under relatively high pressure (column "B-P", 240 – 250 kPa). Thermal integration between columns reboilers and condensers are possible, as represented on Figure 3.

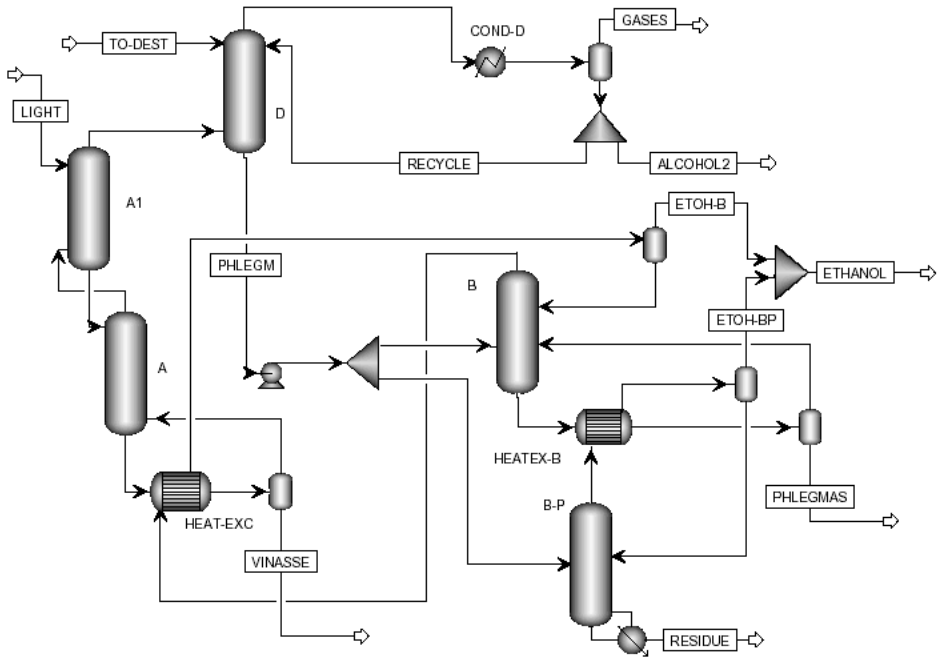


Figure 3. Configuration of the triple effect distillation.

4. Simulations and results

Six different configurations were simulated and evaluated: conventional fermentation and distillation with liquid phlegm only (CFCD), conventional fermentation and distillation with both liquid and vapour phase phlegms (CFCDP), vacuum extractive fermentation and conventional distillation with liquid phlegm only (VFCD), vacuum extractive fermentation and conventional distillation with both liquid and vapour phase phlegms (VFCDP), vacuum extractive fermentation and double effect distillation (VFDD) and vacuum extractive fermentation and triple effect distillation (VFTD). Sobočan and Glavič (2000) also performed simulations of distillation columns for ethanol production and verified that thermally integrated columns provide a reduction in annual costs. In this work, energy consumption of the operations involving heat exchange was evaluated, considering hydrous ethanol production on each case. Results are shown in Table 1, in which heat duty of thermally integrated streams, such as that of the equipments "HEAT-EXC" shown in Figures 2 and 3, is equal to 0. It can be inferred from Table 1 that changes in the distillation process configuration are needed when the alternative fermentation process is considered: the withdrawn of a vapour phase phlegm

on the conventional distillation column does not provide a significant reduction on energy consumption (4 %), as opposed to the conventional fermentation process (18 %).

Table 1. Energy consumption for each process.

	Energy consumption (kJ/kg hydrous bioethanol)					
Operation	CFCD	CFCDP	VFCD	VFCDP	VFDD	VFTD
Reactor cooling	1132	1121	0	0	0	0
Wine heating	3014	2985	0	0	0	0
Column A reboiler	6704	6197	1475	1156	0	0
Column D condenser	6016	4456	2651	2064	2125	2118
Liquid phlegm heating	0	0	19	39	498	0
Column B reboiler	4790	2784	3339	3438	3488	2714
Column B condenser	3669	2720	2346	2727	1304	0
Heating – total	14507	11966	4833	4633	3986	2714
Cooling - total	10817	8297	4997	4791	3429	2118

Since the alternative fermentation process operates under vacuum pressures, the "light" stream must be compressed before being fed to the distillation columns. In the conventional distillation configuration, at least three compressors must be used to raise stream pressure; in the double and triple effect, only one compressor may be used, since the pressure difference between the "light" stream and its feed stage is relatively small. Thus, electricity consumption in the distillation processes increases when the alternative fermentation process is considered: around 600 kJ/kg for the double and triple effect configurations, and around 1200 kJ/kg for the conventional distillation configurations.
Considering exclusively the consumption of thermal energy, the configuration that presents the lowest energy demand is the triple effect configuration (VFTD), providing a reduction in energy consumption equal to 44 % when compared to the VFCD process and 77 %, when compared to the CFCDP process, which is the configuration most commonly employed in Brazilian biorefineries. Nevertheless, the VFTD configuration presents the highest ethanol losses, along with the VFDD process, as shown in Table 2.

Table 2. Ethanol losses on each process.

Operation	CFCD	CFCDP	VFCD	VFCDP	VFDD	VFTD
Ethanol losses (kg/h)	3.6	3.3	1.8	1.7	5.1	5.1
Ethanol losses (%)	10.0	9.2	5.0	4.7	14.2	14.2

In order to reduce ethanol losses in the double and triple effect configuration (VFDD and VFTD), an absorber column may be employed to recover ethanol present in the "GASES" stream, since it accounts for more than 90 % of the losses.
Different stream flows are observed for each process configuration. Because of the increased wine concentration in the vacuum extractive fermentation processes, vinasse volumes are larger in the conventional fermentation process, and phlegmasse (stream "PHLEGMAS" in Figures 2 and 3) is also a significant residue generated in the distillation process. Values of flow for the main process streams are shown in Table 3.

Table 3. Stream flows for each process.

	Stream flows (kg/h)					
Stream	CFCD	CFCDP	VFCD	VFCDP	VFDD	VFTD
FEED	460.7	460.7	169.0	169.0	169.0	169.0
LIGHT	-	-	84.9	84.9	84.9	84.9
TO-DEST	460.7	460.7	69.6	69.6	69.6	69.6
GASES	37.1	37.1	12.4	12.4	17.0	17.0
ALCOHOL2	1.2	0.9	0.6	0.4	0.3	0.3
VINASSE	312.7	319.7	78.1	83.2	22.7	22.7
PHLEGMAS	74.8	67.8	26.5	21.3	81.1	81.0
ETHANOL	35.0	35.3	37.0	37.2	33.4	33.5

5. Conclusions

The combined simulation study of the fermentation and distillation processes for bioethanol production provides important resources that allow the comparison of different process configuration and their impacts on bioethanol losses, energy consumption and residue generation. In this work six process configurations were evaluated, based on conventional and vacuum extractive fermentative processes and conventional, double and triple effect distillation systems for product purification. The optimum configuration regarding energy consumption (the triple effect configuration, VFTD) requires 77% less thermal energy, when compared to conventional fermentation and distillation processes. However, this configuration presents the largest ethanol losses, which may be minimized by the addition of an absorber column, since one of the streams accounts for more than 90 % of the losses.

6. Acknowledgments

The authors acknowledge FAPESP and CNPq for financial support.

References

D.I.P. Atala, 2004, PhD Thesis, School of Food Engineering, State University of Campinas, 2004

M. Balat, H. Balat, C. Öz, 2008, Progress in bioethanol processing, Progress in Energy and Combustion Science, 34,551-573

S. Bui, X. Verykios, R. Mutharasan, 1985, In situ removal of ethanol from fermentation broths.1. Selective adsorption characteristics, Ind. Eng. Chem. Process Des. Dev, 24, 4, 1209-1213

G. Franceschin, A. Zamboni, F. Bezzo, A. Bertucco, 2008, Ethanol from corn: a technical and economical assessment based on different scenarios, Chemical engineering research and design, 86, 488-498

I. C. Macedo, J.E.A. Seabra and J. E. A. R. Silva, 2008, Green house gases emissions in the production and use of ethanol from sugarcane in Brazil: The 2005/2006 averages and a prediction for 2020, Biomass and Bioenergy, 32, 7, 582-595

F.L.H. Silva, M.I. Rodrigues, F. Maugeri, 1999, Dynamic modelling, simulation and optimization of an extractive continuous alcoholic fermentation process, J Chem Tech Biotech, 74,176-182.

G. Sobočan, P. Glavič, 2000, Optimization of ethanol fermentation process design, Applied Thermal Engineering, 20, 529-543

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

OPERATIONAL CONDITIONS IN OIL RECOVERY WITH BLANKET HEATING IN SHALLOW RESERVOIRS

Elthon J. R. Medeiros[a], Janusa S. Araújo[a], Tommy A. Pinto[a], Jennys L. M. Barillas[a], Tarcilio V. Dutra Jr[a,b], Wilson da Mata[a,c]

[a]*Programa de Pós-Graduação em Ciência e Engenharia de Petróleo, Universidade Federal do Rio Grande do Norte,Campus Universitário, Natal - RN 59078-970, Brasil*
[b]*Departamento de Engenharia Química, Universidade Federal do Rio Grande do Norte,Campus Universitário, Natal - RN 59078-970, Brasil*
[c]*Departamento de Engenharia Elétrica, Universidade Federal do Rio Grande do Norte,Campus Universitário, Natal - RN 59078-970, Brasil*

ABSTRACT

The most innovative proposals in the field of petroleum research refer to the oil recovery previously considered economically unviable in the market. Included in this premise is low °API oil, also known as heavy oil. Statistics show that, currently, the world reserves could be greatly improved by heavy oils contained in formations with depths between 50 m and 300 m, which are classified as shallow or ultra-shallow reservoirs. Thermal methods have been the most effective alternative for heavy oil recovery, and among them, there is the steam flood, the technique most often applied to reduce oil viscosity. A recently developed method, called Blanket Heating, combines some of the fundamental characteristics of thermal methods, adapting them to the particularities of the shallow reservoirs. This process works introducing steam in horizontal metal conduits, meaning that the heated fluid does not come into direct contact with the formation, working as a classical heat exchanger. The heating occurs indirectly, avoiding problems such as the recovery of large amounts of water and the insertion of excessive volume of steam, especially in cases where the depth is minimal and the overlying pressure is insufficient to contain the fluid. The primary focus of the article is to examine the influence of parameters involved in the operation of blanket heating, in order to find an optimal operational configuration. The choice of the horizontal direction for conduits reflects the need to maintain the greatest possible contact area between conduits and oil reserves. The results show that the blanket heating may be a viable process for heavy oil recovery in extremely shallow formations. They also show that oil recovery can be maximized in proportion to the increase of temperature in the conduits and the number of conduits. The cumulative oil production is reduced when the distance between the conduits and producer wells or between the own conduits is greater. In addition, the results were better when the completion interval of producer wells and the position of the conduits in relation to the vertical are arranged between the center and base of the reservoir. The inner diameter of the conduits was the only parameter that had minimal influence, showing no significant alterations in the production of oil. The study also showed that the blanket heating does not produce significant emissions of steam to the surface, confirming the reduction in the amount of water produced. In the face of the growing importance of heavy oil on the world market, together with the successful exploration of this resource, and knowing the current

applicability of the steam flood as the main alternative to the economic recovery of this type of oil, justify the need to a study that will expand the options to use this method in reservoirs that retain great amount of viscous oil at lower depths.

Keywords: Steam injection, heavy oil, shallow reservoirs, Blanket Heating, computer simulations.

1. INTRODUCTION

It is unlikely to apply the steamflood process successfully in shallow reservoirs, especially the ultra-shallow ones, since the overburden sediment does not provide sufficient barriers against the steam to ensure that the steam does not migrate to the surface, and it conditions high production of sand and the injected fluids. The steam exhaust would reduce the pressure needed to the reservoir, and also could take with itself the lighter fractions of hydrocarbons, causing environmental impacts. A possible solution to these problems is to use methods of indirect heating.

2. BLANKET HEATING

2.1. Method Description

In the Blanket heating process, the steam produced in a conventional generator flows through horizontal conduits to provide heat throughout the formation. The steam contained in conduits avoids the inconvenience of inserting large amounts of fluid directly into reservoir, and help to maintain the temperature through a injected steam control - and the pressure on the conduits. Producer wells are drilled in reservoir to establish the flow of fluids and can also be used to maintain pressure (Osterloh and Jones, 2001).

The conduits must return to the surface so that steam can be reheated in boilers. An alternative would be the steam coming back internally through the conduits. In this case, an external compartment, call annulare, is used to lead the high-quality steam until tubing end, transmitting heat by conduction to the formation, and an ICCT (Insulated Concentric Coiled Tubing), located internally to the annulare, is used so that the fluid return to the boilers.

2.2. Mechanisms Process

The process involves conducting heat - from steam to oil through conduits. In this sense, an alternative material for the conduits would be one able to improve heat conduction, and resistant enough to withstand the pressure and other factors. Once forwarded to the oil, the heat spreads through hydrocarbon by convection and also by conduction (to a lesser extent). Oil in contact with conduits suffer reduction in density and tends to rise, mobilizing the oil located above, inducing vertical flow and easing the thermal energy transfer and also the wells production. The conduits are placed below the oil, to condition the convection process. The process is also privileged by high permeability (vertical and horizontal) of reservoir shallows.

Besides heat transfer, the success of the process is due to the existence of additional process of vaporization or in situ condensation. Steam formed around conduits rises up passing through sediments, and accelerating the heat transfer to lower temperature zones. When steam is finally condensed, liquid will be drained down, filling the warm zone around conduits, where it will be vaporized again (Osterloh and Jones, 2001).

3. MODELING DESCRIPTION

The selected system uses Cartesian coordinates in "i", "j" and "k", directions with dimensions of 60 m x 60 m x 32 m, totalizing 4,440 blocks (Fig. 1 A). The conduits were laid out along the axis "j", and correspond to the horizontal columns of blocks with radial refinement, and diameter equal to 5 cm. Only blocks representing conduits were heated at different temperatures in order to evaluate the capability of this configuration to transfer heat to the reservoir in different situations. Fig. 1 B is a top view of the radial refinement applied to conduits blocks.

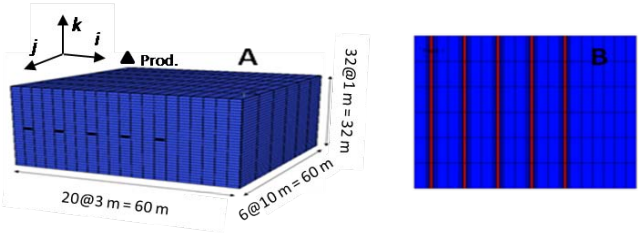


Figure 1.A. 3D representation of base case. B. Conduits (flat “i” -“j”).

3.1. Rock-Reservoir Parameters

Table 1 shows reservoir properties, which were based on real reservoirs data and some literature data.

Table 1. Reservoir Properties; Rock Properties.

Reservoir		Rock	
Gas cap - (m)	2	Effective Compressibility of Rock (1/Pa)	$4{,}4 \cdot 10^{-7}$
length - i (m)	60	Thermal conductivity of rock (J/m.s.°C)	1,73
Width - j (m)	60	Thermal conductivity of water (J/m.s.°C)	0,61
Thickness - k (m)	32	Thermal conductivity of oil (J/m.s.°C)	0,13
Reservoir depth (m)	50	Thermal conductivity of gas (J/m.s.°C)	0,04
Initial water saturation (%)	30	Horizontal permeability - Kh (mD)	1000
Initial oil saturation (%)	70	Vertical permeability - Kv (mD)	100
Initial reservoir temperature (°C)	37,8	Porosity - φ (%)	28
Original oil in place (m^3std)	21509		

Operational parameters used in studied model are shown in Tab. 2. Legend used was: "B", "C" and "T" represents completion and location of conduits, which represent, "Base", "Center" and "Top", respectively. First letter located to the left refer to completion, while the letter to the right is the position of the conduits in relation to the direction "k" (vertical).

Table 2. Levels of Operating Parameters.

Parameter	Levels									
Completion Interval of the well producer and the location of vertical conduits	BC	CB	TT	TB	TC	CC	BB	BT	CT	All B
Distance between conduits (m)	6	9	12	-	-	-	-	-	-	-
Distance between prod. and conduit (m)	3	6	9	-	-	-	-	-	-	-
Quantity of Conduits	1	2	3	4	5	6	7	-	-	-
Radius of the Conduits (cm)	2,5	3,75	5	-	-	-	-	-	-	-
Temperature in the Conduits (°F)	150	200	250	300	350	400	450	500	550	600

4. RESULTS AND DISCUSSION

4.1. Analysis of Conduits Radius

In this section, it was performed an analysis of conduits radius influence on cumulative oil. All parameters shoed in Tab. 2 were considered constants, except the conduits radius, which was studied at three levels (2.5, 3.75 and 5 cm). Fig. 4 shows cumulative oil, cumulative water and cumulative gas for all three radius studied.

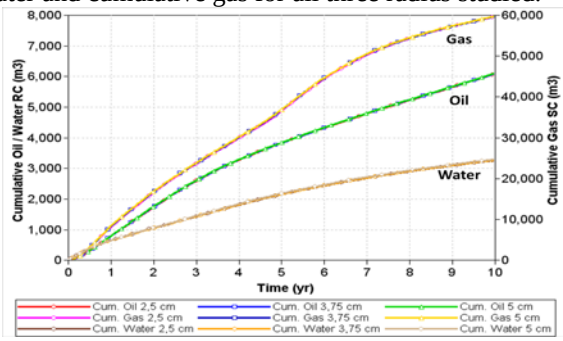


Figure 2. Cumulative Oil Production *versus* Time. Comparative of Radius of the Conduits.

The same behavior can be observed for the three levels as the internal radius does not significantly alter the area of contact of the conduits with the heated oil, changing minimally the heat transfer. Tab. 3 represents a summary of the models with different conduit radius. Oil recovery is given by relation between cumulative oil (CO) and original oil in place volume (OOIP), so OR = CO / OOIP. It can be observed that there was no considerable change in the oil recovery as discussed in the graphics of cumulative production.

Table 3. Oil Recovery e Cumulative Oil at Ended of Production for all Three Conduits Radius.

Radio	CO (m^3)	OR (%)
2,50 cm	6066,51	28,20
3,75 cm	6077,56	28,26
5,00 cm	6090,19	28,31

4.2. Analysis of Conduits Temperature

Fig. 3 shows oil recovery (OR) *versus* temperature, for annual production times. Oil recovery increase with time and temperature. A growth of the recovery factor with the rising of the temperature occurs until the fifth year of project, and after that time there is almost no change in the recovery factor at temperatures above 450 °F (232.22 °C). In the eighth year, it was observed that the highest recovery factor occurs to the temperatures between 400 and 500 °F (204.44 and 260 °C).

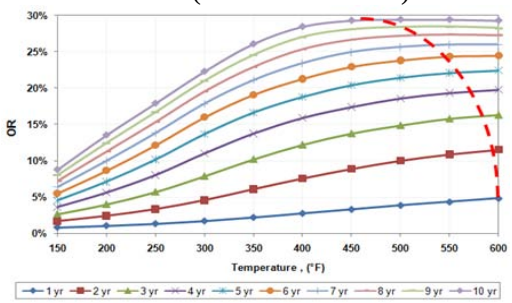


Figure 3. Graph of the Oil Recovery at Each Temperature According to Time.

The red dashed line represents approximately the temperature at which the higher recovery is obtained (qualitative analysis).

4.3. Analysis of Conduits Numbers

It was held in this section a variation in the amount of heating conduits. The analysis was developed using one to six conduits, with fixed distances between them and between the producer well, to check the influence of tubes number on oil recovery. In Tab. 4 there is a growing pattern every years, that is, with more heated conduits, greater cumulated oil production is and it there was no change during the time it was studied.

Table 4. Comparison of Conduits Numbers.

Quantity of Conduits	CO (m3)	OR (%)	Quantity of Conduits	CP (m3)	OR (%)
1	3763	17,50	4	5752	26,74
2	4577	21,28	5	6302	29,30
3	5209	24,22	6	6878	31,98

4.4. Analysis of Distance Between Conduits

In Fig. 4 may be noted that great distances between the heated tubes promote greater cumulative oil due to better distribution of heat inside reservoir. Smaller distances result initially in greater oil rate. In other words, better results are obtained with smaller distances and bigger number of conduits.

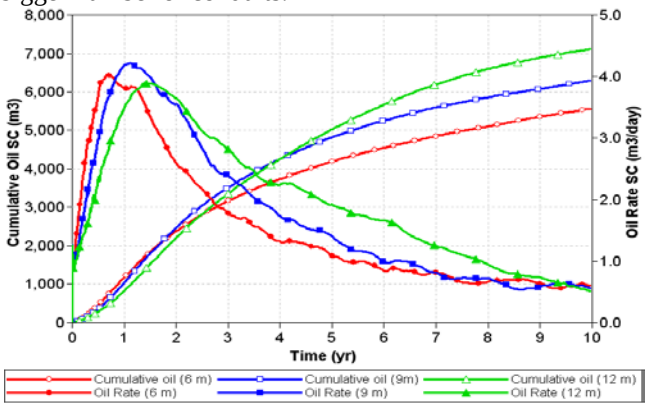


Figure 4. Cumulative Oil and Oil Rat *versus* Time. Distance between Conduits.

4.5. Analysis of Distance between producer and conduit

The closer the conduits are to the producer well the higher is the cumulative oil, resulting on oil viscosity reduction facilitating their mobility, as shown in Tab. 5.

Table 5. Comparison of Distance between Conduits and Producer Well.

Distance between prod. and conduit	CO (m^3)	OR (%)
3 m	6468,86	30,01
6 m	6357,67	29,56
9 m	6066,51	28,20

4.6. Analysis of Completion Interval of the Producer Well and Location of Vertical Conduits

It can be concluded that the best combinations between the completion interval and location of conduits vary over time, between the center and reservoir base. The top had not, at any time considered, a good combination, due to heat loss to the overburden layers of the reservoir, with no significant heating in reservoir bottom (fluids density), and also, less influence of the convective process, because heating occurs on reservoir. The best settings were observed in cases 1, 6 and 7, respectively, producing and heating

at the base (BB), producing at the base and heating the center (BC) and producing and heating in the center (CC) (see Fig 5).

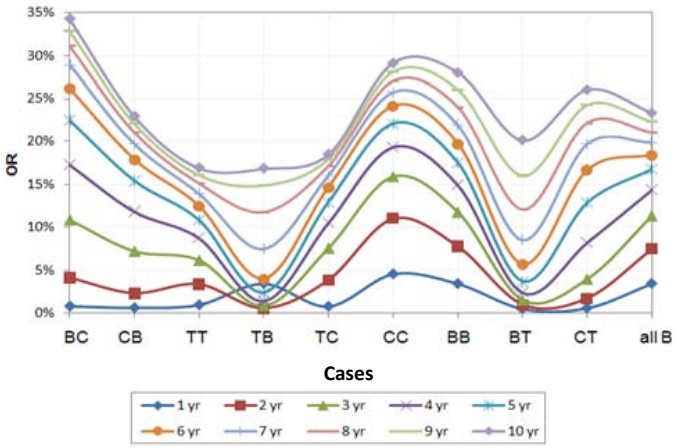


Figure 5. Graph of the Oil Recovery in Each Case in Accordance with the Period.

5. CONCLUSIONS

During the time of study there was increase in the recovery factor with the increasing of temperature in the conduits. The more conduits were heated at high temperatures, the more oil recoveries are obtained.

For the parameters of completion interval and location of conduits, results ruled out the setting at the top due to irrelevant recoveries. The best settings: the base of the reservoir for both parameters (BB), or the base for the completion, with heat in the center (BC).

Oil recovery increases with the reduction of the distance between producer well and conduits, and when the distance between the own conduits is greater, due to better distribution of heated area in the reservoir.

ACKNOWLEDGEMENTS

The authors want to thank PETROBRAS and Laboratório de Estudos Avançados em Petróleo (LEAP-UFRN) for the support received in the execution of this work.

REFERENCES

1. STARS User's Guide, Advanced process and thermal reservoir simulator, Computer Modelling Group Ltd, 2007.
2. Osterloh, W.T.; Jones, J. Novel thermal process of recovery of extremely shallow heavy oil. SPE 83733, p. 127-134. Março, 2003.
3.Osterloh, W.T.; Jones, J. Process for recovery of extremely shallow heavy oil. In: SPE Western Regional Meeting, SPE 68809, Bakersfield, California, EUA, Março, 2001.
4. Medeiros, E. J. R. Estudo do Aquecimento Geral Indireto como método de recuperação em reservatórios rasos de óleo pesado, Master Thesis, Universidade Federal do Rio Grande do Norte, 226 p., Natal, 2008. (In Portuguese).
5. Dunn-Norman, S.; Gupta, A.; Summers, D. A.; Koederitz, L. F.; Numbere, D. T. Recovery methods for heavy oil in ultra-shallow reservoirs. In: SPE Western Regional/AAPG Pacific Section Joint Meeting, SPE 76710, p. 1-4. Alaska, USA, 20-22, Maio, 2002.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A New computational Tool for Falling Film Molecular Distillation Performance Prediction

Evandro S. Mallmann[a], Caliane B. B. Costa[a], Maria R. W. Maciel[a] and Rubens Maciel Filho[a]

[a]*School of Chemica Engineering, University of Campinas, UNICAMP, P.O.Box 6066, 13083-970, Campinas – SP, Brazil, engqui.butia@gmail.com, maciel@feq.unicamp.br*

Abstract

In a previous work, a software named DISMOL, was developed in order to simulate molecular distillation. However, due to the restricted access to this tool and the easiness to make use of data bank and physical properties, in the present work, it is proposed the development of general procedure for this highly specific process in the commercial simulator Aspen Plus®. Since no single unit operation is available in the commercial simulator that can appropriately simulate a falling film molecular distillator (MD), the proposal, in a preliminary approach makes use of a sequence of flash vessels in order to accomplish the task of MD simulation. Experimental data of a binary system distillated in a MD are used for the validation of the developed tool. The results indicate the high potential of this tool in representing a MD equipment, making possible the evaluation of different operational policies in conducting this high vacuum distillation to achieve products with the desired quality and properties. The development of the proposed computational tool is an important step for the investigations of operational policies for MD application in oil refining since no industrial data of such equipment is available.

Keywords: Process Model, Molecular Distillation, Aspen, Model Validation

1. Introduction

The petroleum obtained from Brazilian reservoirs is of heavy and ultra-heavy types, presenting high viscosities that make difficult its exploration (oil removal from submarine reservoirs) as well as the refining operations for fuel and other derivatives production. The refining processes currently in activity in the Brazilian refineries do not allow the heavy and ultra-heavy oil processing, so that they are mixed with light oils (of the Arab type) for their processing. In fact this is usually the general strategy adopted worldwide to deal with such type of crude oil. This procedure leads to the dependence of Brazil on oil importation, as well as to the not complete valuation of the oil found in the country. In this context, an alternative to the initial heavy and ultra-heavy oil operation is the use of molecular distillation, a process conducted at high vacuum. Molecular distillation technology, also known as short-path distillation, is an operation for effective separation or purification of products of high molecular weight. It is a special vaporization operation at very low pressures and, consequently, relatively low temperatures, which make this operation also suitable for thermally sensitive products [1], and in particular for the oil refining, avoiding cracking and undesired reactions. Besides that, no external component needs to be introduced into the system in order to perform the operation that can be an advantage for cost reduction in further separation units and in the quality of final product.

Previous experiments [2, 3-4] showed the viability of the MD operation also for heavy and ultra-heavy oil refining. Nevertheless, a deep knowledge of the process is required in order to obtain the desired product streams: small process variations may result in product streams with completely different characteristics [5]. A wide investigation for the development of operational polices that lead to high performance of the molecular distillation (MD) operation for oil refining is necessary. In this context, it is interesting to develop a simulation tool based on commercial simulator since components data bank and facilities to integrate with other unit operation are available.

Bearing all these in mind, the main objective of this work is the development of a simulator of a falling film MD piece of equipment in Aspen Plus® for heavy and ultra-heavy oil processing. The developed tool turns possible an evaluation of the process variables impact on process behavior and characteristics of obtained product streams (residue and distillate). As Aspen Plus ® does not include a MD piece of equipment, it is necessary to choose, among the available unit operations, the one that better suits to the MD parameters. A sequence of flash vessels coupled with process efficiency is selected to represent the process and experimental data are used to validate the model.

2. Methodology, Simulator Development and Validation

In this investigation, data from the binary system Dibutyl phthalate – DBP (278.35 g/mol) and Dibutyl sebacate – DBS (314.14 g/mol) were used for MD process simulation [5]. This binary system was chosen due to its simplicity when compared to the petroleum (multicomponent system), to its high molecular weight and to the availability of published experimental data. Since Aspen Plus® does not possess a MD operation tool, a sequence of flash vessels, operating at equilibrium, were proposed to emulate the molecular distillation. As MD process is ruled by mass transfer rates [6] (non-equilibrium) [7], efficiency considerations must be done in the developed tool in order to fit the simulator outputs to the experimental data. The idea to use a sequence of flashes corrected with process efficiency is a preliminary approach since relatively well established mass-transfer models could be used into Aspen environment to develop the simulator. However, the proposed approach has the advantage to be simpler and be dependent upon only the efficiency values rather than the uncertainties of the mass transfer models. Additionally, to use mass-transfer approach is necessary to identify which could be the more suitable one. This is an important issue to develop the whole simulation methodology and, once it has been considered suitable, the next step should be the use of rate-based approach.

The parameters used for the efficiency and simulator validation were distillation mass ratio (distillate mass/inlet mass) and the molar compositions of distillate and residue, since such data are ready available. The operation temperature was the analyzed parameter, since it is the limiting factor for the great majority of MD applications [8]. These parameters were chosen because independently of the equipment model they are always available and with good accuracy.

2.1. Simulator Development

In the simulations performed, 50 kg/h of an equimolar mixture were fed to the system, operated at 0.001 mmHg. For the MD operation, this process conditions lead to temperature operation of 95.85°C and a distillation mass ratio of 21.2% (i.e., 21.2% of the mass fed to the falling film molecular distillation unit is obtained as distillate), with molar fraction of DBP equal to 0.775 in distillate and 0.429 in residue streams (DISMOL results). The MD simulation first built in Aspen Plus® is shown in Fig. 1, in which, three flash vessels are connected.

The condensed vapor streams obtained with the simulated process depicted in Fig. 1 have, together, a mass rate greater than that obtained with DISMOL simulation. Also, the molar fraction of DBP in "Distill" stream (Fig. 1) is higher than the real MD one (simulated with DISMOL).

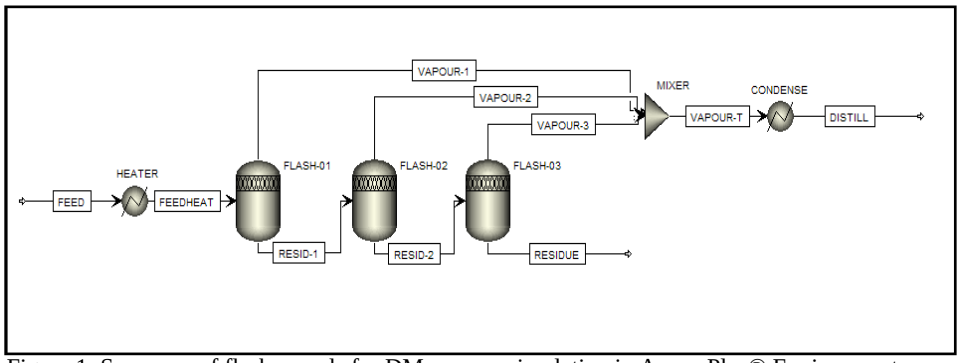


Figure 1: Sequence of flash vessels for DM process simulation in Aspen Plus® Environment.

Furthermore, the operation temperatures of each flash vessel (61.5 °C, 63.0 °C and 64 °C for FLASH-01, FLASH-02 and FLASH-03 respectively) were lower than the real MD temperature for the defined operation pressure. These results are presented in Table 1. The differences in simulation results are due to the fundamental differences with the two unit operations: while a MD process is governed by mass transfer limitations, an ideal flash vessel, as the ones illustrated in Fig. 1, is an equilibrium stage and with an efficiency of 100%.

Table 1. Simulation results for DISMOL and the developed Aspen Plus® tool.

	DISMOL	ASPEN PLUS®	
Distillation mass ratio	0.2120	0.2854	
Distillate DBP molar fraction	0.7750	0.8700	
Residue DBP molar fraction	0.4290	0.3420	
Operation temperature (°C)	95.85	FLASH-01	61.50
		FLASH-02	63.00
		FLASH-03	64.00

The flash sequence leads to a better separation of components since it is an equilibrium operation. However, since MD process is a mass transfer ruled process, it is expected that just one flash vessel may be used (since the more number of flash vessels, the greater the difference between simulation results from non-equilibrium processes results) coupled to an adjusting equation, used to translate flash simulation results to MD process results. Therefore, the final MD tool built in Aspen Plus® possesses just one flash vessel and an adjustment equation to take into consideration that a rate-based process is being emulated by an equilibrium approach.

2.2. Adjusment Tool

Using just one flash vessel, with the selected operation pressure, operation temperature was manipulated in order to evaluate in which operation temperature a distillation mass ratio equals to the one of a real MD process can be obtained. Figs. 2 and 3 show the vapor (which emulates MD distillate) mole fractions and mass flows both for DBP and DBS as a function of temperature. Fig 2 represents the mass flow of the DBP and DBS

components on the vapor stream as a function of flash temperature, while Fig. 3 depicts the molar fraction variation of DBP and DBS on the vapor stream when the flash temperature is changed. From these Figures it is possible to observe that the better separation of system components occurs when flash temperature is 61.5° C. If the flash temperature gets higher, the DBS molar fraction increases and the DBP molar fraction decreases, lowering the vapor purity grade. The condenser turns into liquid phase all the vapor stream coming from the flash vessel. However, in order to obtain the same distillation mass ratio of the DISMOL (21.2 %), the flash equipment has to operate at temperature of 63.1° C. This temperature, nevertheless, results in a higher molar fraction of DBP in the distillate (0.875) and a lower molar fraction of DBP in the residue (0.329).

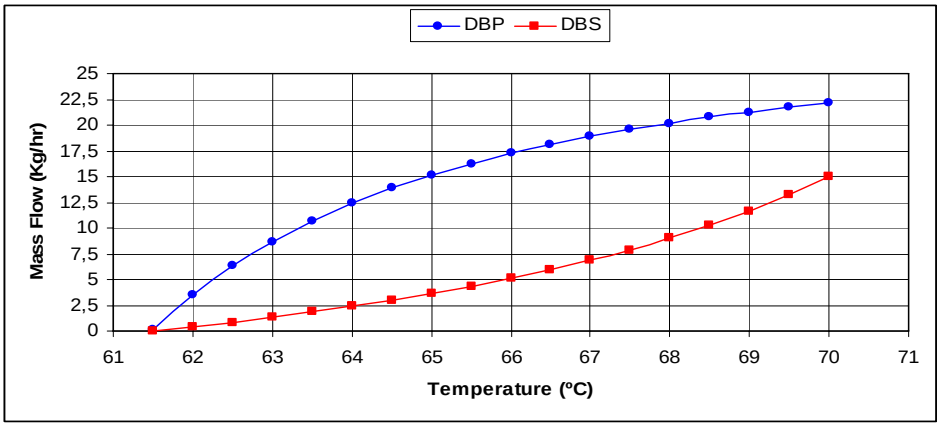


Figure 2. Mass flow of distilled stream as a function of flash vessel operation temperature.

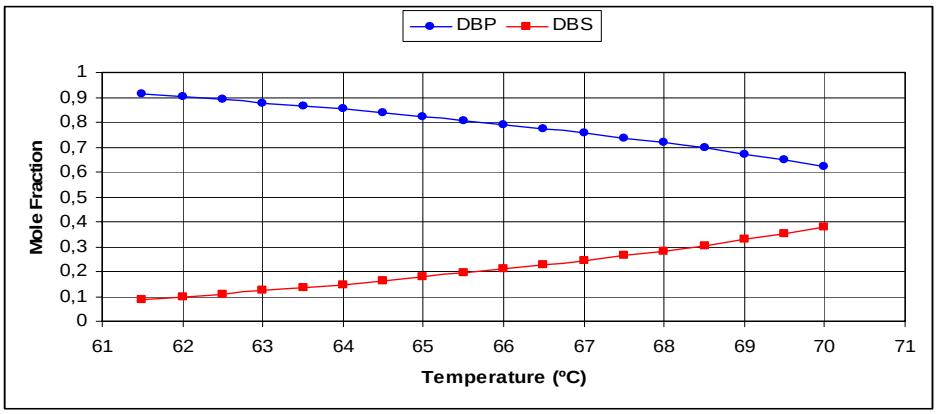


Figure 3. Mole fractions of distilled stream as a function of flash vessel operation temperature.

An adjusting equation, which quantifies an efficiency measurement, is then formulated in order to convert flash results into MD ones, that means to correct the fact that a rate-based process is represent by an equilibrium operation unit. This equation correlates the MD unit operation temperature (given by DISMOL) to the flash operation temperature, so as to provide the same emergent flows composition. The efficiency is calculated, then, dividing flash by DISMOL operation absolute temperatures (Kelvin scale), as depicted in Eq. (1).

$$\eta = \frac{T_{flash}(K)}{T_{DISMOL}(K)} = \frac{336{,}26}{369} = 0{,}91 = 91\% \quad (1)$$

The adjusted molar fraction is then calculated multiplying the efficiency factor 'η' to the DBP molar fraction of flash exit vapor stream. The new (corrected) DBP molar fraction in residue is found by mass balance. Table 2 brings DISMOL results compared to the results generated by the developed tool (one flash vessel coupled to an efficiency factor calculated by Eq. (1). In principle, the proposed procedure is general and may be used to any system including multicomponent ones.

Table 2. DISMOL and corrected simulated data (by Eq. 1) for the binary mixture MD.

	DISMOL	ASPEN PLUS®
Distillation mass ratio	0.2120	0.2120
Distillate DBP molar fraction	0.7750	0.7900
Residue DBP molar fraction	0.4290	0.4090

2.3. Heavy Petroleum Simulation

For the evaluation of the heavy oil behavior (multicomponent) in the flash system, the experimental data reported by Santos [9] was used, through the MD. The oil used is called 'gama' petroleum (fantasy name), proceeding from residue at 420ºC+. From the distillate and residue percent data produced in MD, the simulation of the flash system could be carried out, determining different operation temperatures for one system to another, as shown in the Table 3.

Table 3. Comparison between experimental data (MD) and the simulation data (Flash).

Santos (MD)			Aspen Plus® (Flash)		
Temp.(ºC)	Distill. (%)	Resid. (%)	Temp. (ºC)	Distill. (%)	Resid. (%)
80	27.16	72.84	101	27.30	72.70
140	30.15	69.85	106	30.10	69.90
200	34.94	65.06	115	34.90	65.10
260	47.14	52.86	139	47.22	52.78
340	63.30	36.70	171	63.49	36.51

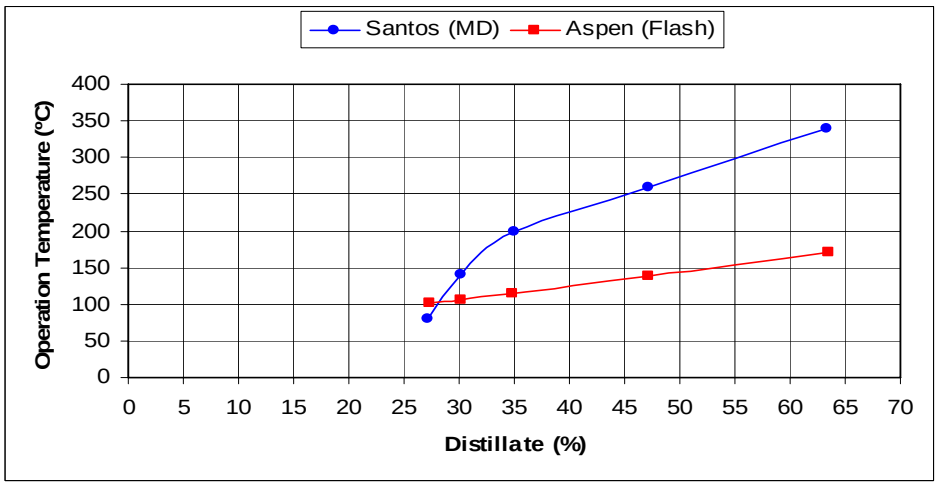


Figure 4. Behavior of the distillate and the residue percent as function of operation temperatures for MD and flash.

From the experimental values, studies and simulations was made for the behavior evaluation of the flash system operation temperature as function of the simulator response variables (distillate and residue percent), similar to the experimental percents.
The difference in the behavior of the curves showed in the Figure 4 is because the very short residence time in the flash system, which needs elevated temperatures to promote the lighter compounds separation. However, the flash needs a variation in the operation temperature lower than MD to produce practically equal distillation indexes.

3. Conclusions

A new tool was developed with the software Aspen Plus® available unit operations to represent MD. As a preliminary approach it is proposed the use of flash operation corrected with process efficiency since MD is a mass transfer limited process. This tool simulates with a good prediction capability literature data for a binary system. The best tool structure was found with the use of one flash vessel coupled to an efficiency factor. The results generated with the developed tool indicated it can be applied to more complex systems, like heavy and ultra-heavy petroleum.

4. Acknowledgements

The authors are grateful to FAPESP (2007/06833-3 and 2006/55177-9) for the financial support.

References

[1] A.T. Erciyes, H. Ishikawa, M. Inuzuka, S. Hiraoka, H. Mori and I. Yamada, I. Chem. E. Symposium Series, 1 (1987) A359
[2] J. A. B. Hernández, L. Z. Linan, A. Jardin, M. R. W. Maciel, R. M. Filho and L. C. M. Oliveira, Rio Oil & Gás Expo and Conference, Rio de Janeiro, Brazil, IBP2712_08 (2008)
[3] E. R. L. Rocha, M. R. W. Maciel, R. M. Filho, C. B. Batistella and L. C. M. Oliveira, Rio Oil & Gás Expo and Conference, Rio de Janeiro, Brazil, IBP2724_08 (2008)
[4] R. S. Rocha, M. R. W. Maciel, R. M. Filho, C. B. Batistella and L. C. M. Oliveira, Rio Oil & Gás Expo and Conference, Rio de Janeiro, Brazil, IBP2730_08 (2008)
[5] C. B. Batistella and M. R. W. Maciel, Computers Chem. Engng., 20 (1996) S19.
[6] J. Lutisan, J. Cvengros, M. Micoslav, Chemical Engeneering Journal, 85 (2002) 225
[7] K.C.D. Hickman, A Review. Chem. Rev., 34 (1943) 51
[8] J. Lutisan, J. Cvengros, M. Micoslav, Chemical Engeneering Journal, 78 (2000) 61
[9] P.S.D. Santos, *Sc. D. These*, Laboratory of Separation Process Development, School of Chemical Engeneering, State University of Campinas, São Paulo, Brasil, 2005

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Morphological Population Balance Models for the Dynamic Evolution of Particle Shape and Size Distribution in Protein Crystallization

Jing J. Liu[1,2], Cai Y. Ma[1], Yang D. Hu[2] and Xue Z. Wang[1]*

[1]*Institute of Particle Science and Engineering, University of Leeds, Leeds LS2 9JT, UK*
[2]*School of Chemical Engineering, Ocean University, Qingdao, PR China, 266001*

Abstract

Protein crystallization is known to be affected by many factors and inherently difficult to control. Being able to model the crystal growth, especially at process scale for the population of particles in a reactor rather than for a single particle, will no doubt greatly help the formulation and manufacture of protein crystals. In this paper, a morphological population balance model is presented which has incorporated the crystal shape information into the population balance process model therefore is able to simultaneously simulate the dynamic evolution of shape as well as size for crystals of tetragonal Hen-Egg-White (HEW) lysozyme within a crystallizer. Morphological population balance models require growth kinetics data for each facet, which was obtained from published data in literature for the two identified independent crystallographic faces, {101} and {110}, of HEW lysozyme.

Keywords: morphological population balance model; protein crystallization; HEW lysozyme, crystal shape control

1. Introduction

Protein crystals have shown significant benefits in the delivery of biopharmaceuticals, but the majority of biopharmaceuticals are not yet marketed in crystalline form with only a few exceptions such as insulin. This situation could be partly due to the fact that proteins are difficult to crystallize and the crystallization process is inherently complicated and difficult to control. Proteins, like many other molecules, can be stimulated to form crystals when placed in the appropriate environment. But protein molecules are large, complex and the nutrient solutions for growing protein crystals differ considerably from the simpler nutrient solutions for small molecule crystal growth. The majority of previous work focused on investigations on crystal growth at the molecular and single crystal scales. For the knowledge to be more useful in the design, optimization and control of industrial protein crystallizers, modeling at the process scale to understand the growth behavior of population of the crystals is clearly important.

In modeling crystallization processes, PB provides the dynamic evolution of crystal size distribution in all stages including nucleation and growth, which is also a powerful tool for studying the effects of various operating conditions such as impurity, solvents, cooling rates and supersaturation. As a result, it is considered as a useful tool for product development, process design, optimization and control. Traditionally, crystallization PB modeling has been conducted using a mono-size dimension by defining the size of a particle as the diameter of a sphere having the same volume of the

particle. Such a simplified treatment obviously misses important information about the evolutionary behavior for the morphology of crystals as a population. Motivated by this observation, some researchers have attempted to develop multi-size dimensional PB models for crystallization processes[1,2,3]. However, the only work that can be found in literature on PB modeling of protein crystallization processes [4] did not take into account of the morphology of protein crystals.

Population balance models, in particular multi-dimensional and morphological PB models for protein crystallization processes, could greatly help the understanding of the growth behaviour for the entire population of crystals in the reactor in terms of morphology as well as size distribution. Such knowledge is vital for the successful formulation and manufacture of protein crystals in industrial scale crystallisers. As a first attempt to build morphological population models for protein crystallization processes, we have chosen a protein that has been well studied in literature on the crystal growth behaviour at single crystal scale, the enzyme hen egg white lysozyme.

2. Lysozyme crystallization

Enzyme Hen Egg White (HEW) lysozyme is an enzyme that crystallises in the tetragonal structure with macromolecules in Figure 1. The morphology of HEW lysozyme is dominated by rhomb-octahedron {101} and hexagon-tetrahedron {110} faces (right). Crystal forms are manifested in the external morphology through the multiplicities of tetragonal symmetry with twelve crystal growth surfaces in total, and determined by the relative growth rates of the different symmetry-related faces involved.

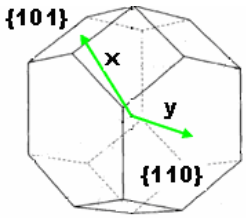


Figure 1. Structure of the tetragonal lysozyme crystal

2.1. Solubility of tetragonal HEW Lysozyme crystals

Solubility is always needed in order to determine the supersaturation at any given operational condition, which is usually used to represent the driving force for crystallisation from solution. Protein solubility can be described as a function of three variables, including pH, precipitant concentrations and temperature. Although efforts were made on measuring solubility, none of the models and techniques is exactly for the operational condition of the current simulation. Therefore in this study we correlated a new solubility equation specifically for the simulated operational condition.

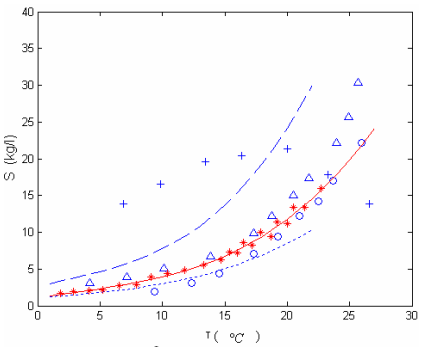


Figure 2. HEW lysozyme solubility

Solubility data plotted for protein HEW lysozyme that was obtained at pH value of 4.5, and different precipitant concentrations of 2.5% from literature, plotted as 'Δ'; 2.5% [5], plotted as 'O'; 2.5% [6], '+'; and 2.0% [7] and 3.0% [8], using blue dashed and dotted curves. Since the simulation condition is pH=4.5 and NaCI concentration of 2.5%, the data of [7] and [8] at 2.0% and 3.0% precipitant concentrations were used to generate data at 2.5% precipitant concentration, plotted as '*'.

It needs to point out that the disparity of the data does not necessarily mean the data is not reliable, because protein crystallisation data is known to have poor repeatability, and minor variation of experiments could give different values. Nevertheless, given the

good agreement between the other three sources, one set of the data as shown in Figure 2 as '*' was used to derive the solubility equation with a third-order polynomial:

$$S = 0.0012096 \times T^3 - 0.010496 \times T^2 + 0.26159 \times T + 1.1408 \qquad (1)$$

where S is the solubility of HEW lysozyme crystal, kg/l, and T is temperature, ^oC. The equation is plotted in Figure 2 as a red solid curve.

2.2. Facet growth rate of tetragonal HEW Lysozyme

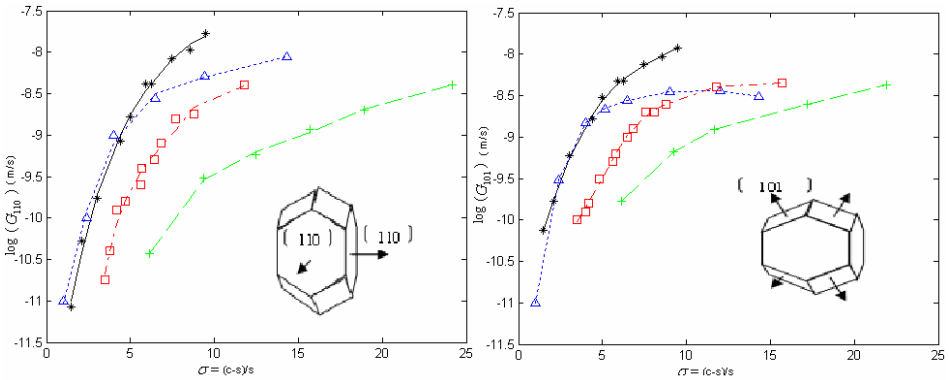


Figure 3. Growth rates of {110} (a), and {101} (b) faces. *– measured at 24 , pH=4.6, and NaCI concentration 3.5%; □– measured at pH=4.5, NaCI concentration 2.5%, and temperature ranging from 12 to 270C; □ – measured at pH=4.5, NaCI concentration 2.5% and temeperature 22 +- obtained under conditions of pH=4.4, NaCI concentration 2.5% and temperature 12.5

There have been experimental studies in literature on growth rates of individual faces of HEW lysozyme crystals. Some of the data is plotted in Figure 3. Although there are some disparities among the data, considering the fact that experiments were conducted in different labs and using different techniques, the growth rate data in Figure 3 is considered comparable and acceptable. The following growth kinetic equations were used to describe the growth rate for each facet:

$$G = k\sigma^n \qquad (2) \qquad\qquad \sigma = (c - s)/s \qquad (3)$$

where G is the facet growth rate m/s, σ relative supersaturation, C the solute concentration *kg/l*, S solubility *kg/l*, k and n are kinetic parameters.

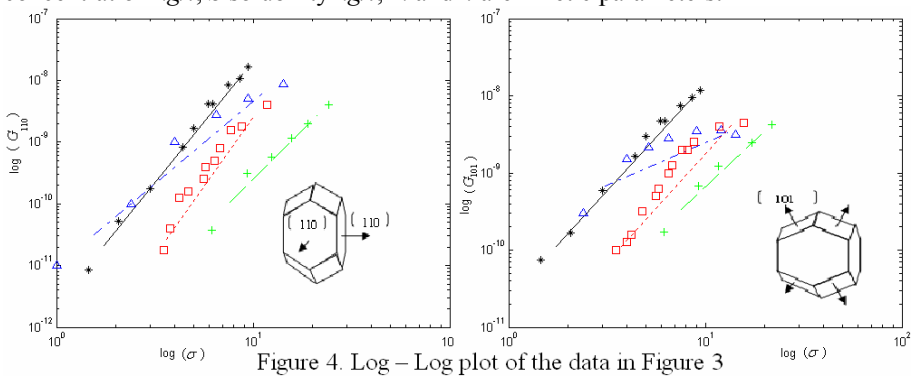


Figure 4. Log – Log plot of the data in Figure 3

Based on the experimental data about facet growth rate, the parameters k and n can be estimated using least square fitting in the log-log scale, as illustrated in Figure 4. Since

n is equal to the slope of the fitting, k can be calculated by the intercept of fitting line with Y axis. It was found that the value of exponent n calculated using the data with dark asterisk of Figure 3 is consistent with the value published [9]. This agreement is proof that the current method for determining the values of the parameters k and n.

Although values for both k and n are estimated for all the data plotted, the only data that is similar to our simulation was collected from [10]. Therefore k and n values obtained from this data were used in the simulation. The k and n values for the two independent facts are $k_{101} = 5.1\times10^{-12}$ m/s, $n_{101} = 2.7$ and $k_{110} = 5.9\times10^{-13}$ m/s, $n_{110} = 3.9$.

2.3. Morphology PB equation for Lysozyme crystallization

As stated earlier, due to the symmetry-related faces, the twelve faces of tetragonal lysozyme crystal can be modelled using two independent faces. Therefore the particle shape can be characterized by two internal length coordinates x, y (Figure 1), which are the distances from the geometric crystal centre to the identified independent faces.

In the current case study, a seeded system with one operating condition (fixed buffer, pH value and salt concentration) is considered and supersaturation is created by reducing the temperature at a constant cooling rate. As is known, in the presence of seed crystals, nucleation can be suppressed not only by different mean seed size and various seed loadings [11], but by keeping the supersaturation within the boundaries where crystal growth occurs, but primary and secondary nucleation do not. So it is appropriate to assume that the amount of solute leaving the solution must be accounted for by crystal growth without nucleated crystals. Under these assumptions, the mathematical formulation for two-dimensional morphological PB modelling can be given by Eq (4), where f is crystal size distribution, x, y referring to each characteristic length scale, G_{101} and G_{110} refers to the growth rate for the length scale, i.e. the growth rate of face {101} and face {110}. The right-hand side stands for the crystal creation mechanisms, indicating zero new crystals being introduced into this system. This is a simplified treatment, in future research, the net rate of new crystals should be considered.

$$\frac{\partial f(x,y,t)}{\partial t}+\frac{\partial[G_{101}(x,t)f(x,y,t)]}{\partial x}+\frac{\partial[G_{110}(y,t)f(x,y,t)]}{\partial y}=0 \qquad (4)$$

2.4. Solution algorithm

Solution algorithms for PB equations have been extensively researched in literature. However, it is well-known that the standard first-order schemes give diffusive solutions while the commonly used second-order schemes give spurious oscillations. So it is of great challenge to simulate these equations with extremely sharp distribution of crystal size (1μm to 500 μm), a wide range of length scales and time scale (20 s to 20000 min). High-resolution algorithms have been specifically developed to provide high accuracy while avoiding the numerical diffusion and numerical dispersion associated with other finite difference and finite volume methods.

The high resolution simulation algorithm has proved to be able to provide short computation time and high accuracy in literature. Therefore it is also used in this study. The two-dimensional population balance distributions were obtained via solving the corresponding equations with the number of size classes being 500×400 and the corresponding mesh size being 1×1 μm within a desktop Dell PC (Pentium 4 2.3GHz processor and 3Gb Ram). No spurious oscillations or numerical diffusion occurred, through simulations of HEW lysozyme crystal growth. The numerical stability of the algorithm is assessed using the Courant-Friedrichs-Lewy condition. Since details of the algorithm can be found in literature[12], they will not be repeated here.

3. Results and Discussion

Linear cooling mode is chosen to produce supersaturation in this study and the starting mean normal distances from crystal centre to the face {101} and {110} have been specified as 50 and 15 μm in *x*, *y* directions. It can be seen in Figure 5 that the shape of population distribution is still kept Gaussian-type with the full width at half maximum almost unchanged during protein crystallization process. At the same time, the mean normal distance to the {110} has increased from 50 to 490 μm, and the mean normal distance to face {101} has changed from 20 to 376 μm.

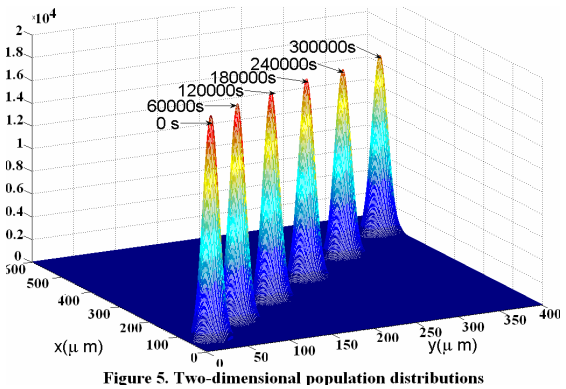


Figure 5. Two-dimensional population distributions

The growth rates for the two individual faces, G_{110} and $G_{101,}$ have different supersaturation dependences, resulting in change in crystal shape, as illustrated in the inset of Figure 6. when the supersaturation is less than 1.8, the growth rates are quite small and almost no growth in both face directions, which is consistent with the conclusion that cessation of the crystal growth might occur at low supersaturation, like a 'dead zone' where little growth occurs on both faces. For protein crystallization, it is important to control the extent of supersaturation because not only the size, shape and solid-state phase of the product crystals are dependent on the supersaturation profile, but also spuersaturation profile can affect the occurrence of different growth mechanisms. There is a concept named as 'critical supersaturation', which is almost two orders of magnitude higher than that required for a typical inorganic system investigated by experiments. At low supersaturation (high temperature) the distance between opposite {101} pyramids far exceeded the distance between opposite {110} faces. At high supersaturation (low temperature) the reverse occurred. It is worthy noting the crossover of the curves for {110} and {101} faces of HEW lysozyme crystals at $\sigma \approx 6.5$, which is the main reason as crystal habits evolution. The crossover resulted mostly from a decrease in the {101} growth rate at low driving force, which is a consequence of the lattice structure. These phenomena have been explained at a molecular level with Monte Carlo method by Durbin and Feher. Additionally, the dashed line in figure 6 shows the averaged growth rate based on a volume equivalent spherical radius from literature. Some variations can be acceptable due to the implementation of different numerical method for values of kinetic parameters. The inset plot in left up corner showed the proportion of growth rate between face {101} and face {110} with supersaturation. It is not linear and has become less and less with the relative growth rate of two faces,

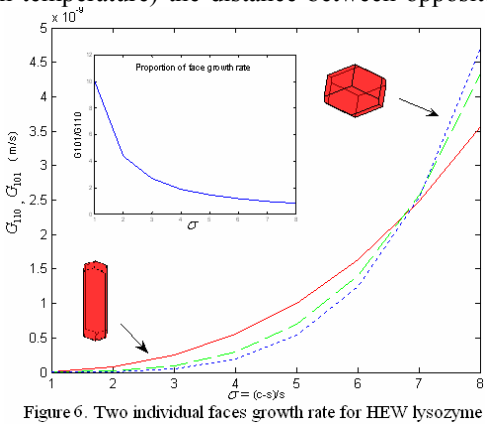


Figure 6. Two individual faces growth rate for HEW lysozyme

resulting in the changes of crystal habits intrinsically with the shape of face {110} from hexagon to rhombus.

As shown in Figure 7 the crystal shape evolution with time for HEW lysozyme crystals can be observed clearly, from needle-like crystals to plate-like crystals during crystallization. Similar results that growth habit changes with σ were obtained by measuring growth rate of {110} and {101} faces of single lysozyme crystal[9]. However, a more quantitative comparison between modelling and these experimental data is not possible for the different operating conditions. The difference of crystal shapes presented here indicates the necessity of using the morphological PB model to capture the habits evolution of protein crystals.

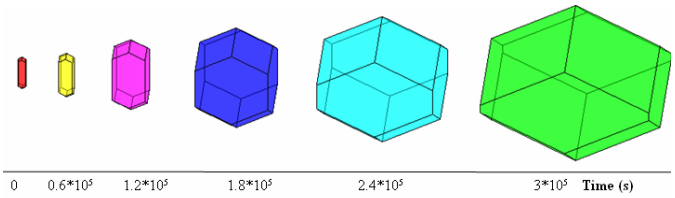


Figure 7. Crystal shape evolution of HEW Lysozyme

4. Conclusion

Although many different experimental techniques have been used to study protein crystallization, the majority has focused on fundamental investigations at the molecular and single crystal levels. Studies on modelling the protein crystal growth behaviour using population balance equations are still scarce and limited to a size definition for a particle as its volume equivalent diameter of a sphere, neglecting crystal morphology information. The morphological population balance model presented in this work for protein crystallization is able to simulate the dynamic evolution of crystal shape as well as size, providing a potentially useful tool for studying the growth behaviour for the whole population of crystals in a crystalliser. The methodology for protein crystallization process modelling has been validated with a HEW lysozyme protein as an example using literature data.

Acknowledgement

Financial support from EPSRC (Grant references EP/C009541/1 and EP/E045707/1) is acknowledged. Thanks are also due to the industrial collaborators including AstraZeneca, Malvern Instruments, National Nuclear Laboratory, Pfizer, Syngeta and 3M Health Care. We also thank China Scholarship Council for sponsoring the first author's visiting study in University of Leeds.

References

1. Ma, C.Y and X.Z. Wang, AIChE J, 2008, **54**, 2321-2334.
2. Ma, C.Y., X.Z. Wang and K.J. Roberts, AIChE J, 2008, **54**, 209-222.
3. Puel, F., G. Fevotte, and J.P. Klein, Chem Eng Sci, 2003, **58**:3715 - 3729
4. Shi, D., P. Mhaskar, N.H. El-Farra, and P.D. Christofides, Nanotech, 2005, **16**: S562-S574.
5. Rosenberger, F., S.B. Howard, J.W. Sowers, and T.A. Nyce, J Cry Grow, 1993, **129**: 1-12.
6. Howard, S.B., P.J. Twigg, J.K. Baird and E.J. Meehan, J Cry Grow, 1988, **90**: 94-104.
7. Cacioppo, E. and M.L. Pusey. J Crystal Growth, 1991. **114**: 286-292.
8. Forsythe, E.L., R.A. Judge, and M.L. J Chem Eng Data, 1999, **44**: 637-640.
9. Durbin, S.D. and G. Feher. J Crystal Growth, 1986, **76**(3): 583-592.
10. Forsythe, E.F. and M.L. Pusey,. Acta Cryst., 1994, **D50**: 614-619.
11. Hojjati, H. and S. Rohani,. Chem Eng Processing, 2005, **44**: 949-957.
12. Ma, D.L., D.K. Tafti, and R.D. Braatz, InD Eng Chem Res, 2002. **41**: 6217-6223.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Pareto Optimization of an Industrial Ecosystem: sustainability maximization

Juliana G. M.-S. Monteiro[a], Patricia A. C. Silva[a], Ofélia Q. F. Araújo[a], José L. de Medeiros[a]

[a]*Dept of Chemical Engineering, Federal University of Rio de Janeiro, Av. Horácio Macedo, n. 2030, Centro de Tecnologia, Bloco E, sala 209, Rio de Janeiro, 21941-909, Brazil*

Abstract

This work aims to design an Industrial Ecosystem for sequestrating of both CO_2 and glycerol in a Chemical Complex with 15 integrated processes to produce: methanol, ethylene oxide, ammonia, urea, dimethyl carbonate, ethylene glycol, glycerol carbonate, β-carotene, 1,2-propanediol and olefins. The Complex is simulated using HYSYS (AspenTech). Processes environmental impact (EI) is calculated using the Waste Reduction Algorithm, while Profit (P) is estimated using classic correlations. The simulation environment is guided by a Matlab (Mathworks) code, which communicates with HYSYS. The performance objective is granting maximum process sustainability, which means finding a compromise between P maximization and EI minimization. Sustainability maximization is understood as a multi-criterial optimization problem, addressed by means of the Pareto optimization methodology for trading off P vs. EI. The general idea behind Pareto optimization is to find a set of solutions in which P can only be improved by compromising on EI and vice-versa. In this work, given the high CPU time demanded for simulating the ecological complex, a procedure for screening the alternatives of CO_2 allocation in the complex is investigated and a reduced Pareto set for the proposed Ecological Chemical Complex is presented.

Keywords: Pareto optimization, sustainability, CO_2 sequestration, process design.

1. Introduction

Global climate change, caused mainly by the increase of CO_2 concentration in atmosphere, is a major concern of the Chemical Industry. Additionally, if biodiesel is to substitute petroleum diesel, glycerol might become a management/environmental problem, since it's generated as a by-product in a 10% (w/w) ratio. The Chemical Industry is currently focusing on environmental performance of processes and products, trying to fit its operation to prevent pollution, using management tools such as Life Cycle Analysis and applying Sustainability Metrics. This work approaches the conceptual design of an Industrial Ecosystem for mitigating the environmental issues attaining chemical and biochemical sequestration of both CO_2 and glycerol to produce: methanol, ethylene oxide (EO), ammonia, urea, dimethyl carbonate (DMC), ethylene glycol (EG), glycerol carbonate (GC), beta-carotene, 1,2-propanediol (1,2-PD), ethylene carbonate (EC) and olefins. Furthermore, a procedure of preliminary screening of the optimal operating region is presented.

2. Chemical Complex Design and Optimization

2.1. Chemical Complex Conceptual Design

2.1.1. Structure

The Chemical Complex proposed in this work is composed by 15 integrated processes, as illustrated in Figure 1. The objective of this complex is to optimize CO_2 chemical and biochemical reuse, by setting the most sustainable CO_2 allocation within the system.

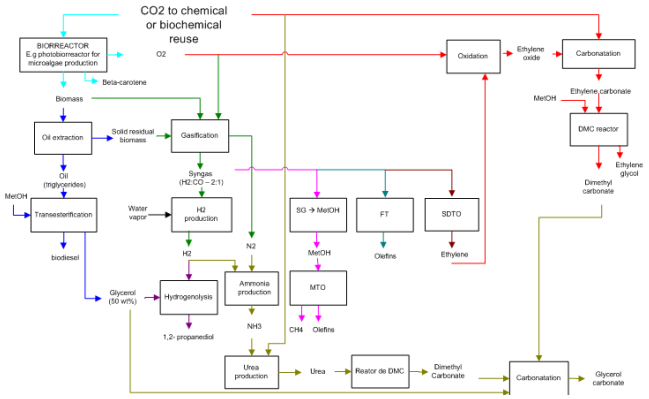


Figure 1: Chemical Complex structure

The processes considered are shortly described as follows: (a) **microalgae biomass and β-carotene production**: CO_2 is biofixed by *Dunaliella salina* cultivation in a photobioreactor. The biomass is then harvested and the β-carotene is separated. The process is simulated using technical data from Chisti (2007) and Araújo *et al*. (2008); (b) **biomass gasification**: biomass is oxidized by O_2 and H_2O to generate syngas with CO:H_2 ratio = 1:2 (Olofsson *et al*., 2005); (c) **biodiesel production**: lipids extracted from biomass react with methanol to produce biodiesel. Glycerol is generated as by-product, as described in González *et al*. (1998); (d) **1,2-PD production**: in a first step, glycerol is dehydrated to acetol; then acetol is hydrogenated to 1,2-PD (Dasaria et al., 2005); (e) **hydrogen production**: vapor reacts with syngas (water-gas shift reaction), generating H_2 and CO_2; (f) **ammonia production**: N_2 from biomass gasification reacts with H_2; (g) **urea production**: ammonia reacts with CO_2 to produce urea (Baal and Lemmen, 2003); (h) **methanol production**: CO and H_2 react under high pressure (7,5 MPa) to form methanol; (i) **methanol-to-olefins (MTO)**: methanol is converted into light olefins, as described by Nouri and Tillman (2005); (j) **syngas-to-DME-to-olefins (SDTO)**: this is a modification of the MTO process, proposed by Cai *et al.* (1995): first, syngas is converted into dimethyl-ether (DME) and then DME is converted to hydrocarbons. The advantage of SDTO route is that syngas conversion to DME is higher than syngas conversion to methanol; (k) **Fisher-Tropsch (FT) synthesis**: FT reaction is modeled according to the Anderson–Schultz–Florey (ASF) model (Trepanier *et al*., 2009). The chain growth probability value was optimized in order to maximize FT selectivity to light olefins; (l) **EO production**: ethylene and oxygen react to produce EO, CO_2 and water. (Coombs et al., 1997); (m) **DMC production from EO**: EO reacts with CO_2, forming EC. EC then reacts with methanol to generate DMC and EG (Cui *et al*., 2004); (n) **DMC production from urea**: methanol and urea reacts forming methyl carbamate, which reacts with methanol to generate DMC and NH_3 (Wang *et al*., 2007); (o) **GC production**: DMC reacts with glycerol to form GC (Rokicki *et al*., 2005).

The thermodynamic equilibriums of the reactions of processes (b), (e), (f) and are calculated by HYSYS.

2.1.2. Modeling

Casavant and Côtè (2004) ascertain the benefits of using chemical process simulation to design industrial ecosystems, a practice that allows "the design of material exchanges and integrated waste treatment to reduce environmental impact; design facilities that maximize energy efficiency; and design facilities that conserve material use". However, rigorous simulation of a Chemical Complex demands a high computational effort for the convergence of all the recycles involved. In the addressed problem, the simulation flowsheet has 228 streams and 115 unit operations. The optimization of the entire complex demands the simulation to be solved several times, raising the computational effort beyond practicability boundary.

In the adopted procedure, the processes are first rigorously simulated in separate, using HYSYS. Kinetic data (whenever available) are used for modeling the reactions, and reactors (PFRs and CSTRs) and distillations columns are properly sized. These preliminary rigorous simulations allow the calculation of technical indexes such as the conversion of a reactant in a given condition, or the recovery of a product in a distillation column. The Chemical Complex is then simulated using simpler operations models, yet employing the technical indexes calculated and reaction conditions. For instance, PFR reactors are replaced by Conversion reactors, and columns are replaced by HYSYS unit operation block called "component splitter", a decision that accelerates the flowsheet convergence, significantly reducing CPU simulation time. The designed system presents several integrations between material and energy flows: mixers and tees dictate the extent of material integration between processes. The energy integration is calculated considering that the utilities flows are linked into a heat exchangers network, as in Bulasara *et al*. (2008). Additionally, heat exchangers are installed within each process whenever energy integration between process flows is possible.

2.2. Procedure for Chemical Complex Performance Analysis

The performance of the Complex is analyzed by investigating its Sustainability, herein understood as a two-dimensional indicator, considering both economical and environmental aspects. The analysis requires that the Profit (P) and the Environmental Impact (EI) of the Chemical Complex be calculated for the evaluated cases.

The Environmental Impact (EI) Function is calculated using the Waste Reduction Algorithm (WAR), which characterizes sustainability with an index that measures Potential Environmental Impacts (PEI) associated to each process emissions. Technical emission factors are obtained from literature or estimated using HYSYS, in terms of quantity of pollutant released for quantity of product obtained. The greenhouse gas emissions due to utilities demand in operations are calculated as a function of the energy inputs/outputs of the equipment used in each plant. The facilities are considered to employ natural gas as fuel. P of each process is calculated according to Equation 1 (Knapp and Doherty, 1990). ISBL is calculated using Douglas' (1988) correlations; utilities costs are summarized in Table 1; the plants operate for 7200h/y; the Marshal and Swift index (M&S) used is 1399, referring to the year 2007; ISBL for urea production plant is estimated from the values reported by van Baal and Lemmen (2003).

$$P = Revenue - ISBL - (Raw\ Material + Utilities) \quad (1)$$

Table 1: Cost of Equipments (Douglas, 1988) and Utilities.

Cost of Utilities	
Vapor	USD 6.98/t
Water	USD 6.90^{-3}/m³
Electrical Energy	USD 0.43/kWh

Both P and EI are affected by material and energy integration decisions, often in different ways – since a more profitable process may have more EI and/or be more energy intensive than a less profitable one. Therefore, the mixers and tees of the process are the keys to investigate various situations, allowing the decisions of whether raw material to a given process is obtained within the system (in closed loops), or from outside system's boundaries, or in a combined internal and external solutions. The same applies to process products: is it better to sell a product (send it outside the system) or to use it as reactant within the system?

Normally, as pointed out by Gibbs and Deutz (2007), cycling of materials and energy is understood as a problem of multi-objective minimization of waste production, energy consumption and raw material consumption. Herein, the focus is given to P and EI, an approach that should include cases in which, for instance, a rise in energy consumption due to higher flows in a given process actually leads to both P rising and EI lowering. After the connections between processes are made and process constraints are applied, the Chemical Complex flowsheet remains with high dimensionality: 15 flow ratios (in 11 tees) are to be set in order to solve a flowsheet case, as shown in Table 2.

Table 2: Chemical Complex degree of freedom

TEE op	Inlet	Outlets	Flow ratios
T1	syngas	to H2 production	*to be set (1)*
		to MetOH production	*to be set (2)*
		to FT	*to be set (3)*
		to DME production	calculated by difference
T2	D. salina	to gasification	*to be set (4)*
		to biodiesel production	calculated by difference
T3	Methanol produced	to DMC production (from urea)	adjusted to set MetOH:urea ratio = 2:1 in DMC production
		to MTO	*to be set (5)*
		to be sold	*to be set (6)*
		to DMC production (from EO)	adjusted to set MetOH:EO ratio = 2:1 in DMC production
		Methanol makeup	calculated by difference
T4	CO_2 from EO production	recycle	*to be set (7)*
		CO_2 purge	calculated by difference
T5	Methanol	to reactor 1 (DMC from urea)	*to be set (8)*
		to reactor 2 (DMC from urea)	calculated by difference
T6	Glicerol produced	to be discharged	*to be set (9)*
		to 1,2-PD production	*to be set (10)*
		to GC production	calculated by difference
T7	Methanol	MetOH purge	*to be set (11)*
		MetOH recycled	calculated by difference
T8	DMC produced	to be sold	*to be set (12)*
		to GC production	calculated by difference
T9	EO produced	to DMC production	calculated by difference
		to be sold	*to be set (13)*
T10	NH_3 produced	to urea production	calculated by difference
		to be sold	*to be set (14)*
T11	urea produced	to DMC production	calculated by difference
		to be sold	*to be set (15)*

Flow ratios must respect the following constraints: each value must be contained in the [0,1] interval and the sum of flow ratios of the same tee must equal 1. The number of combinations of flow ratios sets is explosive. Therefore, MATLAB is connected to HYSYS and set to automatically generate random values for each flow ratio, with the given constrains, and save each feasible case. Cases' feasibility depends on HYSYS's capacity of solving the flowsheet for the proposed flow ratio values. The solutions generated by HYSYS are plotted in a P vs. EI graphic, for mapping the Chemical Complex Sustainability performance. The multi-criterial optimization problem of Sustainability maximization is then addressed by means of the Pareto optimization methodology. The general idea behind Pareto optimization is to find a set of solutions in which P can only be improved by compromising on EI and vice-versa.
An initial set of converged flowsheet solutions, 25 cases in the present work, is analyzed for the calculation of the linear dependence of P and EI to each flow ratio. The search for solutions within the Pareto Set is oriented to progressively vary only the most impacting flow ratios, progressively reducing the pool of flow ratios "to be set".

3. Pareto Set Results and Conclusion

Figure 2 shows the map of feasible solutions for the Chemical Complex. Both P and EI are reported in relative terms, divided by the quantity of CO_2 sequestrated in each case. The Pareto set points are market with a red circled and joined by a line – the Pareto frontier. Every point above the Pareto frontier is dominated by a point of the frontier, while the region below the frontier is unfeasible. In a previous work, Monteiro et al. (2009) presented a 2-D Sustainability Function (SF) for process sustainability assessment, as shown in equations 2 and 3:

$$SF = \omega_P P - \omega_{EI} EI_m \tag{2}$$

$$EI_m = EI \frac{M_{CO_2}}{M_{product}} \tag{3}$$

where: SF = sustainability function; P = profit; EI_m = modified environmental impact; ω_k = weighting factor associated to function k; EI = environmental impact; M_{CO2} = mass of CO_2-equivalent emitted; $M_{product}$ = mass of product obtained.

This metric was employed then for choosing the best eco-technology between two alternatives. As discussed in the referred work, the ratio ω_P/ω_{EI} is the relevant measure of such metric. This metric is applied to the cases herein studied varying ω_P from 0 to 10 and keeping ω_{EI} constant and equal to 1. The points with maximal SF coincide with the points of the Pareto frontier, meaning that the Sustainability Function can be used as objective function for Sustainability optimization.

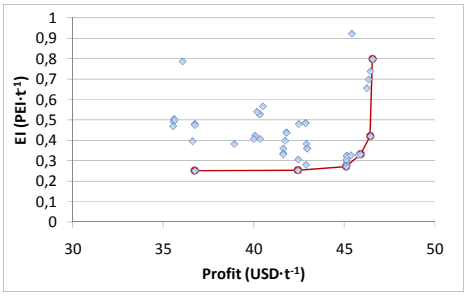


Figure 2: Pareto set (SF maximization solutions)

References

Y. Chisti, 2007, Biodiesel from microalgae, Biotechnology Advances, 25, 294-306.

O. Q. F. Araújo, C. N., Gobbi, R. M. Chaloub, M. A. Z., Coelho, 2008, Assassment of the impact of salinity and irradiance on the combined carbon dioxide sequestration and carotenoids production by Dunaliella salina: a mathematical model. Biotechnology and Bioengineering, Vol. xxx, No. xxx, DOI 10.1002/bit.22079.

I. Olofsson, A. Nordin, U. Söderlind, 2005, Initial review and evaluation of process technologies and systems suitable for cost-efficient medium-scale gasification for biomass to liquid fuels, ETPC Report, University Umea and Mid Swedish University, Sweden.

M. J. I. González, A. R. Medina, E. M. Grima, A. G. Giménez, M. Crastens, L. E. Cerdána, 1999, Optimization of fatty acid extraction from Phaeodactylum tricornutu UTEX 640 biomass, JAOCS, 12.

M.A. Dasaria, P. Kiatsimkula, W.R. Sutterlinb, G.J., Suppes, 2005, Low-pressure hydrogenolysis of glycerol to propylene glycol, Apllied Catalysis, 281, 225-231

H. van Baal, W. Lemmen, 2003, The advantages of a large capacity single line urea plant, Stamicarbon bv Technical Report. Available at http://www.stamicarbon.com/publi_presen/brochures/_en/index.htm

S. Nouri, A.-M. Tillman, 2005, Evaluating synthesis gas based biomass to plastics (BTP) technologies, Chalmers University of Technology, Göteborg.

G. Cai, Z. Liu, R. Shi, C. He, C. Yang, C. Sun, Y. Chang, 1995, Light alkenes from syngas via dimethyl ether, Applied Catalysis, 125, 29-38.

M. A. Trepanier, A. K.Tavasoli, N. Dalai, Abatzoglou, 2009, Co, Ru and K loadings effects on the activity and selectivity of carbon nanotubes supported cobalt catalyst in Fischer–Tropsch synthesis. Applied Catalysis A: General. 353. 193–202.

J. Coombs, K. Daniel, L. Palombo, 1997, Celanese Clearlake Ethylene Oxide Reactor Revamp, JChem Inc. Available at http://www.owlnet.rice.edu/~ceng403/ethox97.html

H. Cui, T. Wang, F.Wang, C. Gu, P. Wang, Y. Dai, 2004, Kinetic Study on the One-Pot Synthesis of Dimethyl Carbonate in Supercritical CO_2 Conditions, Ind. Eng. Chem. Res., v. 43, p. 7732-7739.

G. Rokicki, P. Rakoczy, P. Parzuchowski, M. Sobiecki, 2005, Hyperbranched aliphatic polyethers obtained from environmentally benign monomer: glycerol carbonate; Green Chem., 7, 529–539

T. E. Casavant, R. P. Côté, 2004, Using chemical process simulation to design industrial ecosystems, Journal of Cleaner Production 12, 901–908.

V. K. Bulasara, R. Uppaluri, A. K. Ghoshal, 2008, Revamp study of crude distillation unit heat exchanger network: Energy integration potential of delayed coking unit free hot streams. Applied Thermal Engineering. Article in press.

J.P. Knapp, M.F. Doherty, 1990, Thermal Integration of Homogeneous Azeotropic Distillation Sequences, AIChE J. vol. 36, n.7, 969-984.

J. M. Douglas, 1998, Conceptual Design Of Chemical Processes. McGraw-Hill.

D. Gibbs, P. Deutz, 2007, Reflections on implementing industrial ecology through eco-industrial park development, Journal of Cleaner Production, 15, 1683-1695.

J. G. M-S. Monteiro, O. Q. F. Araújo, J. L. Medeiros, 2009, Sustainability Metrics for Eco-Technologies Assessment, Part II: Life Cycle Analysis. Clean Technologies and Environmental Policy. Article in press.

F. Wang, N. Zhao, J. Li, W. Zhao, F. Xiao, W. Wei, Y. Sun, 2007, Modeling of the Catalytic Distillation Process for the Synthesis of Dimethyl Carbonate by Urea Methanolysis Method. Ind. Eng. Chem. Res, 46, pp. 8972-8979.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Exergy and Sustainable Development for Chemical Industry Revisited

Moises Teles dos Santos,[a] Song Won Park[a]

[a]*LSCP/CESQ - Department of Chemical Engineering, Polytechnic School - University of São Paulo Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil. e-mail address: sonwpark@usp.br*

Abstract

It becomes necessary the standardization of environmental sustainability indicators for energy conversion systems as well as the measurement of the impact of their emissions on the environment. The exergy concept has been evaluated as a useful tool for this purpose. It is based on thermodynamic criteria and reveals the transformation capacity of a given energy carrier in the environment and the consumption and depletion flows of the useful resources available on nature. Natural and artificial systems are supported by energy useful potential (exergy) and the evaluation of this exergy consumption in relation to natural resources is an indication of process sustainability. The aim of this work is review the environmental sustainability concepts from the perspective of exergy analysis, show the contribution of exergy to the Life Cycle Analysis and clarify the relations between exergy and environmental impacts measurements. In the final section, we illustrate with two real industrial examples.

Keywords: exergy, energy, sustainability, natural resources.

1. Introduction

The search for sustainable development must be based on three pillars: economic viability, social concerns and ecological issues. Therefore, despite the fact that sustainability is a multidisciplinary concept involving different areas, process engineering plays a fundamental role in this context as it can promote more sustainable process and products (Bakshi and Fiksel, 2003). Application of this concept, however, has claimed for development of metrics: criteria and sustainability indicators definitions that can be used for measurement, evaluation and comparison of industrial activities impacts on the environment (Yi *et al*., 2004; Ness *et al*., 2007). In this context, Green Engineering is the design, discovery and implementation of engineering solutions for sustainability in a multi-scale approach involving molecular, product, process and systems level (Anastas; Zimmerman, 2006; Garcia-Serna *et al*., 2007; Charpentier, 2008). Anastas and Zimmerman (2006) stated the so-called The Twelve Principles of Green Engineering and the goal of this review is highlight how the use of Second Law of Thermodynamics by means of exergy concept can be used to address some of these principles (as listed below), besides its usefulness as a metric for sustainability.

- **Principle 4**: System components should be designed to maximize mass, energy and temporal efficiency.

- **Principle 5**: System components should be output pulled rather than input pushed through the use of energy and materials.
- **Principle 6**: Embedded entropy and complexity must be viewed as an investment when making decision choices on recycle, reuse or beneficial disposition.
- **Principle 10**: Design of processes and systems must include integration of interconnectivity with available energy and material flows.

2. Exergy

2.1. Introduction

Energy is always a conserved quantity. However, it becomes less available to perform useful work and drive processes. This useful work is quantifiable by means of exergy. The exergy is not conserved and this fact can be used for evaluate processes efficiency and consumption of natural resources. Exergy provides qualitative measure of material and energy streams and can identificate process irreversibility. Exergy is based on simultaneous energy and entropy balances and can be defined as the maximum amount of work when a system is brought to thermodynamic equilibrium with environment, in terms of temperature, pressure and composition, through reversible process (Szargut *et al*., 88). Thus, exergy is a measure of the maximum useful level of an energy carrier, as imposed by Second Law.

2.2. Exergy Balance

A linear combination between First and Second Laws of Thermodynamics (energy and entropy balances) for an open system with *m* material streams entering, *n* material streams leaving the system and *p* heat transfers, with no significant changes in kinetic and potential energy leads, in steady state, to:

$$\sum_{i=1}^{m}[\dot{m}_e(\bar{h}_e - T_0\bar{s}_e)]_i - \sum_{i=1}^{n}[\dot{m}_s(\bar{h}_s - T_0\bar{s}_s)]_i = \dot{W} + T_0\,\dot{S}_g + \sum_{i=1}^{p}\dot{Q}_i(1-\frac{T_0}{T_i}) \quad (1)$$

The second term of this equation is the maximum amount of available work due to the change from the initial to the final conditions of temperature and pressure. In this term, the quantity T_0S_g is the lost work (exergy destruction). This loss is due to entropy generation (S_g), caused by process irreversibilities. In the particular case where an entering stream achieves physical equilibrium with environment (T_0, P_0) through reversible processes (S_g=0), the work obtained is the maximum possible and represents the total physical exergy of the stream. For systems in physical equilibrium with environment but with different composition, there is a remaining potential, which can be used to cause changes. The chemical exergy (b_{ch}) is therefore defined as:

$$\bar{b}_{ch} = \sum_{i=1}^{q} x_i(\mu_{0i} - \mu^*_{0i}) \qquad \text{[kJ/kg]} \quad (2)$$

Where x_i is the mole fraction of the component *i* in the system, μ_{0i} is the chemical potential of each component and μ_{0i}^* the chemical potential of the substance (or reference substances) on defined environmental conditions.

2.3. Applications

2.3.1. Industrial use

For energy conversion systems, recent exergy analysis has been used successfully for evaluation and optimization (Hammond, 2007; Kanoglu and Dincer, 2009) and evaluation of environmental impacts (Meyer *et al*., 2009). The literature presents many cases in different types of industries: sugarcane industry (Pellegrini and Oliveira Jr., 2007; Serra *et al*., 2008; Kamate and Gangavati, 2009), petrochemicals (Rivero, 2002; Bernardo *et al*., 2006) and hydrogen production (Orhan *et al*., 2009). Improvements on process energy efficiencies can lead to the same level of production and services with less consume and depletion of natural resources, matching principles 4,5 and 10. Therefore, exergy, as an auxiliary tool for industrial efficiency optimization also represents a potential tool to achieve sustainable systems. For energy systems, it can be noted the potential of exergy as auxiliary tool for process integration aiming to maximize the use of available process energy streams as highlighted by Principle 10. Some cases in this area can be found in the works of Serra *et al*. (2008) and in a methanol plant (Anantharaman *et al*., 2006).

2.3.2. Impacts on ecosystems evaluation

In a general way, the larger the exergy of a pollutant, the larger the disturbance caused on environmental equilibrium (Kotas, 95). Being a measure of the difference between a given amount of matter and the environment, the exergy content of an emission represents the maximum potential environmental changes until the matter reaches total equilibrium with the environment. Environmental pollution can be interpreted as any interference in the ecosystem equilibrium through material or energy streams. Such streams cause unbalances, due to differences concerning the environmental parameters (temperature, pressure and composition). Exergy, better than energy, reflects these differences. Further discussions about ecosystem evaluation by exergy can be found in the book of Dincer and Rosen (2007): Exergy: Energy, Environment and Sustainable Development.

2.3.3. Depletion of natural resources evaluation

Resources are materials in disequilibrium with the environment, what leads to an exergy content. This is the useful value of resources, not complete described by energy or mass. Minimize depletion of these natural resources is one of the principles of sustainable engineering (Abraham, 2006). Exergy can be used as an auxiliary tool for sustainability indicators of renewable resources as highlighted by Omer (2008). Regarding Principle 6 (Table 1), a measure is required that is able to account for the quality losses during recycling. These quality losses cannot be measured by mass balances, as the quality degradation cannot be translated by mass measures alone (Amini *et al*., 2007); exergy is then proposed as a measure of the efficiency of resources use. An interesting concept that arises from this discussion is the thermo-ecological cost, defined as the cumulative consumption of non-renewable exergy due to the production of a particular product (Szargut and Stanek, 2008). Further discussions of how thermodynamics plays a role on economic level and not only on industrial process level can be also found in Ayres (2008) and in Jorgensen (2006).

3. ELCA: Exergetic Life Cycle Assessment

Life Cycle Analysis (LCA) is a holistic approach to cope with environmental performance of a product, based on its entire life cycle, from the raw material acquisition to ultimate disposition. LCA is very useful to identify all material and energy flows consumed and produced during the entire life period of products.

However, one of the limitations of this method is that it cannot evaluate different emissions in regard to environmental effects in the same basis. As exergy can clarify and quantificate resource depletion, waste emission and process losses when comparing different substances, there is an increasing number in literature reports of application of the so-called Exergetic Life Cycle Analysis (ELCA) (Cornelissen and Hirs, 2002), as reported for electricity generation (Bakshi and Ukidwe, 2006), hydrogen production (Granovskii et al., 2007) and coal gasification and aluminium production (Ulgiati et al., 2006). There are three major advantages of using exergy in LCA: 1. Provide a common measure of inputs and outputs and allows the estimation of exergetic efficiency that is an indication of potential improvements. 2. Facilitate the comparison between different materials in relation to environmental impacts. 3. Facilitate reporting and monitoring of companies and countries of environmental indicators over time (Ayres *et al*., 1998).

4. Cases

4.1. Efficiency in use of resources in a cogeneration plant

We applied the exergy efficiency concept in a cogeneration plant in a pulp and paper industry (Figure 1). The resources are biomass, black liquor (sub-product of Kraft papermaking) and fossil fuel-based oil. By mean of physical exergy of different streams, two assessment of efficiency can be done: energy basis or utilizable part of energy (exergy). An exergy efficiency, ψ, can be defined as $\psi=E_{out}/E_{in}$, which represents the portion of resources that is delivered in useful form by the products. The energetic efficiency is about 81.55%, while in exergy basis, is 24,90%. Thus, exergy approach clearly reveals that most part of exergy stocked in fuels is not converted, but destroyed by irreversible phenomena taking place in the system. As noted by Rosen et al. (2008), the use of energy as a measure for identifying and measuring the benefits of energy systems can be misleading and confusing, while exergy identify possible efficiency improvements.

4.2. Ammonia process

Figure 3 shows the typical mass and energy flows of an ammonia process, taken form Bakshi and Ukidwe (2006). It reveals the ability of exergy analysis for cope with different materials and energy flows in the same and rational basis. In the first case, different units are used leading to a lack of information about the quality of the resources. In contrast, exergy flows shows the streams in terms of consistent units of exergy. It is also showed the value of the so-called cumulative exergy consumption (436 MW) which considers exergy losses in the process and its supply chain up to natural resources.

5. Concluding Remarks

Alternatives thermodynamic approaches for energy efficiency assessment and depletion of natural resources can provide additional and useful information seeking for sustainable processes. Create engineering solutions beyond current or dominant technologies, improving and innovate to achieve sustainability is one of the principles of sustainable engineering (Abraham, 2006) and exergy analysis (a Second Law based concept) can match this principle as been an auxiliary and non-conventional tool. Matter and energy are always conserved in any process they are used. The real indication of natural resources depletion is related to their useful value (exergy) in conducting process and maintaining suitable conditions of life on Earth.

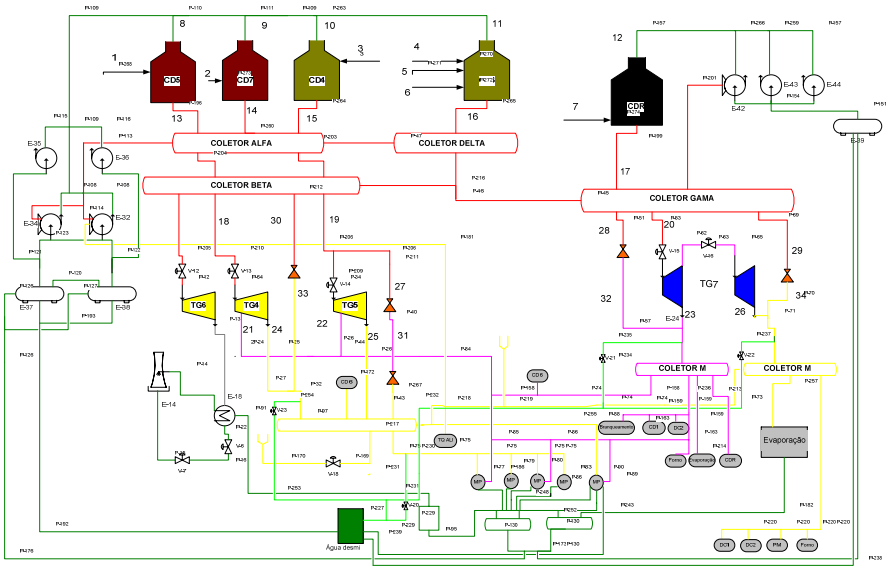


Figure 1: Cogeneration system in a pulp and paper industry.

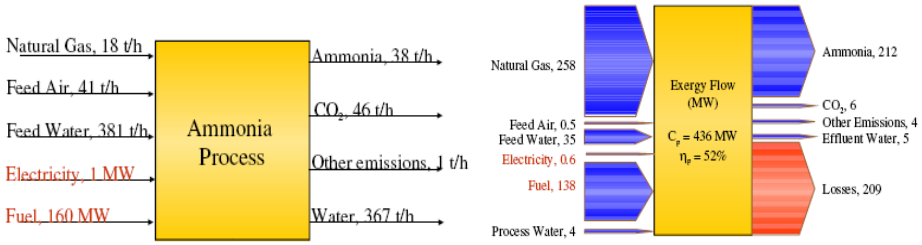


Figure 2: Energy and exergy flows form a typical ammonia process. Source: Bakshi and Ukidwe (2006).

References

Abraham, M.A., 2006, Principles of Sustainable Engineering, Sustainability Science and Engineering: Defining principles, M.A. Abraham, Elsevier B.V., Amsterdam, The Netherlands.

Amini, S.H., Remmerswaal, J.A.M., Castro, M.B., Reuter, M.A., 2007, Quantifying the quality loss and resource efficiency of recycling by means of exergy analysis, Journal of Cleaner Production, 15, 907-913.

Anantharaman, R., Abbas, O. S., Gundersen, T., 2006, Energy Level Composite Curves - a new graphical methodology for the integration of energy intensive processes, Applied Thermal Engineering, 26, 1378–1384.

Anastas, P.T., Zimmerman, J.B., 2006, The Twelve Principles of Green Engineering as a Foundation for Sustainability, Sustainability Science and Engineering: Defining principles, M.A. Abraham, Elsevier B.V., Amsterdam, The Netherlands.

Ayres, R.U., 2008, Sustainability economics: Where do we stand?, Ecological Economics, 67, 281-310.

Ayres, R.U., Ayres, L.W., Matinás, K., 1998, EXERGY, WASTE ACCOUNTING AND LIFE-CYCLE ANALYSIS, Energy, 23, 355-363.
Bakshi, B.R., Fiksel, J., 2003, The Quest for Sustainability: Challenges for Process Systems Engineering, AIChE Journal, Vol. 49, No. 6, 1350-1358.
Bakshi, B.R., Ukidwe, N.U., 2006, The Role of Thermodynamics in Life Cycle Assessment of Existing and Emerging Technologies, Proceedings of the 2006 IEEE International Symposium on Electronics and the Environment.
Bernardo, P., Barbieri, G., Drioli, E., 2006, AN EXERGETIC ANALYSIS OF MEMBRANE UNIT OPERATIONS INTEGRATED IN THE ETHYLENE PRODUCTION CYCLE, Chemical Engineering Research and Design, 84, 405–411.
Charpentier, J.C., 2008, Perspective on multiscale methodology for product design and engineering, Computers and Chemical Engineering, In Press.
Cornelissen, R. L., Hirs, G.G., 2002, The value of the exergetic life cycle assessment besides the LCA, Energy Conversion and Management, 43, 1417-1424.
Dincer, I., Rosen, M.A., 2007, Exergy: Energy, Environment and Sustainable Development, Elsevier.
Garcia-Serna, J., Perez-Barrigon, L., Cocero, M.J., 2007, New trends for design towards sustainability in chemical engineering: Green engineering, Chemical Engineering Journal, 133, 7–30.
Granovskii, M., Dincer, I., Rosen, M.A., 2007, Exergetic life cycle assessment of hydrogen production from renewables, Journal of Power Sources, 167, 461–471.
Hammond, G.P., 2007, Industrial energy analysis, thermodynamics and sustainability, Applied Energy, 84, 675–700.
Jorgensen, S.E., 2006, Eco-exergy As Sustainability, WIT Press, Southampton.
Kamate, S.C., Gangavati, P.B., 2009, Exergy analysis of cogeneration power plants in sugar industries, Applied Thermal Engineering, 29, 1187–1194.
Kanoglu, M., Dincer, I., 2009, Performance assessment of cogeneration plants, Energy Conversion and Management, 50, 76–81.
Kotas, T., 1995, The Exergy Method of Thermal Plants Analysis, Melbourne, Krieger.
Meyer, L., Tsatsaronis, G., Buchgeister, J., Schebek, L., 2009, Exergoenvironmental analysis for evaluation of the environmental impact of energy conversion systems, Energy, 34, 75–89.
Ness, B., Urbel-Piirsalu, E., Anderberg, S., Olsson, L., 2007, Categorising tools for sustainability assessment, Ecological Economics, 60, 498 – 508.
Omer, A.M., 2008, Energy, environment and sustainable development, Renewable and Sustainable Energy Reviews, 12, 2265–2300.
Orhan, M. F., Dincer, I., Rosen, M. A., 2009, The oxygen production step of a copper–chlorine thermochemical water decomposition cycle for hydrogen production: Energy and exergy analyses, Chemical Engineering Science, 64, 860 – 869.
Pellegrini, L. F., Oliveira Jr, S., 2007, Exergy analysis of sugarcane bagasse gasification, Energy, 32, 314–327.
Rivero, R., 2002, Application of the Exergy Concept in the Petroleum Refining and Petrochemical Industry, Energy Conversion and Management, 43, 1199-1220.
Rosen, M.A., Dincer, I., Kanoglu, M., 2008, Role of exergy in increasing efficiency and sustainability and reducing environmental impact, Energy Policy, 36, 128–137.
Serra, L.M., Lozano, M., Ramos, J., Ensinas, A.V., Nebra, S.A., 2008, Polygeneration and efficient use of natural resources, Energy, In Press.
Szargut, J., Morris, D., Steward, F., 1988, Exergy Analysis of Thermal, Chemical and Metallurgical Processes, New York, Hemisphere.
Szargut, J., Stanek, W., 2008, Influence of the pro-ecological tax on the market prices of fuels and electricity, Energy, 33, 137–143.
Ulgiati, S., Raugei, M., Bargigli, S., 2006, Overcoming the inadequacy of single-criterion approaches to Life Cycle Assessment, Ecological Modelling, 190, 432–442.
Yi, H., Hau, J.L., Ukidwe, N.U., Bakshi, B.R., 2004, Hierarchical Thermodynamic Metrics for Evaluating the Environmental Sustainability of Industrial Processes, Environmental Progress, Vol.23, No.4, 302-314.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Evaluation of Adsorbed Polyampholyte Layers by Using Quartz Crystal Microbalance

Deusanilde J. Silva[a], Orlando J. Rojas[b], Song Won Park[a], Martin A. Hubbe[b]

[a] *Department of Chemical Engineering - USP Av. Prof. Luciano Gualberto, n. 380, trav. 3, CEP 05508-900, São Paulo - SP, Brazil.* deusanilde@gmail.com
[b] *Department of Forest Biomaterials, NC State University, Raleigh, NC, USA*

Abstract

Viscoelastic properties of layers of polyampholytes adsorbed on charged surfaces were studied by quartz microgravimetry. By applying the Voigt viscoelastic model the effective mass and thickness of layers after adsorption from solution at different salt concentrations were calculated. The obtained results were compared with the Sauerbrey equation, which applies to the case of thin, rigid adsorbed layers. The estimates of mass and thickness from the Voigt model were typically larger, and were more strongly affected by variations in the ionic strength of adsorbing solution. Since the Voigt model uses multiple frequencies and dissipation overtones, it was found that the calculated adsorbed layer mass was closer to the actual values, while the Sauerbrey approach resulted in underestimation. This observation was explained by the fact that adsorbed layers of polyampholytes were soft and highly dissipative. It was noted that the observed changes in dissipation of the adsorbed polyampholyte layers were comparatively large, which suggests a large amount of coupled water.
Keywords: Quartz microgravimetry; Sauerbrey equation; Voigt model; Adsorbed mass; Layer thickness.

1. Introduction

Polyampholytes are natural or synthetic organic polymers that have both, positive and negative groups in their macromolecular structure (Bohidar, 2002). A distinctive feature of these macromolecules is their anti-polyelectrolyte behavior, which involves low solubility in water near a pH value corresponding to net-neutral charge, and the fact that their solubility at such pH increases with increased concentration of salt in solution. This behavior is explained by the balance of electrostatic interactions between positive and negative groups in the molecule. On the other hand, with increasing concentration of low molecular weight electrolytes, the electrostatic attraction between the opposing groups of charges is reduced, resulting in increased expansion and solubility of the polymer in solution. Polyampholytes have found increased utilization in a wide range of applications in medicine, wastewater treatment, oil recovery and papermaking, to name only a few. Adsorption of charged polymers depends on bulk solution environmental conditions, the charge density of the polymer and the surface charge (Rojas, 2002). Evaluation of the polymer adsorbed layer can be carried out by using the Quartz Crystal Microbalance (QCM) (Rodahl and Kasemo, 1996; Höök et al., 2002; Munro and Frank, 2004; Silva, Yamaguchi et al., 2008; Silva, Hubbe et al., 2008). This device enables studies of the dynamics of polymer adsorption processes by measuring the shift in resonance frequency of a quartz crystal electrode with time. The energy dissipation, related to the damping of the sensor oscillation, can be used to study the viscoelastic properties of the adsorbed layer, i.e., its softness or rigidity. Typically, a soft layer is able to more effectively dampen the crystal oscillation. Information from QCM is

relevant to understand the relationship between the structure of the adsorbed layer and the hydration behavior of macromolecules at interfaces. The thickness and the mass of an adsorbed polymer, with its coupled water can be estimated by using Sauerbrey equation. This equation assumes that the amount of adsorbed mass is proportional to the changes in QCM frequencies, which is the case for adsorbed rigid layers (Sauerbrey, 1959). The QTool software of the Quartz Crystal Microbalance with Dissipation (E4 QCM-D unit from Q-Sense, Sweden) was used to evaluate the adsorption of polyampholyte layers from aqueous solution. The changes in frequencies and energy dissipation at different overtones were used to quantify the adsorbed layer properties. Overall, the main objective was to evaluate the adsorbed polyampholyte layers with regard to the adsorbed mass and thickness, under different salt concentration. This was accomplished by using the Sauerbrey equation and the Voigt viscoelastic model.

2. Experimental

2.1. Materials

2.1.1. Polymer

The amphoteric polymer was provided by Harima Chemicals Inc., Japan. It was prepared by random, free-radical polymerization. The cationic monomer was dimethylaminopropilacrylamide, a tertiary amine, the anionic monomer was itaconic acid and the neutral was acrylamide. Molecular weight of the polyampholyte synthesized was 2.93 x 10^6 Da with cationic to anionic group ratio of 5:4, as measured by NMR (Wang et al., 2002). Aqueous solution (1g/l) viscosity of 2,400 mPa.s (25 ^oC) (Wang et al., 2002) and isoelectric point of 7.3 (Silva, Hubbe et al., 2008).

2.1.2. Model surfaces

Cellulose films were developed by following the protocol described by (Gunnars et al., 2002). Gold-coated quartz sensors (for QCM-D measurements) were used as base substrate for the cellulose films. The surfaces obtained by spin coating deposition consisted of flat, uniform and ultrathin films. The roughness of the dry model films were about 5 nm (root-mean-square value) as measured by atomic force microscopy (AFM). Silica substrates bearing different charge density than that of cellulose, was also used for comparison purposes. Gold and silica sensors were provided by Q-Sense, Sweden.

2.2. Density and viscosity

The densities of the polyampholyte solutions for each salt concentration were obtained by measuring the weight of a fixed volume of 5000 µl with 0.0001 g of accuracy (Adventurer SL from Ohouas). Viscosity measurements were performed according to the TAPPI T230 om-89 protocol by using a viscosimeter from Cannon Instrument Company. Both viscosity and density were measured at room temperature, 25 °C. Four replicates were made of each solution to obtain both parameters.

Table 1 - Fixed fluid density and viscosity parameters for Q-Tools input

Parameter	Salt concentration, mM [NaCl]				
	0.1	1	10	100	1000
Fluid density (ρ_f), kg/m3	1002.94	1000.23	1002.31	1005.74	1043.72
Fluid viscosity (η_f), kg/m.s	0.00107	0.00109	0.00100	0.000968	0.00187

2.3. Models

The thickness and the mass of adsorbed polyampholyte were calculated from the QCM-D data by using two approximations, namely, the Sauerbrey rigid model and the Voigt viscoelastic model. For the Voigt model we used *Q-Tools* software to carry out the respective determination, where three harmonics were used for frequency and

dissipation responses under five ionic strength and two substrate conditions. The changes in QCM frequencies were assumed to be proportional to the adsorbed mass, as is the case of adsorbed rigid layers (Sauerbrey, 1959). Alternatively, we used the Voigt model (Voinova et al., 1999) to relate the adsorbed mass and the viscoelasticity of the film. The fundamental frequency fo (oscillating frequency without adsorbed mass) was 4.95 MHz. The sensitivity constant (C) in the Sauerbrey equation has the value 17.7 ng.Hz^{-1}.cm^{-2}. Thus the added or adsorbed mass Δm (ng/cm^2) was calculated from Eq. 1, where n is the overtone number (n = 1, 3, 5, 7). If a rigid layer is evenly deposited on one or both sides of the electrodes, the resonant frequency will decrease proportionally to the mass of the adsorbed layer (Garg et al., 2008).

$\Delta m = -\frac{C\Delta f}{n}$; and rearranging to the Sauerbrey equation: $\Delta f_m = -\frac{2\times f_0^2 \times \Delta m}{A\times\sqrt{\rho_q \mu_q}}$ *[Eq.1]; [Eq.2]*

where Δf_m is the measured frequency shift, $\rho_q = 2.648$ g/cm^3 is the density of the quartz, $\mu_q = 2.947\times10^{-11}$ dyne/cm^2 is the shear modulus of quartz, and A is the piezo-electrically active area. This equation is not valid if the deposited mass is not rigidly deposited, if it slips on the surface, or if it is not evenly deposited on the electrode surface. The change in the resonant frequency of the QCM crystal also depends on the viscosity and density (η_f, ρ_f) of the gas or liquid medium which is in contact with the crystal according to:

$$\Delta f_{aqueous} = -\frac{n\times f_0^{3/2}}{\sqrt{\pi\rho_q\mu_q}}\left(\rho_f\eta_f\right)^{1/2} \qquad \textit{[Eq.3]}$$

The viscoelastic properties of the adsorbed polymer layer were evaluated by measuring the energy dissipation of the crystal oscillation, D. Typically, a soft layer is able to quickly dampen the resonator oscillation, while a rigid layer would produce a slower dampening effect. The energy dissipation, or damping, was thus defined by Eq. 4 (Rodahl and Kasemo, 1996), where f is the fundamental resonance frequency and τ is the decay time constant of the oscillation amplitude,

$$D = (\pi f\tau)^{-1} \qquad \textit{[Eq.4]}$$

The QCM-D frequency and dissipation data were monitored with time before and after polymer injection on the different substrates at different conditions of ionic strength. Rinsing with buffer solution at the end of the adsorption experiments was typically performed to determine the amount that was irreversibly adsorbed. During the measurements, the QCM liquid chamber was temperature-stabilized to 25 °C and buffer solution was injected at a flow rate of 130µl/min. All experiments with the QCM were repeated at least two times. See the data modeling procedure in Figure 1.

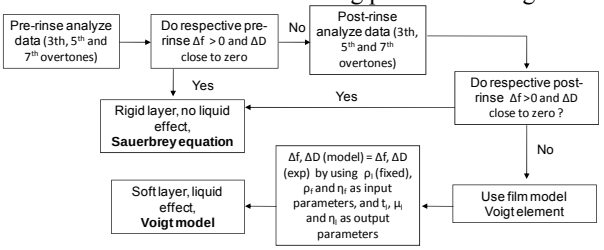


Figure 1 – Flowchart of QCM-D data modeling, subscript *l* is for "layer" and *f* is for "fluid". After: (Munro and Frank, 2004).

The Figure 2 shows the Voigt model where the polymer layer is subjected to an oscillating shear stress σ and behaves like a Voigt element with shear viscosity η_l as a dashpot and shear modulus μ_l as a spring. (Voinova et al., 1999; Tammelin et al., 2004).

Figure 2 – Geometry of the surface model covered by a viscoelastic mono-layer (left) and a schematic depiction of Voigt viscoelastic element (right).

The derivation of Voigt model (Voinova et al., 1999) as described by (Vogt et al., 2004) is

$$\Delta f \approx \frac{t_l \rho_l f}{2\pi\rho_0 t_0}\left(1+\frac{2t_l^2\chi}{3\delta^2(1+\chi^2)}\right);\ \Delta D \approx \frac{2t_l^3\rho_l f}{3\pi f_0\rho_0 t_0}\frac{1}{\delta^2(1+\chi^2)};\ \chi=\frac{\mu_l}{\eta_l f};\ \delta=\sqrt{\frac{2\eta_l}{\rho_l f}} \qquad [Eqs.5\text{-}8]$$

where χ is the ratio of the storage modulus μ_l and the loss modulus η_l. The length δ is the viscous penetration depth.

3. Results and discussion

3.1. Suitability of the viscoelastic model

The dissipation values of the polyampholyte for both silica and cellulose surfaces for all salt concentrations were larger than zero for both pre-rinsing and post-rinsing steps (Figure 3). Substantial deviations from Sauerbrey equation could be expected to occur (Rodahl and Kasemo, 1996; Höök et al., 2002), and a viscoelastic model is therefore recommended (Rodahl and Kasemo,1996; Voinova et al., 1999).

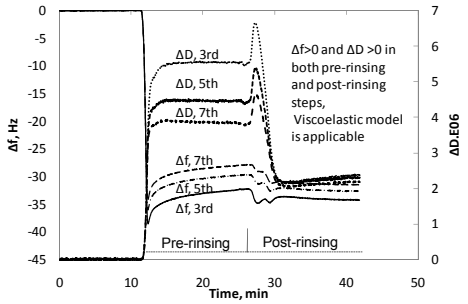


Figure 1 – Changes in frequency and dissipation for polyampholyte adsorbed layer on silica surface at pH 4.3 and 10 mM of salt concentration.

3.2. Experimental and fitted results

The Voigt model was used to fit the thickness and the amount of adsorbed polymer mass. Figure 4 shows the experimental and fitted results for the three harmonics studied (3^{rd}, 5^{th} and 7^{th}), and the comparison between the Sauerbrey and fitted thickness.

A viscoelastic model to evaluate the thickness and the polymer adsorbed mass was more applicable than the Sauerbrey model for rigid films. For example, the thickness of the absorbed polymer layer estimated by using Voigt model was 5 times higher than that estimated by using the Sauerbrey equation, as shown in Figure 4 right. Figure 5 shows the adsorbed mass and the thickness of the polyampholyte layers for silica (left) and cellulose (right) surfaces at five salt concentrations. Higher values of adsorbed mass and thickness were observed at intermediate salt concentrations. It is noted that the observed

changes of dissipation of the adsorbed polyampholyte layers were comparatively large, which suggests a large amount of coupled water.

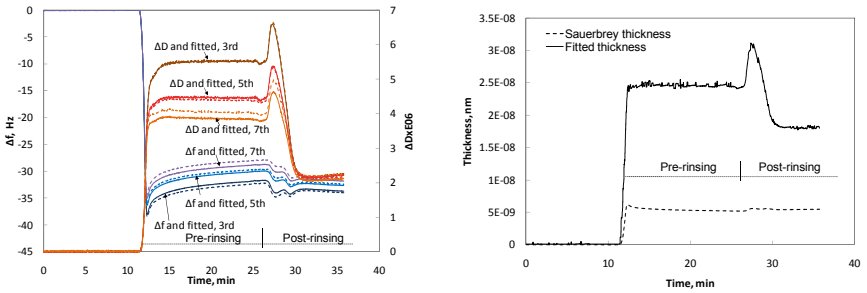


Figure 2 – Comparison between experimental and fitted results by using Voigt model (left), and thickness results by using Sauerbrey Equation and Voigt Model (right). Both graphics were for polyampholyte layer adsorbed on silica surface at pH 4.3 and 10 mM [NaCl].

At low electrolyte concentration, the polyelectrolyte adsorption can be attributed to an ion-exchange mechanism that happens in the electrostatic double layer (Rojas et al., 2001; Rojas, 2002). This mechanism is thermodynamically favored due to the net gain of entropy due to release of counter-ions at the interphase. At high electrolyte concentration, the electrostatic interaction between the polyelectrolyte and the charged surface is reduced. Higher fitted values for polymer adsorbed amount and thickness at the highest salt concentration (1000 mM) on the cellulose surfaces were not expected (Figure 5 B). These results can be explained by the large variability of the results recorded in QCM trials for this salt concentration condition that was computed in the fitting analysis.

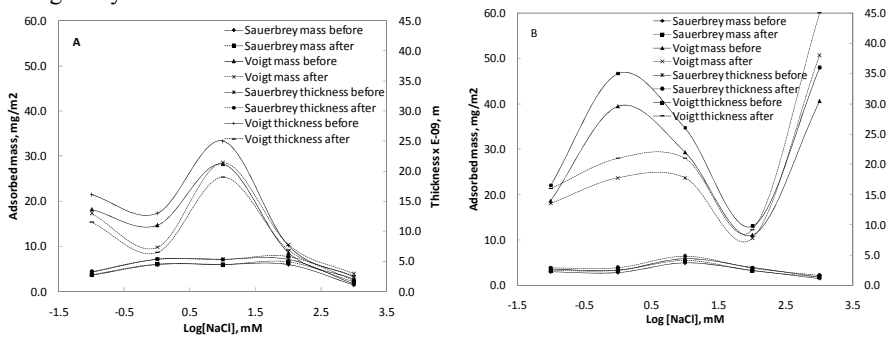


Figure 3 - Results of adsorbed amount of mass and thickness of the polyampholyte layer for silica (A) and cellulose (B) model surfaces.

Surface and polymer charge densities, and ionic strength of adsorbing solution, are factors that affect the interactions forces at the solid - liquid interface (Rojas, 2002). The conformation of these polymers in bulk solution also affects the characteristics of the adsorbed layer, its rigidity and viscoelasticity. Although we expected that these parameters increase with salt concentration (as well as the dynamic light scattering and turbidity, not presented here), we show that at higher salt concentration the electrostatic screening do not allow effective interactions between the polymer and the surface.

4. Conclusion

In the present work we evaluated the properties of adsorbed polyampholyte layers under different salt concentrations. Calculation were performed following the Sauerbrey Equation and a viscoelastic model, for both silica and cellulose surfaces. According to the results, the following conclusions can be drawn: (1) Larger polymer adsorbed mass and thickness were obtained with the Voigt model; (2) The Voigt model showed larger effects of the variations from changes in the ionic strength of adsorbing solution that otherwise is not accounted by Sauerbrey approach; (3) Since the Voigt model uses multiple frequency and dissipation overtones, it was found that the calculated adsorbed layer mass is closer to the actual mass while the Sauerbrey approach underestimates the respective values. Such underestimation is explained by the fact that the adsorbed polyampholyte layers are soft and dissipative.

References

Bohidar, H. B., Ed. (2002). Characterization of polyelectrolytes by dynamic laser light scattering. Handbook of Polyelectrolytes and their Applications. California, USA, American Scientific Publishers.

Garg, A., Heflin, J. R., Gibson, H. W., Davis, R. M. (2008). "Study of film structure and adsorption kinetics of polyelectrolyte multilayer fims: effects of pH and polymer concentration." Langmuir **24**: 10887-10894.

Gunnars, S., Wågberg, L., Cohen Stuart, M.A. (2002). "Model films of cellulose: I. Method development and initial results." Cellulose **9**: 239–249.

Höök, F., Vörös, J., Rodahl, M., Kurrat, R., Böni, P., Ramsden, J.J., Textor, M., Spencer, N.D., Tengvall, P., Gold, J., Kasemo, B. (2002). "A comparative study of protein adsorption on titanium oxide surfaces using in situ ellipsometry, optical waveguide lightmode spectroscopy, and quartz crystal microbalance/dissipation." Colloids and Surfaces B: Biointerfaces **24**: 155-170.

Munro, J. C., Frank, C. W. (2004). "Polyacrylamide Adsorption from Aqueous Solutions on Gold and Silver Surfaces Monitored by the Quartz Crystal Microbalance." Macromolecules **37**(3): 925-938.

Rodahl, M., Kasemo, B. (1996). "Frequency and dissipation-factor responses to localized liquid deposits on a QCM electrode." Sensors and Actuators A-Physical **54**(1-3): 448-456.

Rojas, O. J. (2002). Adsorption of polyelectrolytes on mica Encyclopedia of Surface and Colloid Science. A. Hubbard. New York, Marcel Dekker**:** 517-535.

Rojas, O. J., Neuman, R.D., Claesson, P.M. (2001). "Desorption of low-charge-density polyelectrolyte adlayers in aqueous sodium n-dodecyl sulfate solution." Journal of Colloid and Interface Science **237**: 104-111.

Sauerbrey, G. Z. (1959). "The use of quartz oscillators for weighing thin layers and for microweighing." Zeitschrift Fur Physik **155**(2): 206-222.

Silva, D. J., Hubbe, M. A., Park, S. W., Rojas, O. J. (2008). Adsorption and viscoelastic properties of polyampholytes monitored by QCM-D. V CONGRESO IBEROAMERICANO DE INVESTIGACION EN CELULOSA Y PAPEL 2008. Mexico.

Silva, D. J., Yamaguchi, T., Song, J., Park, S.W., Hubbe, M., Rojas, O. (2008). Swelling and water-holding ability of adsorbed polyampholytes. 235th ACS National Meeting an Exposition, New Orleans, LA.

Tammelin, T., Merta, J., Johansson, L.-S., Stenius, P. (2004). "Viscoelastic properties of cationic starch adsorbed on quartz studied by QCM-D." Langmuir **20**(25): 10900-10909.

Vogt, B. D., Lin, E. K., Wu, W.-L., White, C. C. (2004). "Effect of Film Thickness on the Validity of the Sauerbrey Equation for Hydrated Polyelectrolyte Films." J. Phys. Chem. B **108**: 12685-12690.

Voinova, M. V., Rodahl, M., Jonson, M., Kasemo, B. (1999). "Viscoelastic acoustic response of layered polymer films at fluid-solid interfaces: Continuum mechanics approach." Physica Scripta **59**(5): 391-396.

Wang, Y., Hubbe, M., Sezaki, T., Wang, X., Rojas, O., Argyropoulos, D. (2002). "The role of polyampholyte charge density on its interactions with cellulose." Nordic Pulp & Paper Research Journal, **21**(5): 158-165.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimization of Scaffolds in Alginate for Biofabrication by Genetic Algorithms

Rodrigo Rezende[a], Mylene Rezende[a], Paulo Bártolo[b], Ausenda Mendes[b], Rubens Maciel Filho[a]

[a]*University of Campinas, P.O. Box 6066, Campinas, 13083-970, Brazil*
[b] *Centre for Rapid and Sustainable Product Development, Polytechnic Institute of Leiria, Portugal*

Abstract

With an increasing in the rate of transplants due to damaged or affected tissues or organs by accidents or diseases and also by the aging of the population in many countries as Brazil, have motivated the research of some novel and alternative ways focused on restoring and replacing tissues. Biofabrication by means of Rapid Prototyping techniques can help in the fashioning and final production of scaffolds devoted to support and stimulate the growth of new tissues. For soft tissues, a biomaterial known as Alginate has been studied and used as raw-material for scaffolds fabrication. A scaffold must guarantee good strength and stiffness at the same time the material degrades gradually. In this work, a single mathematical model experimentally obtained that describes an interesting mechanical behavior of the degradation of alginated-scaffolds is developed. The optimization process scheme using Genetic Algorithms to maximize the elastic modulus and therefore to aid the design of scaffolds in alginate is proposed. The optimization is very welcome to tissue engineering and Biofabrication.

Keywords: Genetic Algorithms (GAs), Scaffolds, Biofabrication, Tissue Engineering, Alginate.

1. Introduction

A Brazil's recent picture of the 2008 waiting list for transplants, according to the Ministry of Health, reveals that there are more than 68,000 candidates waiting the availability of organs to therefore undergo a transplant. Based on those data, alternative methods of tissue and organ recovering have been studied and many developments have been proposed and applied successfully in the Tissue Engineering. Tissue Engineering as an interdisciplinary field combines the use of living cells with either natural or synthetic extra-cellular structures (scaffolds) to develop body parts or devices that will enable the restoration, maintenance or enhancement of living tissue and organs (Rezende et al., 2007a). Usually, the scaffolds have high porosity (macro-porosity), appropriate surface morphology (micro-porosity), large surface area, suitable pore size and highly connected pore structure. They must also be biocompatible and biodegradable. A variety of biodegradable and biocompatible hydrogels, as the alginate, have been used for Tissue Engineering. The alginate is one of the most popular materials due to its relatively low cost, natural origin and easy handling besides other physical properties advantages. Biofabrication is a new class of Rapid Prototyping that

represents a group of non-conventional techniques with great potential to produce scaffolds with customised external shape and predefined internal morphology (Leong et al., 2003). These techniques also allow controlling both pore size and distribution. A scaffold must own very dynamical and adaptive characteristics in order to be implanted and to take its main roles which are to carry the stem live cells inside it, to back the growth of these cells and besides this to biodegradate appropriately since the minimum material should remain after the tissue is reconstructed. In this sense, it is fundamental to be aware of the mechanical and chemical properties since the scaffold must guarantee good strength and stiffness at the same time the material degrades gradually. To know how the mechanical behavior of the scaffold will be, some time later, is the keyword. And the understanding about the match between biodegradation and young modulus is mandatory. The present and future of biomedical materials development requires this degree of control prediction in the design, synthesis, and function of next-generation materials. A prediction job is possible and it has already been used so that the scaffold state can be forecasted before its fabrication and, as a good alternative, to know how and how much alginate should be used. Other future analyses can be around the best geometry to be adopted during rapid prototyping technique actuation. This paper presents a single mathematical model experimentally obtained that describes an interesting mechanical behavior of the degradation of alginated-scaffolds. The deal of this work is the optimization of scaffolds in alginate making use of Genetic Algorithms (GAs). GAs represent a class of stochastic optimization procedures based on natural systems according to Darwin's observations. GAs have been successfully applied to a range of problems and have characteristics of easiness of implementation and capability of escaping local optimal solution. The objective of GA is to find out the best values of alginate amount for scaffold fabrication that maximize the elastic modulus.

2. Optimization Problem Formulation

The aim of the optimization is to determine optimal features for the fabrication of optimized alginate scaffolds for Tissue Engineering. The optimization objective in this work is to find out optimal values of alginate composition and initial porosity in order to fabricate scaffolds with, at a pre-determined time, a high mechanical behavior (elastic modulus), since those parameters are two of the most important ones and they need to be well combined in order to reach the best value for the elastic modulus. Then, the optimization problem can be written as:

$$\begin{aligned} &\underset{[\alpha,\phi_0]}{\text{Maximize}} && E(\phi_0,\alpha,t) \\ &\text{Subject to:} && 1\% \le \alpha \le 8\% \\ & && 30\% \le \phi_0 \le 80\% \end{aligned} \qquad (1)$$

where E is the elastic modulus (shear effects are not considered), a is the alginate composition and f_0 is the initial porosity.

The variation of the mechanical properties of alginate is due to degradation and porosity changes. The degradation of alginate structures was determined through the analysis of the shrinkage variation along time as shown in Figure 1 (Rezende et al., 2007a):

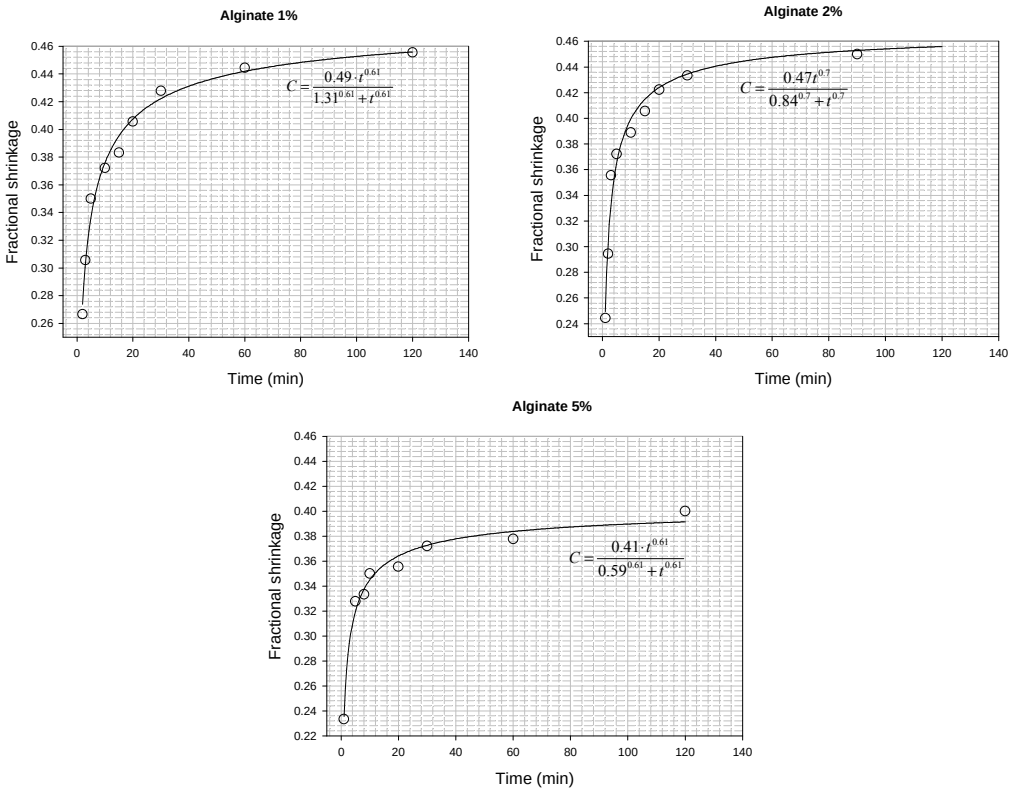


Figure 1 – Fractional shrinkage along time for different alginate compositions (1%, 2%, 5%).

There are various studies modeling tissue scaffolds behaviors in the literature (Nair et al., 2008). In the present case, a sigmoidal model of three parameters represents the shrinkage process that is given by the following equation:

$$C(\alpha, t) = \frac{\varsigma(\alpha) \cdot t^{\vartheta(\alpha)}}{\lambda^{\vartheta(\alpha)} + t^{\vartheta(\alpha)}} \tag{2}$$

where t is the time and $\varsigma, \vartheta, \lambda$ are variables that depend on the alginate composition (a). Porosity at each time is also a function of alginate composition and shrinkage:

$$\phi(\phi_0, \alpha, t) = \phi_0 + \zeta(\phi_0, \alpha) \cdot C(\phi_0, \alpha, t) + \psi(\phi_0, \alpha) \cdot C^2(\phi_0, \alpha, t) \tag{3}$$

where ζ, ψ are constants depending on alginate composition and C is the shrinkage.

The dependence between the elastic modulus and porosity for different alginate compositions is given by the following equation:

$$E(\phi_0, \alpha, t) = E_0(\phi_0, \alpha) + k_1(\phi_0, \alpha) \cdot \phi(\phi_0, \alpha, t) + k_2(\phi_0, \alpha) \cdot \phi(\phi_0, \alpha, t)^2 + k_3(\phi_0, \alpha) \cdot \phi(\phi_0, \alpha, t)^3 \tag{4}$$

with E_0 being the initial elastic modulus, k_1, k_2, k_3 constants dependent on both the alginate composition and the initial porosity and ϕ the final porosity of the scaffold.

The optimization problem presented on Equation 1 is a single constrained optimization problem. In this paper, two cases of constraints are considered: constraints at shrinkage and final porosity: 1) shrinkage higher than 25% and 2) final porosity higher than 80%.

In order to solve the constrained optimization problem, a constraint handling method based on the penalty function approach was used, not requiring any penalty parameter (Rezende et al., 2007b). In this case, the expression of the fitness function for a minimization problem, where infeasible solutions are compared based only on their constraint violation, is given by the Equation 5:

$$F(\mathbf{x}) = \begin{cases} f(\mathbf{x}) & \text{if } g_j(\mathbf{x}) \geq 0 \quad \forall j=1,2,...,nc \\ f_{\max} + \sum_{j=1}^{m} \langle g_j(\mathbf{x}) \rangle & \text{otherwise} \end{cases} \tag{5}$$

where $f_{\max}$ is the objective function value of the worst feasible solution in the population.

3. Optimization of Scaffolds using GAs

The Genetic Algorithms approach starts with a random population of chromosomes that are a set of solutions for the optimization problem. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration. Usually, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population.
The Genetic Algorithm used in this research work to solve the scaffolds optimization problem is a Fortran binary code. The employed genetic operators are the tournament selection, the uniform crossover, the creep and the jump mutation. Niching and elitism are also employed. The input parameters, chosen by a trial and error method, are indicated in Table 1.

Table 1 - The GA input parameters.

GA input parameters	Value
Population size per generation	50
Maximum number of generations	30
Crossover probability	0.60
Jump mutation probability	0.077
Creep mutation probability	0.077
Initial random number seed for the GA run	-1000

The following results of the scaffolds optimization using Genetic Algorithms for constrained problem considering two cases of constraints are presented.

3.1. Constraint 1: Shrinkage > 25%

In this case, the objective of the optimization problem is to maximize the elastic modulus subject to a shrinkage higher than 25% and 2) final porosity higher than 80%.
The results obtained for this case are shown in Table 2:

Table 2 – Optimization results for the constrained problem (shrinkage > 25%)

Optimization Variables	Initial Alginate composition (%)	7.06
	Initial Porosity (%)	30.00
Objective Function	Elastic modulus (KPa)	17.52
Constraint	Shrinkage (%)	25.22
Output Variable	Final Porosity (%)	70.97

Figure 2 shows the evolution of the objective function along all the generations. Figure 3 presents the obtained profile of the objective function from the best combination of the optimization variables at each generation. The output variables (shrinkage and final porosity) dependent on the best combination are also indicated.

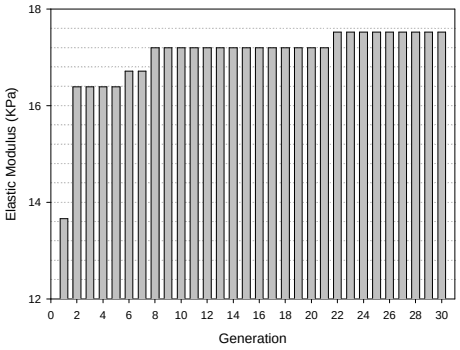


Figure 2 – Evolution of the elastic modulus along all the generations (shrinkage > 25%).

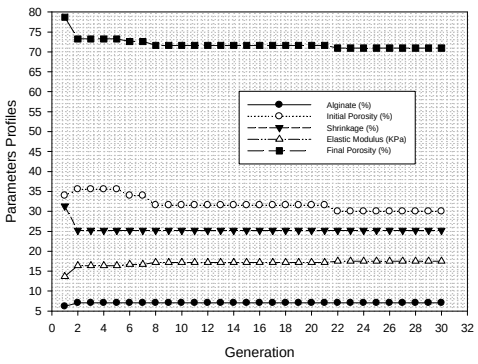


Figure 3 – Profiles of the objective function, shrinkage and final porosity obtained with the best values of the optimization variables at each generation (shrinkage > 25%).

3.2. Constraint 2: Final porosity > 80%

In this case, the objective of the optimization problem is to maximize the elastic modulus subject to a final porosity higher than 80%.

Results obtained for this case are shown in Table 3:

Table 3 – Optimization results for the constrained problem (final porosity > 80%).

Optimization Variables	Initial Alginate composition (%)	5.79
	Initial Porosity (%)	32.38
Objective Function	Elastic modulus (KPa)	12.99
Output Variable	Final Porosity (%)	80.01
Constraint	Shrinkage (%)	33.29

Figure 4 shows the evolution of the objective function along all the generations. Figure 5 presents the obtained profile of the objective function from the best combination of the optimization variables at each generation. The output variables (shrinkage and final porosity) dependent on the best combination are also indicated.

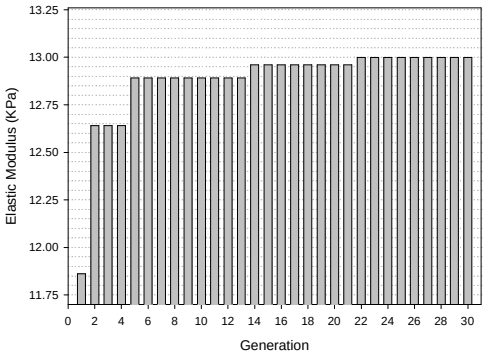


Figure 4 – Evolution of the elastic modulus along all the generations (final porosity > 80%).

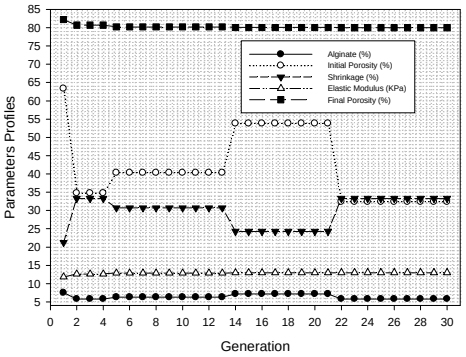


Figure 5 – Profiles of the objective function, shrinkage and final porosity obtained with the best values of the optimization variables at each generation (final porosity > 80%).

Observing the two cases presented, it can be seen that the constraints were respected. In spite of the obtained value of young modulus, the order of this value is compatible with results found in the literature (Khalil, S.E.D., 2005). Throughout this study is incipient and new boundaries and adjustments need to be done. One of the main characteristics of the alginate-scaffold is to degradate along the time where the growing of final porosity indicates this phenomenon. The Initial Alginate composition (a) is the real range that has been used experimentally.

4. Conclusions

This research employs Genetic Algorithms to optimize the mechanical behavior of alginate scaffolds for Biofabrication. The mathematical model was experimentally obtained and the best values for both alginate composition and initial porosity of the scaffold allowing the constrained maximization of the elastic modulus were determined through the optimization by Genetic Algorithms.

References

R.A. Rezende, P.J. Bartolo, A. Mendes, H. Almeida, R. Maciel Filho, 2007a, Experimental Characterisation of the Alginate Gelation Process for Rapid Prototyping. The eighth International Conference on Chemical & Process Engineering, v. 11, p. 509-514.

K.F. Leong, C.M.Cheah, C-K.Chua, 2003, Solid freeform fabrication of three-dimensional scaffolds for engineering replacement tissues and organs. Biomaterials, 24, 2363-2378.

Nair, K., Yan, K. and Sun, W. 2008. A computational modeling approach for the characterization of mechanical properties of 3D alginate tissue scaffolds, Journal of Applied Biomaterials and Biomechanics, 6: 35 – 46.

M.C.A.F. Rezende, R.A. Rezende, A. Mendes, P.J. Bartolo, A.C. Costa, R. Maciel Filho, 2007b, Evolutionary Optimization for Soft Tissue Engineering. Advanced Research in Virtual and Rapid Manufacturing, v.3 , p. 81-85.

Khalil, S.E.D. 2005. Deposition and Structural Formation of 3D Alginate Tissue Scaffolds. Drexel University. (PhD Thesis).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Risk Analysis and Environmental Impact Analysis in a Chemical Processing Facility

Maristhela P. de A. Marin[a], Elias Basile Tambourgi[b]

[a]*Departmentof Chemical Engineering, Centro Universitério da FEI, Av. Humberto de A. C. Branco 3972, São Bernardo do Campo 09850-901, Brasil*
[b]*Faculdade de Engenharia Química,UNICAMP, Av. Albert Einstein 500, Campinas 13083-852, Brasil*

Abstract

The present work provides a method of risk analysis in chemical processing facilities, which consists primarily of the following steps: description and study of the system, hazard analysis and identification, identification of the accidental scenarios, calculation of the consequences of the scenarios and risk characterization. To illustrate the application of the method, an uranium hexafluoride production facility was selected. The main hazards identified were: fire, explosion, spill of toxic chemical substances and ground contamination due to problems in the waterproof blanket of the waste contention bund. Some cases of environmental impact due to uranium hexafluoride (UF_6) release were quantitatively analyzed for different scenarios, considering two categories of atmospheric stability: Pasquill D and Pasquill F. It was found that the impact which reached a great external area around the plant corresponded to the condition of atmospheric stability F.

Keywords: Risk Analysis, Environmental Impact, Consequences Analysis.

1. Introduction

The industrial accidents in the past few years, particularly in the 80's, combined with environmental problems, contributed significantly to gather the attention of government authorities, industry and society, to seek means for the prevention of such episodes that damage people's safety and the quality of the environment.

Currently, there is a strong trend for quantitative analysis of risks, which is considered to be a useful tool in identifying hazards, to be held more frequently on the parts of industrial systems considered most critical with risks of accidents with environmental consequences.

The risk analysis consists on the systematic examination of industry facilities in order to identify hazards and quantify the consequences of an accidental event.

The method of risk analysis proposed in this work is similar to that proposed in World Bank (1985), specifically for chemical processing plants and applicable, with some changes, in nuclear facilities. It was applied to a uranium hexafluoride (UF_6) production plant. In the present work the main hazards were identified and the consequences of accidental release of UF_6 were quantitatively analyzed and the calculations were made using the software Conv_aci (Molnary, 1993).

2. Method

The plant is divided into functional units and for each one the systems and the components of interest are identified. Each component is analyzed individually and therefore determined its possible modes of failure. The inventory of hazardous materials are selected and grouped, being analyzed quantitatively only those with the largest inventories in terms of toxicity, flammability, and quantities. The hazards are identified, the cases of release are classified and the source terms are determined, following the evaluation of consequences. If probability data is known, one can estimate the frequency of undesirable events. Finally, mitigating risks measures are proposed.

An historical analysis of accidents at similar facilities is considered a part of risks analysis. This research aims to identify the hazards in industrial activities, determine the common causes of accidents, raise historical trends and probabilities of occurrence of accidents, assess the results of calculation models simulation and to assist in developing the event trees.

2.1. Case study

The plant chosen for the consequence analysis was Kerr Mc Gee Sequoyah Hexafluoride Plant. This installation purified the concentrated uranium, yellowcake, through the process of extraction by solvent for subsequent conversion into uranium hexafluoride (UF_6). The report (NRC, 1975) describes both plant and chemical processes in detail which allows their analysis. For the application of the method, the plant was divided into four functional units*:*

- Area I: UF_6 production
- Area II: Elemental fluorine production
- Area III: Chemicals storage
- Area IV: Waste contention basins

For all areas components, modes of failure and inventory of materials were identified.

2.1.1. Theoretical background

The assessment of consequences of toxic substances releases requires knowledge of their physicochemical properties and their toxicity, the adoption of a model of atmospheric dispersion and the establishment of criteria for the classification of undesirable events. Although the UF_6 is a stable compound, its manipulation is complex, since the substance is highly reactive with water, ether and alcohol, forming stable products. The reaction of UF_6 with water is highly exothermic and its hydrolysis products are uranyl fluoride (UO_2F_2) and hydrofluoric acid (HF) according to the reaction:

$$UF_{6(g)} + 2H_2O_{(g)} \rightarrow UO_2F_{2(s)} + 4HF_{(g)} \quad (1)$$

Of its products of hydrolysis, the HF presents only chemical toxicity while UO_2F_2 has chemical and radiological toxicity. The uranium chemical toxicity surpasses the radiological at any enrichment level (NRC, 1991).

In order to analyze the consequences of UF_6 release, it was considered that the release occurs at ground level, that the individual exposed to the effects of the released substance remains in the center line of the plume during the dispersion and that the particles are highly respirable and transportable. The incorporation of uranium via inhalation is calculated using the expression (NRC, 1986):

$$IU=(\frac{\chi}{Q}).Mu.BR.f_d \quad (2)$$

Where:

IU is the mass of radioactive substance incorporated by inhalation (mg); $\frac{\chi}{Q}$ represents factor of atmospherical dispersion depending on the distance (s/m^3); *BR* is the breathing rate of an average individual, as 2,66 x10^{-4} m^3/s; *Mu* is the mass of released uranium (mg); *fd* represents factor of dry deposition correction.

It is considered that the particles resulting from the release of UF_6 have about 10^{-6} m of diameter, and therefore, are fully inhaled.

Concentration of HF:
The concentration of HF (C_{HF}) in air is given by the expression (NRC, 1986):

$$C_{HF}=(\frac{MHF}{t})(\frac{\chi}{Q}) \quad (3)$$

Where:
MHF is the mass of HF formed in the hydrolysis of released UF_6; *t* represents time of release (s)

Criteria for classification of the event for the substance release:
To classify the event of accidental release of UF_6, two different criteria were used. The first one, proposed by Hanamann et al. (1995) is a quantitative method to classify the impacts on human health shown in Table 1. The groups are defined in terms of tolerance limits on exposure to soluble uranium (*Iu*) and the concentration of hydrofluoric acid gas (C_{HF}), as shown in Table 2. The following indices for the tolerance limits are: *Emergency Response Planning Guideline* (ERPG), *Threshold Limit Value - Short Term Exposure Limit* (TLV-STEL), *Threshold Limit Value - Ceiling* (TLV-C) and *Immediately Dangerous to Life or Health* (IDLH).

Table 1**:** Classification of the event according to the toxicological consequence

Category	Toxicological Consequence
Despicable	1. Exposition of public inferior than GROUP 3
	2. Exposition on property limits inferior than GROUP 2
	3. Exposition on event location inferior than GROUP 1
Marginal	1. Exposition of public greater than GROUP 3
	2. Exposition on property limits greater than GROUP 2
	3. Exposition on event location greater than GROUP 1
Critical	1. Exposition of public greater than GROUP 2
	2. Exposition on property limits greater than GROUP 21
Catastrophic	1. Exposition of public greater than GROUP 1

Table 2: Values of tolerance limits that define the groups

	Soluble Uranium (*Iu*)	Concentration of HF (C_{HF})
GROUP 3	TLV-STEL: 0,6 mg/m^3	ERPG 1: 4,1 mg/m^3
GROUP 2	TLV-C: 1,0 mg/m^3	ERPG 2: 16,4 mg/m^3
GROUP 1	IDLH: 1,0 mg (incorporation)	ERPG 3: 25 mg/m^3

The second criterion used for classification of UF_6 release is proposed in the report NUREG 1391 (NRC, 1991). In this criterion the event is classified as abnormal or accident due on the individual is exposure to soluble uranium (*Iu*) and to the concentration of HF (C_{HF}) as shown in Table 3, referenced on the IDLH.

Table 3: Criteria for classification of the event as abnormal or accident

Classification	C_{HF} (mg/m^3)	*Iu* (mg)
Abnormal Event	< 25 (1800/t)^2	2<Iu<10
Accident	> 25 (1800/t)^2	Iu ≥ 10

3. Results and Discussion

3.1. Major undesirable events

Depending on the method application, the following hazards were identified: fire in the area of extraction by solvent, and UF_6 release from the area of fluorination and collection on Area I; explosion in the fluoride cells in Area II; rupture of substance storage tanks, as anhydrous hydrofluoric acid (HF), anhydrous ammonia (NH_3) and nitric acid (HNO_3), which are used in the process; failure on the sealing blanket of the waste contention basins in Area IV.

This study quantitatively analyzed only the case of UF_6 release and consequently release of HF gas and solid UO_2F_2 as depicted in equation (1).

3.2. UF_6 release

The event analyzed is characterized by the release of liquid UF_6 from the product withdrawal area. It is considered that during the transfer operation of liquid UF_6 from the primary cold trap to the storage cylinder, a rupture of a transport pipeline occurs. It is admitted that all of the liquid UF_6 contained in the cold trap (9492 kg) was released to the process building during 900 seconds (15 min). As a result of the release, some of the product will be solidified and some vaporized. In the plant operating conditions, the vaporized fraction of liquid UF_6 is 49% (Siman-Tov, 1984).

Three scenarios (C) were studied:

• C1: Due to a failure in the ventilation system, the entire amount of UF_6 vaporized (4651 kg, equivalent to 3139 kg of uranium and 1057 kg of HF) is released into the atmosphere, at ground level, during 900 s (15 min).

• C2: The ventilation system failed, but of the total amount of UF_6 vaporized (4651 kg), at ground level, only a portion of the product that reacts with air is released. It is admitted that 80% of HF (846 kg) and 50% of uranium (1560 kg) are released, the rest being confined within the facility.

• C3: Credit is given to the exhaust system, and it is assumed that 90% of the UF_6 vaporized is retained by the filters and gas scrubbers. Therefore, at ground level, it is released to air 314 kilograms of uranium and 106 kg of HF.

According to the distance from the release point, the incorporation of uranium (*Iu*) and the concentration of HF (C_{HF}) were calculated, considering two classes of atmospheric stability: Pasquill D (neutral atmospheric stability with wind speed equal to 3m/s) and Pasquill F (stable atmospheric stability with wind speed equal to 1m/s) using the software Conv_aci. For event classification, the following areas were defined: Area 1: internal dependencies of the building, where the health impacts of workers of the plant are analyzed; Area 2: within the site, that is, the area bounded by the limits of building

and by the facility property limits, where the health impacts of workers of the plant are analyzed, Area 3: area outside the limits of the facility, where the health impacts on local individuals are analyzed.

It was considered that the limit of the property is located 1175 m from the point of release. Table 4 shows the distance in meters, from which C_{HF} becomes less than 35 mg/m^3, the tolerance limit given by the IDLH for 900 second of leaking and the distance from which *Iu* is less than 10 mg.

Table 4: Results of UF_6 release

	Distance (mts) for C_{HF}≤35mg/m^3		Distance (mts) for *Iu*≤10mg	
	Class D	Class F	Class D	Class F
C1	1600	7600	1400	1500
C2	Property Limits	6700	Property Limits	1300
C3	350	1600	500	960

Scenario C1

According to the criterion 1, the toxicological consequences are classified catastrophic for both class D and class F of atmospheric stability, as C_{HF} and *Iu* to the individual of the public (Area 3) surpass the values of GROUP 1, as indicated on both Tables 1 and 2. Considering the classification by the criterion 2, it is observed that for the Pasquill D stability class, both C_{HF} and *Iu* are below the reference levels from 1600 m and 1400 m from the release point, which classifies the case as an accident. Considering the atmospheric stability condition of Pasquill F, the values of C_{HF} and *Iu* are smaller than the reference levels from 7600 m and 1500 m of the release point, what also classifies the event as an accident.

Scenario C2

According to criterion 1 for Class D of atmospheric stability, the event is classified as critical, because C_{HF} and *Iu* for individuals of the public (Area 3) exceed the values of GROUP 2 and at location (Area2) the values of C_{HF} and *Iu* surpass the values of GROUP 1. For class F of atmospheric stability, the event is classified as catastrophic, once C_{HF} and *Iu* for individuals of the public (Area 3) exceed the values of GROUP 1. Regarding the criterion 2, for the condition of atmospheric stability Pasquill D, near the limit of the property, C_{HF} is below 35 mg/m^3 and *Iu* is below 10 mg, likely to be classified as an abnormal event. In condition F, the results characterize accident, because C_{HF} is less than 35 mg/m^3 from 6700 m and *Iu* is below 10 mg from 1300 m to the point of release.

Scenario C3

In relation to criterion 1, it is observed that the toxicological consequences due to exposure to C_{HF} and *Iu* are classified as: marginal for Class D of atmospheric stability, as C_{HF} and *Iu* on the leakage area (Area 1) are higher than the value of GROUP 1; critical for the class F of atmospheric stability, because C_{HF} and *Iu* on the locality (Area 2) is higher than in GROUP 1. Considering the classification criterion 2, the results indicate abnormal event for the Class D of atmospheric stability, as C_{HF} is less than 35 mg/m^3 from 350 m and *Iu* less than 10 mg from 500 m. In class F of atmospheric stability, there are attained levels of more than C_{HF} 35 mg/m^3 at distances up to 1600 m, characterizing accident. *Iu* is less than 10 mg to 960 m from the point of release.

4. Conclusions

The broadest criterion for consequences evaluation on human health due to the event of accidental release of UF_6, as: despicable, marginal, critical or catastrophic, is the widest because it involves the evaluation of consequences in three different areas: the area of occurrence of leakage (Area 1), the location (Area 2) and the area outside the limits of the property (Area 3). The toxicological consequences for scenario C1, Class D of atmospheric stability and scenarios C1 and C2, considering the class F of atmospheric stability, are classified as catastrophic, once they represent situations where individuals of the public are subject to C_{HF} or *Iu* exceeding the values in GROUP 1. It should be noted, however, that the scenario C1 is extremely conservative, as it does not take into consideration the performance of the ventilation system filters, disregards the confinement exerted by the building itself and does not include the phenomena of elevation of the plume and deposition of particles of UO_2F_2 on the floor of the building. If these factors are added with the low rates of failure of UF_6 pipes and filters, plus the fact that operators of the facility could serve to isolate the area affected by the leakage, the C1 scenario can be considered not credible. It follows, therefore, that the risk of C1 associated with individuals of the public is almost negligible.

The second criterion used to classify events of accidental release of UF_6, only classifies the toxicological consequences on the health of individuals from the public. It has only two categories in the classification of the consequences: abnormal event or accident. It is verified therefore, that the worst conditions of release correspond to the F condition of atmospheric stability, since for all examined cases the consequences are classified as accidents. The scenario C3 illustrates how the adoption of impact attenuators measures (such as the performance of the ventilation system filters) may benefit the plant and reduce the risks to individuals from the public. From the obtained results, it can be observed that the criteria 1 and 2 are coherent, as there are equivalences between the catastrophic or critical categories on criterion 1 and the accident classification on criterion 2.

5. Acknowledgements

The authors acknowledge and thank the institutional support offered by Centro Universitário da FEI.

6. References

Hannaman, G. W., Kryskas, P., Mahan, J., 1995. Qualitative Methods for Assessing Risk. ASME, USA.

Molnary, L.,1993. Caracterização de um modelo de camada limite planetária para avaliar liberações de radionuclídeos em instalações nucleares. Dissertação (Mestrado)–IPEN. São Paulo: Universidade de São Paulo.

Nuclear Regulatory Commission, 1975. Sequoyah Uranium Hexafluoride Plant: Final Environmental Statement; NUREG 75 007.

Nuclear Regulatory Commission, 1986. Assessment of the Public Health Impact from the Accidental Release of UF_6 at the Sequoyah Hexafluoride Fuel Corporation Facility at Gore; NUREG 1189, v. 1.

Nuclear Regulatory Commission, 1991. Chemical Toxicity of Uranium Hexafluoride Compared to Acute Effects of Radiation, NUREG 1391.

Siman-Tov, M. et al., 1984. Scenarios an Analytical Methods for UF_6 Releases at NRC Licensed Fuel Cycle Facilities. Oak Ridge, TN, USA, NUREG /CR 3139.

World Bank, 1985. Manual of Industrial Hazard Assessment Techniques. London.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Modeling of Kinetics of Water Droplets Coalescence in Crude Oil Emulsion Subjected to an Electrical Field.

Antonio E. Bresciani[a], Candido F. X. de Mendonça[b], Rita M. B. Alves[a] and Claudio A. O. Nascimento[a]

[a]*LSCP/CESQ - Department of Chemical Engineering - Polytechnic School – University of São Paulo, São Paulo, SP, Brazil.*
[b] *Escola de Artes, Ciências e Humanidades - University of São Paulo, São Paulo, SP, Brazil.*

Abstract

Water is used in petroleum desalting units to dilute and remove the salted water droplets that the crude oil contains. The basic processes promote the coalescence of small droplets of conducting water dispersed in a crude oil emulsion. In order to make separation easier, the emulsion is distributed horizontally between two electrodes and subjected to an electrical field, which generates an attractive force among the droplets, promoting coalescence phenomena and further sedimentation. The main purpose of this study is to reduce the demand of fresh water and the liquid effluent generation in refineries. This paper presents a new model developed in order to calculate the droplets velocity by using the balance of the acting forces. The model is able to determine the droplets trajectory in order to define if they can be separated from the continuous phase. Besides the deterministic approaches based on traditional equations, the model uses also the concept of cellular automata. Thus it is possible to solve the problem in a stochastic way and to show visually the sequence of droplets collisions and coalescence phenomena. This methodology enables to calculate the amount of water that can be separated of the emulsion for a number of different operating conditions and then to optimize the process. Comparisons between the obtained results by the developed model and the operational performance of a real desalting unit are carried out. A good accuracy is observed, which shows that the real process is very well represented by the developed model.

Keywords: water use optimization, emulsion, coalescence, automata cell

1. Introduction

Studies in the optimization of water use in industrial process plants involve profound knowledge of all processes that use water (Bagajewicz, 2000). It is necessary to formulate equations to calculate water flow, contaminants concentrations at process input and output and the maximum and minimum limits for these concentrations. Desalting is an important process in petroleum refining that uses water to remove salt that is present in petroleum when it is extracted in oil fields. Water is added to dilute Sodium, Calcium and Magnesium Chlorides. An emulsion is formed, comprising droplets with an average diameter in the range of 10 to 20μm (Bhardwaj, 1994). The water separated from this emulsion removes salt in the petroleum, thus avoiding corrosive wear on the equipment further downstream in the refining process, as well as

residual deposits and catalyst contamination. In order to facilitate the separation, an external electric field is set up which stimulates the agglutination of the smaller droplets into larger droplets that settle and are thus separated from the continuous phase. The industrial petroleum desalting system comprises of two desalting stages, involving the recycling of water, as demonstrated in Figure 1

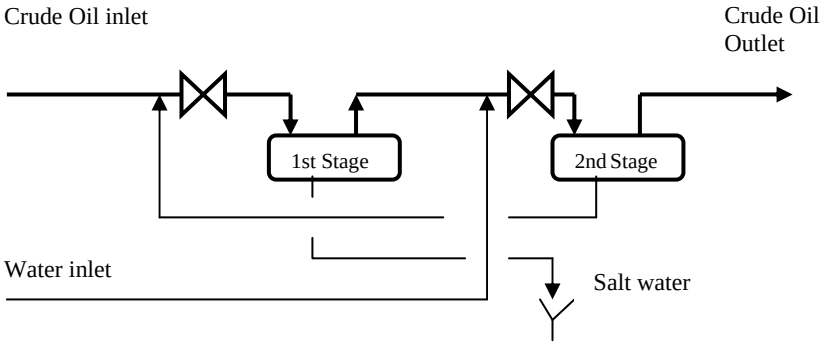


Figure 1. Desalting system employed by petroleum refineries.

These are normally modular, high-speed desalters, to which an external electric field is applied, and where the emulsion is injected into the space between the electrodes. Discharge from the continuous phase is considered to be horizontal in the volume between the electrodes and vertical after it passes by the electrodes, as demonstrated in Figure 2. The small droplets are affected by both the action of the electric field and gravitational force. They then agglutinate and are separated from the continuous phase. The smaller droplets that are not large enough to be decanted remain in the petroleum that exits the desalter.

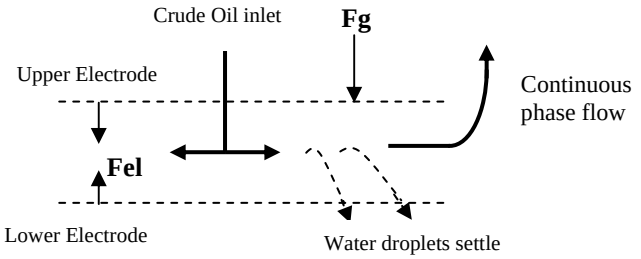


Figure 2. Discharge of the emulsion between the electrodes and forces that act upon the droplets.

The time in which the emulsion remains under the influence of the electric field is calculated by the equation for the speed of the horizontal discharge from the continuous phase, from the moment the emulsion exits the central distributor until it leaves the group of electrodes.

The balance of forces that act upon the droplets leads to the formulation of a mathematical model which enables the calculation of the time between collisions, the speeds immediately before collisions and the trajectory of the droplets.

The operational variables that have the most influence in this process are the operating temperature, the electric power gradient, the quantity of water in the mixture and the density and viscosity of the petroleum. The formulated model enables the study of the influence of these variables on the efficiency of the separation of the mixture, hence permitting process optimization.

The influence of the quality of water in the mixture is studied in laboratory experiments, where the separation time and the quantity of water separated are measured with analytical equipment that ascertains the intensity of light transmitted or refracted by the droplets.

2. Formulation of the mathematical model.

2.1. Horizontal discharge from the continuous phase

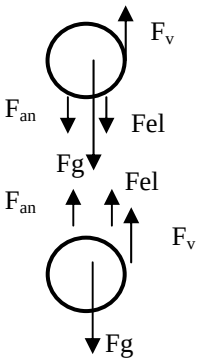
Equation (1) calculates the horizontal speed (v_h) of the continuous phase over the distance (z) from the distribution valve outlet to the desired position, for the flow and dimensions of the desalter studied.

$$v_h = 0.0278/(0.0235 + 0.322\ z + 1.10\ z^2)\ (m/s) \quad (1)$$

The time in which the emulsion is under the influence of the electric field is obtained through the integration of Eq. (1), and the result is 20 seconds.

2.2. Forces acting between droplets

The forces acting between droplets are the electric force (Fel), resulting from the effect of the electric field, gravitational force and viscous force. Natural attraction/repulsion forces (Van der Waals) and the Brownian force are not considered, as their intensities are much smaller than the others under consideration (Eow, 2001). Figure 4 demonstrates how the forces act upon the droplets.



F_{an} : Natural attraction inter-droplets force (van der Walls and repulsion)

F_{el} : Attraction inter-droplets force by electrical field

F_g : Gravity

F_v : Viscous force

Figure 3 Forces acting on the droplets

The balance of forces acting upon each droplet is established, resulting in equation. (2), for the upper droplet (subscript 1), and equation. (3), for the lower droplet (subscript 2).

$$F_g + F_{el} - F_v = m_1\ dv_1/dt = 4/3\ \pi\ a_1^3\ \rho_a\ dv_1/dt \quad (2)$$

$$F_g - F_{el} - F_v = m_2\ dv_2/dt = 4/3\ \pi\ a_2^3\ \rho_a\ dv_2/dt \quad (3)$$

Through the force equations and the parameter values for the oil studied (viscosity, density and dielectric constant), for T=120 ^oC, which can be considered to be the operating temperature of the desalter, we get Eq. (4), for the upper droplet, and Eq. (5), for the lower droplet.

$$1.57 + 0{,}134\ 10^{-3}\ E_0^2\ a_1\ a_2^2\ \ell^4 - 7.3\ 10^{-6}\ a_1^{-2}\ v_1 = dv_1/dt \quad (4)$$

$$1.57 - 0{,}134\ 10^{-3}\ E_0^2\ a_1^2\ a_2\ \ell^4 - 7.3\ 10^{-6}\ a_2^{-2}\ v_2 = dv_2/dt \quad (5)$$

where E_0 is the applied electric power gradient (kV/cm), a_1 and a_2 are the radii of droplets 1 and 2, m and ℓ are the mass and the distance between the droplets.

Equations (5) and (6) can be solved numerically for the pair of droplets of radii a_1 and a_2 under the effect of the electric power gradient E_0. The integration time must be very small ($2\ 10^{-6}$ seconds), requiring a lot of time for computational processing.

Considering that it is an almost stationary regime, with $dv_1/dt = dv_1/dt = zero$, for 120°C, we get equation. (6), for the upper droplet, and equation (7), for the lower droplet.

$$v_1 = 2.15\ 10^5\ a_1^2 + 1.83\ 10^1\ E_0^2\ a_1^3\ a_2^2\ \ell^4 \quad (6)$$

$$v_2 = 2.15\ 10^5\ a_2^2 + 1.83\ 10^1\ E_0^2\ a_1^2\ a_2^3\ \ell^4 \quad (7)$$

Equations (6) and (7) are also solved numerically, with a longer integration time (10^{-3} seconds). Table 1 shows the results of the solution of the systems of equations 4/5 and 6/7 for the temperature of 120°C. The errors resulting from the consideration of an almost stationary state are small, which justifies its use to obtain shorter computation processing times.

Table 1. Time between collisions t_c for each power gradient, 7% water, for 10μm drops.

	Eo = 0.35 kV/cm	Eo = 0.35 kV/cm	Eo = 0.35 kV/cm	Eo = 0.35 kV/cm	Eo = 0.35 kV/cm
t_c (s) (7/8)	1.16	0.57	0.142	0.063	0.035
t_c (s) (num sol 5/6)	1.20	0.59	0.147	0.065	0.037
Error (%)	3.3	3.4	3.4	3.1	5.4
Absolute error (s)	0.04	0.02	0.005	0.002	0.002

2.3. The mathematical model employing the concept of cellular automata.

A model based on the concept of cellular automata was developed, which enables the application of equations of the process being studied to the whole group of droplets, in a stochastic manner, in order to visualize the collision process and optimize the process variables.

The concept of cellular automata was introduced by von Neumann (1970) and consists of a sequence of many cells carrying discreet values arranged in a matrix of these cells and the alteration of these values depends on the values of the variables of the neighboring cells. Wolfram (1983), in his discussion of cellular automata as systems of statistical mechanics, established that physical systems containing many discreet elements with local interactions are conveniently modeled as cellular automata.

The application of cellular automata to more complex systems has only come into use more recently, due to the greater availability of computational resources in recent times. For example, we can cite Ohgai et al (2007), who applied this concept to study the effects of accidents (explosions) on a certain physical environment. The simulation space consists of a group of cells that comprise a geometric solid where the dimensions are parameters. In this case, the height is fixed and corresponds to the application space of the electric power gradient, and is divided into 1143 cells, each with 0.1 mm. Thus, a cell corresponds to a horizontal slice of the geometric solid, as shown in Figure 4. Initially, an emulsion of water in oil, with a distribution of droplet size and the concentration of water defined, is randomly injected into the first 254 cells, which are in front of the nozzle of the distribution valve. The emulsion is subjected to the attraction forces between the droplets for 20 seconds, which is the time for the emulsion to drain

horizontally and to exit the space between the electrodes and the effect of the power gradient. Eqs. (6) and (7) are used to calculate the speed of displacement of the droplets due to the great reduction in the computational processing time and the small loss of precision. Collisions lead to the occurrence of larger droplets, and these drain vertically downwards, and then decant and are separated from the mixture.

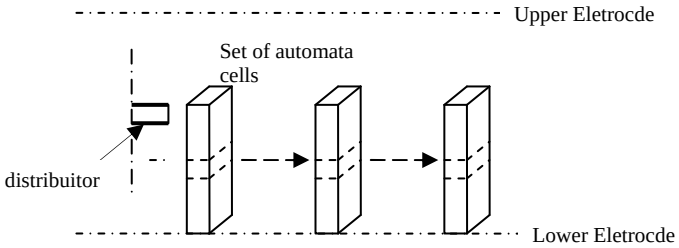


Figure 4. Diagram of the arrangement of the group of automaton cells in the space between the electrodes and their displacement in the Lagrangian perspective.

The state of any one cell depends on the location and the size of each droplet within the cell. The next state of a cell is calculated, and the size and location of each droplet of the cell is obtained. One droplet at a time is moved, and all the other droplets of the cell and the neighboring cells are "frozen". Each iteration consists of two steps. First, the displacement along axes *x, y* and *z* is determined for each droplet of a given cell. Vertical displacement on axis *y* is calculated from the speed resulting from the application of equations. (6) and (7) on all the surrounding droplets. The horizontal displacement results from the force introduced to represent the effects of alignment provoked by the electric power gradient and the turbulence resulting from the discharge of the mixture. Next, which droplets have suffered collision is ascertained, and the new size and position of the resulting droplet is determined.

In this study, each iteration corresponds to an interval of 1 millisecond, and twenty thousand iterations are necessary. Lastly, the volume of the small droplets remaining in the mixture is calculated, thus obtaining the efficiency of separation. Moreover, the group of cells is maintained geometrically fixed, in other words, horizontal and vertical displacement, due to the increase in drainage area resulting from the real geometrical shape of the desalting model, is not considered.

The model considers the variables of operational temperature, concentration of water and distribution of the size of the droplets as parameters, enabling the study of the influence of each one of these in the separation of the emulsion. Due to low kinetic energy the droplets with a radius of less than 5μ do not agglutinate when they hit others that are also less than 5μ. For the case under study, it is also considered that the droplets formed with a radius more than 88 μm may be separated from the mixture by gravity.

3. Results

Figure 5 shows that the efficiency of dehydration, which is the calculation of the percentage of separated water, compared with the volume of water in the emulsion at entry to the desalter, increases with the increase in the concentration of water. The BSW (Basic Sediment and Water), that is the content of water in the mixture that leaves the desalter, reduces with the increase in the concentration of water. Both results are in accordance with expectations for the process because, with the increase in the concentration of water, the number of droplets increases and the distance between the droplets is reduced, resulting in the collision between the droplets becoming more

intense. Figure 6 shows that the increase in the intensity of the power gradient applied has as a consequence the reduction in the BSW and increase in dehydration efficiency.

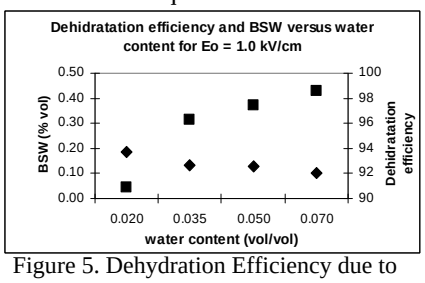


Figure 5. Dehydration Efficiency due to water concentration

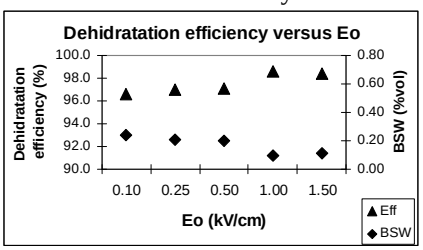


Figure 6. Dehydration efficiency due to electric voltage gradient.

Figure 7 shows emulsion situations, recorded from a computer screen and application stages of the model, thus enabling the determination of the evolution of the agglutination process and the separation of droplets.

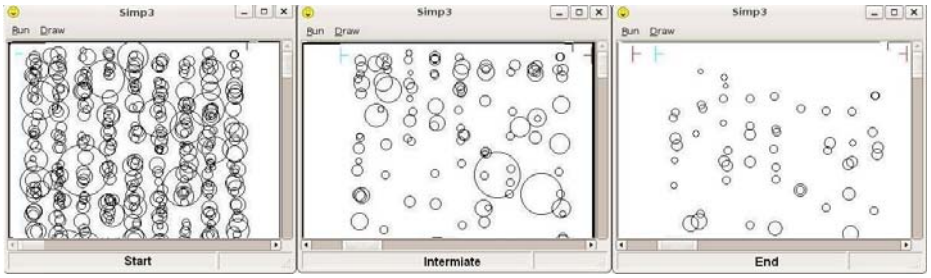


Figure 7. Recorded from process computer screens showing the application of the model.

4. Conclusion

The model developed based on cellular automata with balance of force equations enables the simulated accompaniment of the physical process involving the sequence of agglutination. The results obtained with application of the model based on cellular automata show that the influence of each operational variable studied is coherent with the fundamental theory of the process and they are similar to the results of tests executed in an industrial plant.

References

Bagajewicz, M. A review of recent design procedures for water networks in refineries and process plants. Computer and Chemical Engineering. 24, 2093 – 2113, 2000.

Bhardwaj, A. and Hartland, S. Dynamics of emulsification and demulsification of water in crude oil emulsions, Ind. Eng. Chem. Res, 33, 1271-1279, 1994.

Eow, J.S., Ghadiri, M., Sharif, A.O., Williams, T.J. Electrostatic enhancement of Coalescence of water droplets in oil: A Review of the current understanding, Chemical Engineering Journal, 84 173-192, 2001.

J. von Neumann, The General and Logical Theory of Automata, in J. von Neumann. Collected Works (ed. A.H. Taub), Vol. 5, p.288; J. von Neumann, ``Theory of Self-Reproducing Automata", (ed. A.W. Burks), Univ. of Illinois Press (1966); ed. A.W. Burks, ``Essays on Cellular Automata", Univ. of Illinois Press (1970).

Ohgai, A., Gohnai, Y., Watanabe, K., Cellular automata modeling of fire spread in built-up areas-A tool to aid community-based planning for disaster mitigation, Computers, Environment and Urban Systems, 31 441-460, 2007

S. Wolfram, Statistical Mechanics of Cellular Automata, Rev. Mod. Phys. 55 (1983) 642.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Published by Elsevier B.V.

On the Optimal On-Line Management of Photovoltaic-Hydrogen Hybrid Energy Systems

Victor M. Zavala,[a,b] Mihai Anitescu,[a] and Theodore Krause[b]

[a]*Mathematics and Computer Science Division, Argonne National Laboratory, 9700 S. Cass Ave, Argonne, IL, USA*
[b]*Chemical Technology Division, Argonne National Laboratory, 9700 S. Cass Ave, Argonne, IL, USA*

Abstract

We present an on-line management strategy for photovoltaic-hydrogen (PV-H_2) hybrid energy systems. The strategy follows a receding-horizon principle and exploits solar radiation forecasts and statistics generated through a Gaussian process model. We demonstrate that incorporating forecast information can dramatically improve the reliability and economic performance of these promising energy production devices.

Keywords: receding horizon, stochastic, Gaussian process, solar, hydrogen.

1. Introduction

Hybrid technologies are attractive alternatives for satisfying increasing energy needs in diverse industrial sectors. The main idea is to couple components that generate power from different sources such as fossil fuels or renewables. With this approach, it is possible to overcome cost and efficiency limitations of the individual components and, in turn, minimize the overall system costs. A promising hybrid is the so-called photovoltaic-hydrogen (PV-H_2) system. A schematic representation is given in Figure 1. The idea is to generate electricity from solar radiation to fulfill a given load. The excess power is stored in a battery bank or in the form of hydrogen produced by electrolysis. The stored hydrogen can be converted back to electric power by using a fuel cell system. Power conditioning devices are used to regulate the voltages of the different devices, which are connected to a common busbar.

An important obstacle affecting the reliability of PV-H_2 systems is the fact that the main energy source is intermittent and highly uncertain. To illustrate this, in Figure 2 we present the total solar radiation for year 2004 at location 41° 59' N/87° 54' W in the Chicago, IL area. Another important issue is the fact that the components might have significantly different efficiencies (giving rise to different levels of power losses). Consequently, it might not be immediately evident which component is the optimal one to store and provide energy at a particular time. Motivated by these issues, researchers have devoted significant effort to developing on-line control or management strategies. Most of the strategies reported so far have been based on fuzzy logic and neural networks techniques (Vosen and Keller 1999, Ulleberg 2004). While these strategies might seem practical at a first glance, they are not general enough to handle economics, forecasts, and operational limits systematically. In this work, we present a general on-line management strategy for PV-H_2 hybrid systems. The strategy follows a receding-horizon (RH) technique and incorporates an economic objective function. With this strategy, we can directly study the effect of using forecast information on the overall operating costs. We demonstrate that using management strategies that neglect the

future radiation trends can severely affect the system costs and reliability. Motivated by this observation, we derive a strategy to obtain approximate forecasts and associated statistics through a Gaussian process modeling technique. We then use this information to derive a stochastic RH strategy that satisfies the load reliably.

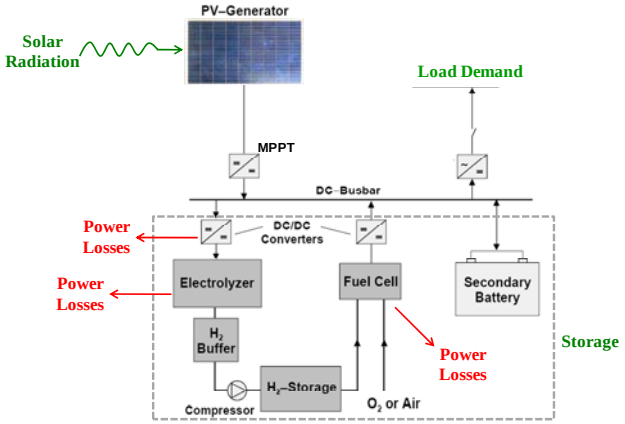


Figure 1. Photovoltaic-Hydrogen Hybrid System

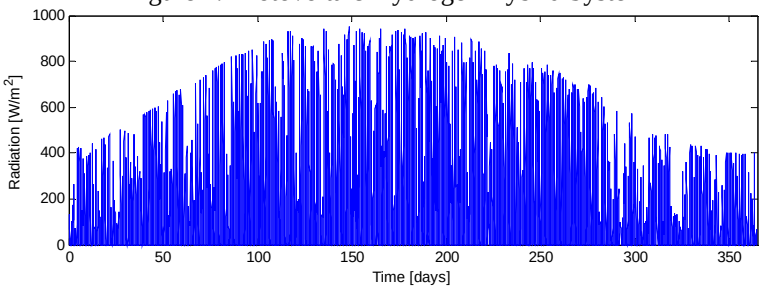


Figure 2. Profile of total solar radiation in Chicago IL, 2004.

2. System Dynamic Model

The dynamic model comprises a systems wide power balance. The power entering through the solar module at time t_k is denoted by P_k^{PV}(kW). This can be calculated by using the measured radiation G_k^T (kW/m^2) and the module design characteristics. The electric current will go through a DC-DC converter that will seek to match the electric current voltage to the voltage of the distribution busbar. This conditioning process has an inherent efficiency θ_{PV} and generates power losses. The remaining power $\theta_{PV}P_k^{PV}$is sent to the busbar to satisfy the current load P_k^{LOAD}. The excess power can be used to produce hydrogen in the electrolyzer and/or to charge the battery. In order to run the electrolyzer, the power extracted P_k^{EL} passes through a buck DC-DC converter, which brings the current voltage down to the operating voltage of the electrolyzer. The efficiency of this step is θ_{BU}. The remaining power $\theta_{BU}P_k^{EL}$enters the electrolyzer. The conversion process to hydrogen has an efficiency θ_{EL}. The *net* price for each power unit produced by the electrolyzer is given by C_{EL}. Since hydrogen can be seen as an asset, the net price can be negative. The produced power $\theta_{BU}\theta_{EL}P_k^{EL}$ in the form of hydrogen is stored in a storage system modeled by a difference equation of the form $E_{k+1}^{H2} = E_k^{H2} + \Delta_k\,(\theta_{BU}\theta_{EL}P_k^{EL} - P_k^{FC})$, where E_k^{H2} is the total energy stored (kWh) at time t_k and $\Delta_k = t_{k+1} - t_k$ (hr). The hydrogen *state of charge* is defined as $SOC_k^{H2} =$

$100\frac{E_k^{H2}}{E_{MAX}^{H2}}$, where E_{MAX}^{H2} is the nominal maximum capacity (kWh). A certain amount of power P_k^{FC} can be withdrawn from the storage to feed a fuel cell and generate electric power. The conversion process has an efficiency θ_{FC}. The cost for each unit of power produced by the fuel cell is given by C_{FC}. The remaining power is then passed through a boost DC-DC converter that brings the voltage of the current up to the operating voltage of the busbar. The process has an efficiency θ_{BO}. The remaining power $\theta_{FC}\theta_{BO}P_k^{FC}$ is sent to the distribution busbar. The system might be able to buy a given amount of power P_k^G from the grid in order to balance the system. This power will have a cost C_G which depends on the location and the degree of independence required by the application (e.g., for a stand-alone system, $C_G = \infty$). Excess power at the busbar can also be dumped to the grid or environment, which is modeled by variable P_k^D. The cost of dumped power is C_D. If the power is dumped to the grid, this cost becomes an asset (set by net-metering rates). The power remaining at the busbar can be used to either charge or discharge the battery. The net battery power P_k^B is calculated as $P_k^B = \theta_{PV}P_k^{PV} + P_k^G + \theta_{FC}\theta_{BU}P_k^{FC} - P_k^{EL} - P_k^{LOAD} - P_k^D$. The battery balance is $E_{k+1}^B = E_k^B + \Delta_k P_k^B$, and the state-of-charge is $SOC_k^B = 100\frac{E_k^B}{E_{MAX}^B}$. The fixed model inputs are P_k^{PV} and P_k^{LOAD}. The degrees of freedom are P_k^{EL}, P_k^{FC}, P_k^G, and P_k^D.

3. Management Strategy

The RH strategy solves, at time t_k, a linear programming (LP) problem of the form

$$\min_{P_j^{EL},P_j^{FC},P_j^G,P_j^D} \sum_{j=k}^{k+N-1} C_{EL}P_j^{EL} + C_{FC}P_j^{FC} + C_G P_j^G + C_D P_j^D$$

$$s.t.\ E_{j+1}^{H2} = E_j^{H2} + \Delta_j\left(\theta_{BU}\theta_{EL}P_j^{EL} - P_j^{FC}\right), \quad j = k, \dots, k+N-1$$

$$E_{j+1}^B = E_j^B + \Delta_j P_j^B, \quad j = k, \dots, k+N-1$$

$$E_k^{H2} = given, \quad E_k^B = given$$

$$P_j^B = \theta_{PV}P_j^{PV} + P_j^G + \theta_{FC}\theta_{BU}P_j^{FC} - P_j^{EL} - P_j^{LOAD} - P_j^D,\ j = k, \dots, k+N-1$$

$$0 \geq SOC_j^B \leq 100, \quad 0 \geq SOC_j^{H2} \leq 100, \quad j = k, \dots, k+N.$$

From the solution of the LP, we obtain the optimal future trajectory for the electrolyzer, fuel cell, grid, and dump powers that minimizes the operating costs, maximizes H_2 production and, simultaneously, satisfies the load and the storage limiting levels. To solve the LP, we need information of the future solar power $P_j^{PV}, j = k, \dots, k+N-1$ expected to be available. Important research questions that, to the best of our knowledge, have not been addressed so far are: What is the economic impact of folding forecast information in on-line management strategies? What is an appropriate forecast horizon? How can we get accurate forecast information? To address these questions, we perform a numerical case study. The efficiencies of the components are obtained from Vosen and Keller (1999). The unit costs are obtained from Stoll and von Linde (2000). A constant load of 1kW is assumed. The maximum peak PV power is 5 kW. We first solve an open-loop optimal control problem using *perfect* forecast information for a horizon of 365 days ($N = 365x24 = 8760$ hours). We use the optimal cost as a reference for the *best economic performance possible* over one year of operation. This is on the order of $1,000/yr. We then run the closed-loop RH strategy spanning the year for different horizons $N = 1, 6, 12, 24, 3x24, 7x24, 14x24$ (hr) with an update time of $\Delta_k = 1$ hr and compute the corresponding relative costs. This required extensive computations. For each scenario, approximately 8,500 LP problems needed to be solved. The 14-day forecast LP contains 2,000 constraints and 1,000 degrees of freedom

and can be solved in less than one second with a state-of-the-art solver. The results are summarized in Figure 3. Several interesting and unexpected conclusions can be drawn from this study: (1) the relative operating costs decay exponentially to zero as the horizon is increased; (2) for a purely reactive strategy ($N = 1$ hr), the relative costs can go as high as 300%; and (3) the *overall best cost can be obtained with a finite forecast* ($N = 24x14$ hr). This last result has important practical implications because it is often difficult to obtain accurate long-term weather forecasts. In addition, note that the economic penalty of using a short forecast of 24 hr is just an increase of 10% in relative costs, whereas the penalty for a forecast of 12 hr is 31%. This implies that a practical horizon should be sufficiently long to capture the periodicity of the daily radiation. The reason for these strong economic penalties becomes evident from Figure 4. Here, we present the power profiles for the fuel cell for both the open-loop and the $N = 12$ hr closed-loop cases. Note that *shorter forecasts induce more aggressive control actions*, which in turn affect the costs. As we increase the horizon, the system is allowed to react more *proactively*, which is reflected in smoother controls.

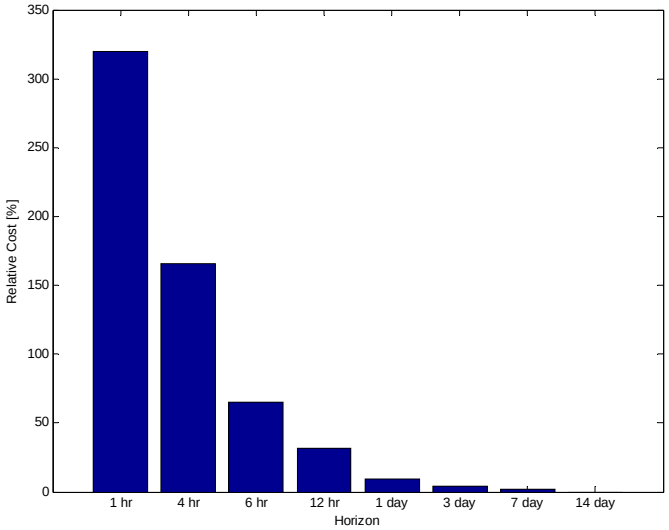


Figure 3. Impact of forecast horizon on economic performance.

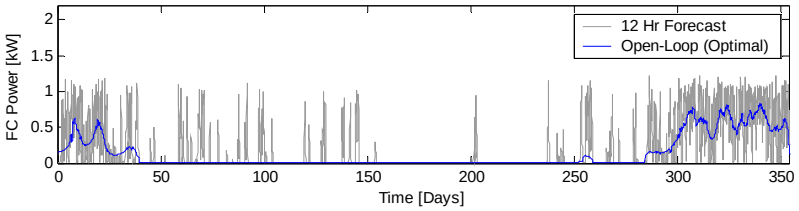


Figure 4. Impact of forecast horizon on power profiles.

4. Gaussian Process Model and Stochastic Management Strategy

Gaining access to solar radiation predictions can be complicated or impractical. In addition, if the forecast is not accurate enough, the management strategy can run out of stored energy prematurely and will be unable to satisfy the load. This situation is particularly critical in stand-alone systems. In the absence of forecasts, we could assume that the radiation profile of the next day will be similar to that of the previous day. Another option is to use historical data to construct a dynamic empirical model. For instance, a time-series approach could be used to build an auto-regressive (AR) model.

An approach that has recently received attention is Gaussian process (GP) modeling (Rasmussen and Williams 2006). The idea is to construct an AR model by specifying the structure of the covariance matrix rather than the structure of the dynamic model itself. We have found that this feature makes the GP modeling particularly flexible. Consequently, this was the approach used in this work. Because of space restrictions, we present only the basics of the GP algorithm. We construct a model by regressing the future radiation (output) y_{k+1} to the current radiation and to the radiation observed T time steps ago (we use $T = 24$ hr to enforce periodicity). We define the inputs $x_k = [y_k, y_{k+1-T}]$ to give $y_{k+1} = f(x_k)$. We collect a number of input-output pairs as training data sets represented by Y and X. We assume that the inputs are correlated through an exponential covariance function

$$K(x, x', \vartheta) = \vartheta_1 \exp\left(-\frac{1}{\vartheta_2}\|x - x'\|^2\right) + \vartheta_3,$$

where $\vartheta_1, \vartheta_2, \vartheta_3$, are parameters estimated by maximizing the log likelihood function:

$$\log p(Y|\vartheta) = -\frac{1}{2} Y K^{-1}(X, X, \vartheta) Y - \frac{1}{2}\log\det\big(K(X, X, \vartheta)\big).$$

Once the parameters are obtained, we can compute mean predictions $\mathrm{Y^T}$ with associated covariance $\mathrm{K^T}$ at a set of test points $\mathrm{X^T}$. In our context, these are the time-varying radiation trends. The predictive equations are

$$\mathrm{Y^T} = K(X^T, X, \vartheta) K^{-1}(X, X, \vartheta) Y$$

$$\mathrm{K^T} = K(X^T, X^T, \vartheta) - K(X^T, X, \vartheta) K^{-1}(X, X, \vartheta) K(X, X^T, \vartheta).$$

In Figure 5, we present the mean forecast and 100 samples drawn from the normal distribution $N(Y^T, K^T)$ at a particular day. We use approximately 400 training data sets. Note that the distribution captures the true radiation values, implying that the assumed covariance structure is reasonable. We use the GP forecast distributions to derive a stochastic RH strategy that minimizes the *expected* cost and satisfies the constraints for all possible realizations of the future radiation. The objective function takes the form:

$$\min_{P_j^{EL}, P_j^{FC}, P_j^{G}, P_j^{D}} \quad \boldsymbol{E}\left[\sum_{j=k}^{k+N-1} C_{EL} P_j^{EL} + C_{FC} P_j^{FC} + C_G P_j^G + C_D P_j^D\right].$$

Symbol $\boldsymbol{E}[.]$ denotes the expectation operator. We solve this infinite-dimensional problem using a sample-average approximation strategy. To test the stochastic strategy, we construct scenarios by sampling the predicted GP distribution. In Figure 6, we compare the performance of the optimal open-loop strategy with perfect forecast (RH optimal) with a strategy that uses the mean forecast (GP mean) and one that uses 100 samples (GP samples). The GP samples LP contains 16,000 constrains and 5,000 degrees of freedom and can be solved in 3 seconds. In addition, we analyze the performance of a simple strategy that propagates the radiation of the last day to the next day (RH simple). In the top graph, we present the state-of-charge of H_2 storage along the year. Note that the GP mean and GP samples approach obtain similar levels to those obtained by RH Optimal. On the other hand, the total H_2 produced with RH Simple is 10% lower. In the bottom graph, note that GP mean fails to satisfy the power demand at one instance during the year (indicated as large or infinite cost). RH simple is the least reliable strategy; failing at 4 instances (overlapping due to time scale). On the other hand, the stochastic GP samples strategy is always able to satisfy the demand.

5. Conclusions

In this work, we have derived a receding-horizon strategy to perform the on-line management of PV-H_2 hybrid energy systems. We conclude that a few days long forecast is sufficient to obtain an acceptable economic performance. In addition, we demonstrate that capturing the uncertainty of the future radiation trends is critical to satisfy the load reliably.

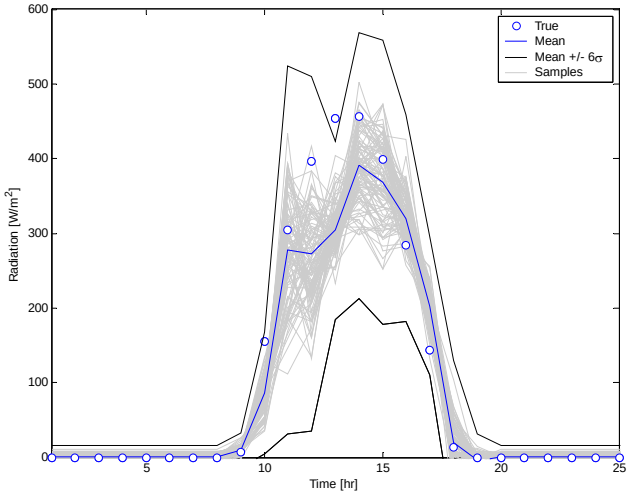


Figure 5. Mean forecast and 100 realizations obtained with the GP model.

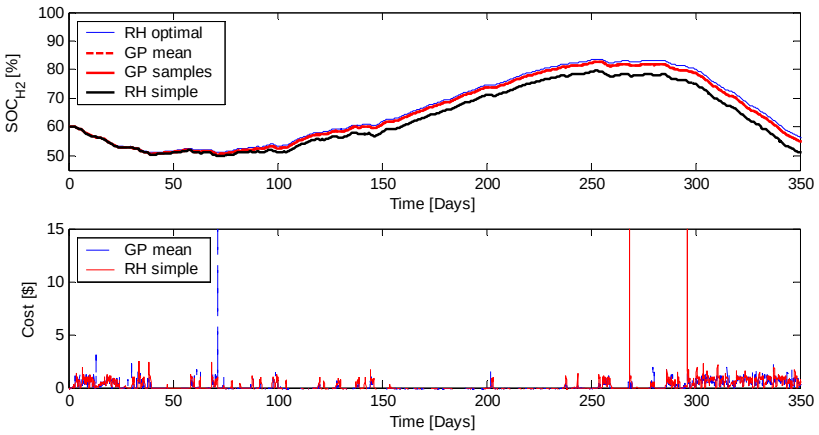


Figure 6. Performance of stochastic RH strategy with GP forecasts.

Acknowledgment

This work was supported by the Office of Advanced Scientific Computing Research, Office of Science, U.S. Dept. of Energy, under Contract DE-AC02-06CH11357.

References

O.Ulleberg, 2004. The importance of control strategies in PV–hydrogen systems, Solar Energy, 76, 323-329.

S. R. Vosen and J. O. Keller, 1999. Hybrid energy storage systems for stand-alone electric power systems: optimization of system performance and cost through control strategies, Int. J. Hydrogen Storage, 24, 1139-1156.

R. Stoll and F. von Linde, 2000. Hydrogen - what are the costs? Hydrocarbon Processing, 79, 42-46.

C.E. Rasmussen and C.K. Williams, Gaussian Processes for Machine Learning. 2006, MIT Press.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Operator Trainer System for the Petrobras P-26 Semi-submersible oil and gas production unit

A. C. Pereira,a A. Riera,b , G. Padilla,b E. Musulin,b N. J. Nakamurac

aPetrobras, Av. Chile 65, Rio de Janeiro 20031-912, Brasil
bSoteica Lationamerica S.A., A. Thomas 796 6B, Buenos Aires C1427CCU, Argentina
cSOteica do Brasil Ltda., Rua Dr. Cid de Castro 108, São Paulo 04064-040, Brasil

Abstract

Operator trainer systems aim to improve operator performance, by simulating scenarios such as emergency conditions, thus reducing accidents and increasing processes economical results. In this paper, we present PETROBRAS' Oil & Gas Production Process and Utilities Simulator Environment called AMBTREI (Training Environment) that mimics the actual Control Room of an E&P semi-submersible Platform at a very high fidelity level. This training environment was created utilizing Soteica's Operator Training System solution (S-OTS). The dynamic process model will be described as well as the Process Control Interface that was implemented. The software used will be explained in detail and the conclusions that have been reached in almost 2 years of use will be presented. **Keywords:** Operator Training, offshore platform, dynamic process simulation.

1. Introduction

1.1. Objectives

Operator trainer systems aim to improve operator performance thus reducing accidents and increasing processes economical results. They allow, among other capabilities:

- simulation of scenarios, such as emergency conditions.
- modeling for operator training purposes.
- studying plant behavior when an upset occurs

Petrobras' Project IND P&G 26.2.2 had the following main objectives:

- to increase capabilities at offshore E&P area, keeping PETROBRAS' Operational Excellence Plan (Plano de Excelência Operacional) guidelines according to its safety rules.
- to provide training, qualification and certification programs for professionals working at oil production areas (process and utilities) in offshore E&P platforms, focusing on normal and emergency operation techniques besides start-up and shutdown platform processes, avoiding on-site training risks.

1.2. Project Highlights

As part of the project, AMBTREI (Training Environment), an Oil & Gas Production Process Simulator for an E&P semi-submersible platform was created. Located at the Núcleo de Treinamento Offshore (NTO) Engenheiro Nelson Stavale Malheiro, in the Technology Center Euvaldo Lodi, Benfica, north part of Rio de Janeiro, this training environment aims to reproduce the actual Control Room of an E&P semi-submersible Platform at the highest fidelity level.

It allows training not only on oil processing but also utilities related processes. Being one of the most modern in the world, its implementation was possible through an agreement between PETROBRAS, Firjan/SENAI-RJ and SOTEICA Ideas and Technologies LLC, with representation in Brasil, acting as the Operator Training Software (S-OTS) provider and dynamic process model developer. Soteica brought to the project past experience in the field of Operator Trainers (Ref. 1) as well as a previous assignment with Petrobras' CENPES in the development of a dynamic model for the same platform (Ref. 2). PETROBRAS also took an active part in the project through its business units UN-BC, E&P-ENGP and CENPES, giving technical and consulting support.
AMBTREI main customer, PETROBRAS, will be training 1824 production operators for the first 7 years of the contract, achieving big economical savings. For Firjan/SENAI-RJ, this deal represents one more step forward its objective of becoming a training solutions provider for the Oil & Gas industry.
As part of the agreement between PETROBRAS and Firjan/SENAI-RJ, the latter expects to train platform operators from other Oil & Gas companies, so that, a partnership with ABERDEEN SKILLS AND ENTERPRISE TRAINING (ASET) was created, through which the Scottish company guarantees Firjan/SENAI-RJ exclusive rights to develop these training activities in Latin America.

1.3. Main Project characteristics
The following items describe the AMBTREI project:

- Production and Utilities Process dynamic model was based on P-26 PETROBRAS Platform. It's possible to simulate other platforms modifying this model.
- This model takes into account:
  - Production and Utilities systems found on actual platforms such us: oil transportation to the platform through risers, three-phase oil-gas-water separation, oil exportation, gas compression, sea water lift, compressed air systems, hot and cold water systems among others.
  - Instrumentation as well as dynamic behavior of input and output variables
- Virtual Platform and Control System bi-directional interaction reproducing actual platform behavior.
- Virtual Control Room (supervisory system) and Control System bi-directional interaction, allows training the operators in using actual supervisory systems, mimicking actual platform behavior.
- Process and Utilities Production training courses are currently available with 2 options: Improvement and Certfication.
- SENAI-RJ instructors, trained and certified by ASET (Aberdeen, Scotland), are located in a trainers' room next to the Virtual Control Room. This room is equipped with Trainer Nodes allowing them to select preconfigured training scenarios, create new ones, simulate participation of field support teams to execute manual movements such as valves opening, etc.
- An ASET certification training course as well as SENAI-RJ certification is available for highly experienced Oil & Gas and Utilities Process Operators.

2. P-26 Platform Dynamic Process Model

The dynamic process model was developed using Hysys Dynamics, a first-principles based simulator. The process model includes:

- 16 production wells.
- Gas Booster compression.

- 10 injection wells.
- Production manifold gas exportation.
- Oil-water separation (electrostatic separation was also considered).
- Oil-gas separation.
- Re-injection Gas Compression.
- Glycol regeneration unit.
- Fuel gas system.
- Hot water system (2 furnaces).
- Sea water lift and distribution.
- Cooling water system.
- Water injection.

The first step in model development involved data gathering and other preparation issues. The following documentation was required:

- P&IDs
- Control system diagrams
- Operational procedures
- Input stream data (pressure, temperature and composition of well fluids arriving on P-26);
- Operating screen snapshots in normal processing condition;
- Main equipment data sheets (vessels, drums, pumps, compressors, valves, etc.);
- Pumps and compressors curves
- Piping data and equipments ground level when needed.
- Feed streams composition analysis data.

Turbine modeling data were especially difficult to obtain. The process data available were related to the global turbine performance only. Data for individual pieces (air compressor, combustion chamber, gas generator and power turbine) were not made available by the manufacturers as these data are considered confidential. A single set of intermediate information (in between pieces of equipment) was obtained. From this information set, additional data for turbine modeling were estimated. To get closer to reality, more intermediate data would be necessary.

Regarding other equipment, almost all required modeling data were gathered. Some minor equipment data sheets were missing and had their data estimated to reproduce actual plant behaviour.

3. Process Control Interface and Details

One of the main tasks of an OTS project is to define the interface between the dynamic process model and the control system. Different alternatives were taken into account:

- Emulation: simulation of the control system with similar look & feel. It eliminates system integration issues with the offline automation system and allows complete control of the simulation environment. However, emulated systems do not protect the integrity or fidelity of automation systems and do not support accurate implementation of the training system components
- Stimulation: usage of actual automation system components that accurately and exactly represent the control system behaviour.
- Virtual stimulation: control configuration and displays are identical to the actual DCS, but control logic runs on virtual controller hardware instead of actual controller hardware.
- Partial stimulation: actual automation system components may be used with some emulation of lower complexity parts of the automation system.

Partial stimulation was the adopted approach, replicating the original installation (hardware and configuration) whenever control hardware was available. Emulation was the choice at the following situations:

- minor equipments control programs run on same model and provider hardware, but to cut costs and reduce installation times, programs were migrated to run in main control hardware
- control hardware was difficult to get.
- some minor equipment logic was so simple that the cost-benefit ratio to add necessary hardware made the effort worthless.

4. Soteica Operator Trainer System (S-OTS)

The Soteica Operator Training System (S-OTS) aims to improve operator performance thus reducing accidents and increasing processes economical results. In particular, it is an excellent tool to train operators under process abnormal conditions.

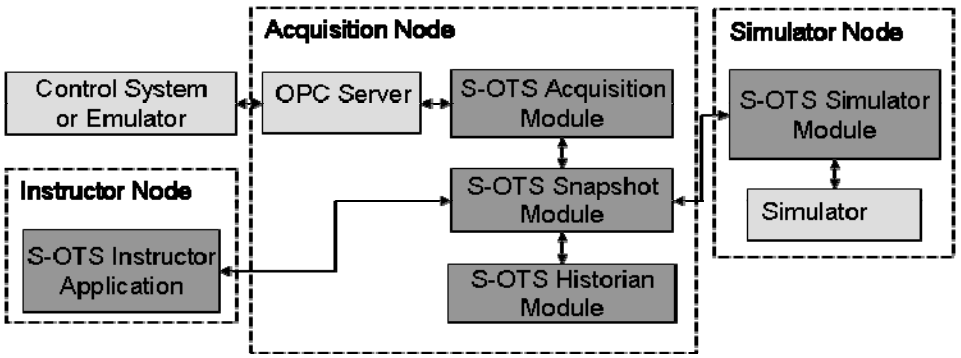


Figure 1: Soteica Operator Trainer System architecture. Arrows shows the flow of information (i.e. the state of the variables) between modules.

The S-OTS is responsible for capturing the operator actions, through the control system or field operated devices (informed to the instructor), send this information to the simulator, execute an integration step of the simulator, take the state of the variables of the simulator and put this information back to the control system. The S-OTS will also save the information related to each training session and keep a database of each user and its training sessions' information. Such information can then be accessed by means of an Excel ® add-in in order to evaluate the operators' performance.
S-OTS was developed following a modular and distributed architecture (see Figure 1) that manages the interaction between every system component (control system, Human-Machine Interface (HMI), Process Simulator and Instructor environment).

S-OTS modules are driven by a special module, the "S-OTS Instructor Application", which has three components:

- the Instructor application administrator
- the Excel history lookup tool
- the Training sessions common database.

It has a graphical user interface that interacts with the instructor to register users as instructors or operators to be trained, keep a list of training sessions developed, start, stop and pause training sessions, enter process perturbations or changes, read process variables values and searches sessions by user or date criteria and access history files from Excel®.

The S-OTS Snapshot module keeps the state of all the configured variables required for the training session at any given time. This module allows the other modules to exchange the state of the variables of the training sessions.
The S-OTS Simulator module reads the controller outputs from the S-OTS Snapshot, puts them into the Hysys model, executes an integration step and puts the new simulated values in the S-OTS snapshot.
The S-OTS Acquisition module is a bi-directional interface from the S-OTS Snapshot to the PLC controller using OPC standard. It writes the PLC controller outputs to the snapshot module and writes the current simulated variables values back to the PLC.
The S-OTS Historian module stores process data (temperatures, flows, etc.), instructor variables and field operated devices values. Once the training session is finished the S-OTS Instructor application gets the archived data and stores it in a file on the Instructor Node. The access to historical data will be available through the Excel® history lookup tool.
As a consequence of its modular architecture, additional functionality could be added to the S-OTS system to satisfy special needs.

5. Conclusions

As result of this project, PETROBRAS has a Production Process and Utilities Simulator Environment called AMBTREI (Training Environment), that mimics the actual Control Room of an E&P semi-submersible Platform at a very high fidelity level, allowing among other benefits:

- to increase capabilities at offshore E&P area
- to provide training, qualification and certification programs, focusing on normal and emergency operation techniques besides start-up and shutdown platform processes, avoiding on-site training risks and achieving important economical savings.
- to improve operators performance reducing accidents and increasing process economical results. In particular, the S-OTS environment is specially suited to train operators under abnormal process conditions.
- Trained people in 2007 (activities started on Jul-2007): 90.
- Trained people until November 2008 (Nov-11-2008): 167.

It's important to emphasize the high level of acceptance that these training activities have, based on the Operators' feedback. In particular they pointed out the following aspects:

- The interactive way in which activities are carried, and
- as this activity covers a lot of very useful topics it serves as a refresher of a great amount of knowledge acquired in past training courses.

6. Acknowledgments

This project is the result of the effort of many participants; we want to extend our special thanks to: Sérgio Ferreira Pinto (made PLC programs and HMI displays modifications), Wanderley de Carvalho Rangel and Carlos Edmundo C. Lima (tested dynamic process model to verify it reproduced actual plant behaviour); José Luis Loureiro Alves and Paulo Elias Bucazzio, AMBTREI instructors who intensively tested the S-OTS software. Also, we want to give a special mention to Rajeev Agarwal who led the development of the initial dynamic model that was used as the base for this OTS.

References

Rajeev Agarwal, Ky Alston, Thomas J. Oliver, Juan Pablo Ruiz, Sebastián Cúneo, Advanced Operator Training and Plant Operations for a Gas-Plant using First Principles Dynamic modeling and Human-Machine Interface Emulation, IAPG Natural Gas Treatment Congress, 2004,<http://www.soteica.com/files/papers2dwnld/2004_OTS_Gas_Plant_IAPG_Natural_Gas_Treatment_Congress.pdf>

E. N. Satuf, J. C. Vaz1, R. F. Pimenta, R. Agarwal, How to operate a complex floating production unit inside a meeting room, 4th Mercosur Congress on Process Systems Engineering, 2005, <http://dpi.eq.ufrj.br/Anais_A/CENPROMER2005/nukleo/pdfs/1061_mpv_p26_enpromer_1061_2005_final.pdf>

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Stochastic Programming with Tractable Mean–Risk Objectives for Refinery Planning under Uncertainty

Cheng Seong Khor, Thi Huynh Nga Nguyen

Chemical Engineering Department, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 31750 Tronoh, Perak, Malaysia

Abstract

The application of information technology (IT) and information systems (IS) have been crucial in enhancing the operating flexibility and resiliency of refineries. In particular, the process systems engineering (PSE) community has been instrumental in carrying out a key role in extending the systems engineering boundaries from mere chemical process systems to the incorporation of business process systems with consideration for risk. Thus, this paper considers a robust framework for the economic and operational risk management of a refinery under uncertainty by extending an existing two-stage stochastic program with fixed recourse via scenario analysis. The problem is mathematically formulated as a two-stage stochastic nonlinear program with a tractable mean–risk structure in the objective function. Two measures of risk are considered, namely the metrics of mean-absolute deviation (MAD) and Conditional Value-at-Risk (CVaR). The scenario analysis approach is adopted to represent uncertainties in three types of stochastic parameters, namely prices of crude oil and commercial products, market demands, and production yields. However, a large number of scenarios are required to capture the stochasticity of the problem. Therefore, to circumvent the problem of the resulting large-scale model, we implement a Monte Carlo simulation approach based on the sample average approximation (SAA) technique to generate the scenarios. A statistical-based scenario reduction strategy is applied to determine the minimum number of scenarios required yet still able to compute the true optimal solution for a desired level of accuracy within the specified confidence intervals. The proposed model is illustrated through a representative numerical example, with computational results demonstrating how risk-averse-and risk-inclined solutions in the face of uncertainty can be attained in a risk-conscious model.

Keywords: two-stage stochastic programming, refinery planning, mean-absolute deviation (MAD), Conditional Value-at-Risk (CVaR)

1. Introduction

Stochastic programming has emerged as one of the most prominent operation research models for optimization involving uncertainty. Refinery planning problems are subject to uncertainty in many factors, which primarily includes fluctuations in prices of crude oil and saleable products, market demand for products; and production yields. The risk terms in Khor (2007) are handled using the metric mean-absolute deviation. After obtaining the first model with MAD as risk measurement, the second model is developed in which the risk terms are performed by CVaR. A comparison between the two models to assess which of these two risk measures is superior, both computationally

and conceptually, in capturing the economic and operating risk in the planning of a refinery. However, there are large numbers of scenarios that create difficulty to handle various circumstances. For example, there may be more than thousands of cases happening. It is hard to predict and control numerous scenarios. Therefore, it is necessary to find the minimum number of scenarios to capture all the circumstances. Monte Carlo simulation approach based on the sample average approximation (SAA) technique is applied in this thesis to generate the minimum number of scenarios which present for thousands cases.

Risk Model III as presented in Khor et al. (2007):

$$\max z = E[z_0] - \theta_1 V(z_0) - E_s - \theta_3 W \tag{1}$$

where $E[z_0]$ is the expectation of the original objective function z_0 with random price coefficients; θ_1 and θ_3 are weights representing the risk factors; $V(z_0)$ is the sampling variance of z_0; E_s is the expected recourse penalty, and W is the MAD-based risk measure. In this work, we extend model (1) by considering MAD as the risk measure in place of variance. Further, we investigate the viability of the risk measure Conditional Value-at-Risk (CVaR), which has gained wide attention in computational finance, within the domain of refinery planning. All the uncertain parameters are assumed to be discrete random variables.

2. Monte Carlo Simulation Approach based on Sample Average Approximation (SAA) Method

In this work, we adopt the Monte Carlo simulation approach for scenario generation based on the Sample Average Approximation (SAA) method (Shapiro, 2000; Shapiro and Homem-de-Mello , 1998; You and Grossmann, 2008). The procedure involved is as follows:

Step 1. A relatively small number of scenarios (for example, 50 scenarios) with their associated probabilities are randomly and independently generated for the uncertain parameters of prices, demands, and yields. This is accomplished by employing the Monte Carlo simulation approach based on the SAA technique. (This data is otherwise obtained from plant historical data.) The resulting stochastic model (a linear program) with the objective function given in (2) is solved to determine the optimal stochastic profit with its corresponding material flowrates.

$$\max E_z = E_{z_0} - E_\xi \tag{2}$$

where $$E_{z_0} = \sum_{i\in I}\sum_{s\in S} p_s c_{i,s} x_i$$

$$E_\xi = \sum_{i\in I}\sum_{s\in S} p_s \xi_{i,s} = \sum_{i\in I}\sum_{s\in S} p_s \left[\left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^- \right) + \sum_{k\in K} \left(q_{i,j}^+ y_{i,k,s}^+ + q_{i,j}^- y_{i,k,s}^- \right) \right]$$

Step 2. The Monte Carlo sampling variance estimator is determined using the optimal stochastic profit and flowrates computed in step 1.

$$S(n) = \sqrt{\frac{\sum_{s=1}^{S} \left(E_z - z_{i,s} \right)^2}{S-1}} \qquad \text{where } z_{i,s} = \sum_{i\in I} \left(c_{i,s} x_i + \xi_{i,s} \right) \tag{3}$$

Step 3. The lower- and upper-confidence limits of the 95% confidence interval H of 1–α are computed as follows:

$$\left[E_z - \frac{z_{\alpha/2} S(n)}{\sqrt{S}}, E_z + \frac{z_{\alpha/2} S(n)}{\sqrt{S}} \right] \tag{4}$$

Step 4. The minimum number of scenarios N that is theoretically required to obtain an optimal solution is determined using the relation below:

$$N = \left[\frac{z_{\alpha/2} S(n)}{H} \right]^2 \tag{5}$$

where the standard normal random variable $z_{\alpha/2} = 1.96$ at confidence interval $(1-\alpha) = 95\%$.

3. Formulation of Stochastic Refinery Planning Model with Mean-Absolute Deviation (MAD) as Risk Measure

The risk metric mean-absolute deviation (MAD) is employed as a measure of deviation from the expected profit (Konno and Koshizuka, 2005; Konno and Yamazaki, 1991). It is defined as follows:

$$\mathrm{MAD}(x) = E\left| \sum_{j=1}^{n} R_j x_j - E\sum_{j=1}^{n} R_j x_j \right| \tag{6}$$

In this work, the rate of return R in (6) refers to unit cost of materials (crude oil and refinery products) and the amount of money x_j invested in an asset j refers to the refinery production amount. Therefore, the formulation of the MAD-based risk measure for price uncertainty becomes:

$$\mathrm{MAD}(z_0) = \sum_{s\in S} p_s \left| \sum_{i\in I} c_{i,s} x_{i,s} - \sum_{i\in I}\sum_{s\in S} p_s c_{i,s} x_{i,s} \right| \tag{7}$$

while for demands and yields uncertainty, it is given by:

$$\mathrm{MAD}_{\xi} = \sum_{s\in S} p_s \left| \sum_{i\in I} \xi_{i,s} - \sum_{i\in I}\sum_{s\in S} p_s \xi_{i,s} \right| \tag{8}$$

4. Formulation of Stochastic Refinery Planning Model with Conditional Value-at-Risk (CVaR) as Risk Measure

CVaR, also termed as Mean Excess Loss, Mean Shortfall or Tail VaR, is a risk assessment technique that is originally intended to be employed for reducing the probability that an investment portfolio will incur high losses. It offers the advantage of a linear programming formulation for determining the optimal solution of financial planning problems that explicitly minimizes loss or risk. CVaR is performed by taking the likelihood (at a specific confidence level, e.g., 0.95 or 0.99) that a specific loss will

exceed the metric known as Value-at-Risk (VaR). From a mathematical point of view, CVaR is derived by taking a weighted average between VaR and the losses exceeding VaR. For a discrete probability distribution function, CVaR can be defined as follows (Rockafellar and Uryasev, 2000; 2002):

$$F_\alpha(x, \mathrm{VaR}) = \mathrm{VaR} + \frac{1}{1-\alpha}\sum_{i\in I}\sum_{s\in S} p_s \left(f\left(x, y_{i,s}\right) - \mathrm{VaR}\right) \tag{8}$$

Using CVaR as the risk metric yields the following form of the objective function:

$$\max\ z = E[z_0] - \theta_1 \mathrm{CVaR}_{z_0} E_{s'} - \theta_3 \mathrm{CVaR}_\xi \tag{9}$$

where CVaR_{z_0} is the risk measure imposed by the recourse costs to handle price uncertainty.

$$\mathrm{CVaR}_{z_0} = \mathrm{VaR}_1 + \frac{1}{1-\alpha}\sum_{s}\sum_{i} p_s \left(c_{i,s}x_{i,s} - \mathrm{VaR}_1\right) \tag{10}$$

where CVaR_ξ is the risk measure imposed by the recourse costs to handle uncertainty in demands and yields.

$$\mathrm{CVaR}_\xi = \mathrm{VaR}_2 + \frac{1}{1+\alpha}\sum_{i\in I}\sum_{s\in S} p_s \left[\left(c_i^+ z_{i,s}^+ + c_i^- z_{i,s}^-\right) + \sum_{k\in K}\left(q_{i,j}^+ y_{i,k,s}^+ + q_{i,j}^- y_{i,k,s}^-\right) - \mathrm{VaR}_2\right] \tag{11}$$

Substituting equations (10) and (11) into (9), we obtain a two-stage stochastic programming model with mean–risk objective in which the risk measures are assessed by CVaR.

5. Numerical Experiments

We illustrate the risk modeling approach proposed in this paper on the numerical example taken from Khor et al. (2008) and provide major details on the implementation using GAMS/CONOPT3.

5.1. Solving Two-Stage Stochastic Program with MAD as Risk Measure

The expectation of the objective function value is given by the original objective function itself: The corresponding expression for expected profit is formulated for the 13 scenarios that has been randomly generated.

5.2. Solving Two-Stage Stochastic Program with CVaR as Risk Measure

The following is the procedure for developing a loss distribution in order to determine the value for the parameter VaR.

Step 1. The objective value of deterministic profit for each of the 13 scenarios is computed (i.e., multiplication of flowrate and the corresponding price per unit flowrate).

Step 2. The probability of each scenario is randomly generated using Monte Carlo simulation based on pseudorandom number generation.

Step 3. The computed values in Step 1 are sorted in ascending order.

Step 4. The plot of cumulative distribution function against the sorted deterministic profit values is developed to obtain a representation of the loss distribution. At confidence interval of $(1-\alpha) = 0.95$, we can read off the value of VaR from the loss distribution plot, as depicted in Figures 1 and 2, which represents the penalty for uncertainty in prices and in both demands and yields, respectively.

Stochastic Programming with Tractable Mean–Risk Objectives for Refinery Planning under Uncertainty

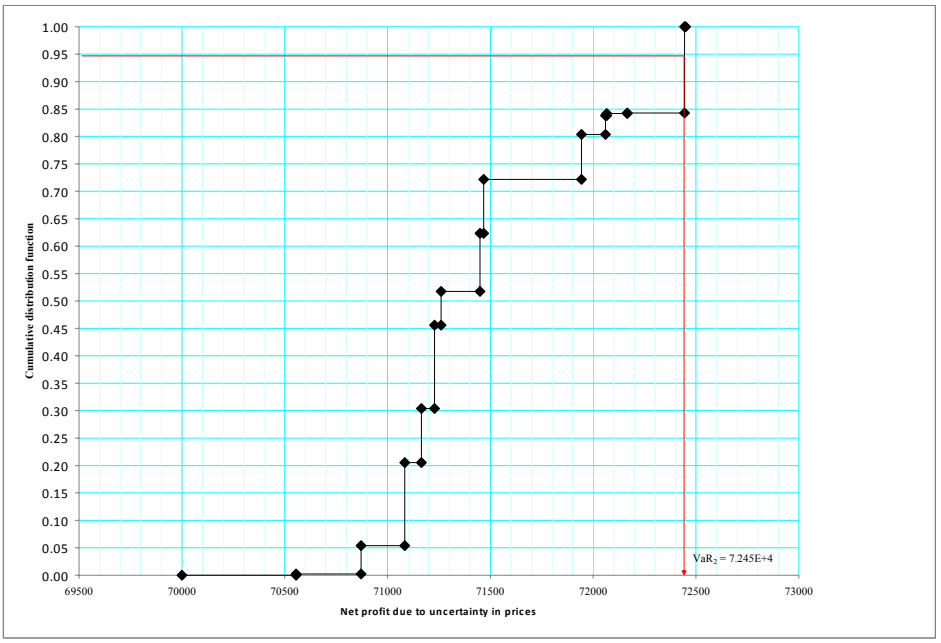


Figure 1: Loss distribution to determine VaR_1

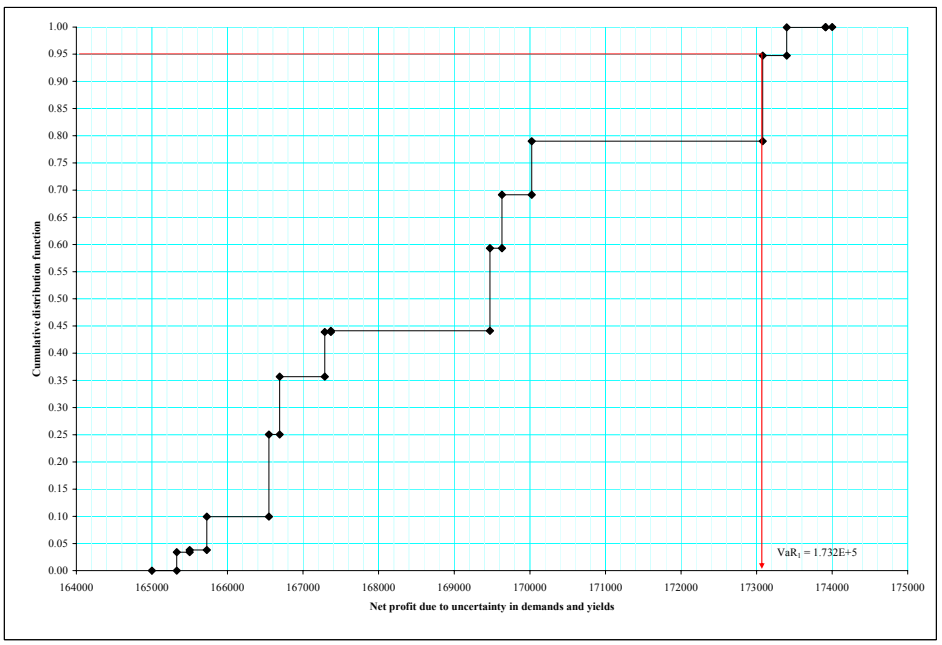


Figure 2: Loss distribution to determine VaR_2

The computational statistics and a summary of the main computational results are provided in Tables 1 and 2 , respectively.

Table 1. Summary of computational results

Monte Carlo sampling variance estimator $S(n)$	489.4
Lower bound of confidence interval H	965.3
Upper bound of confidence interval H	1237
Range of confidence interval H	271.3
Minimum number of scenarios N	13
Optimal solution for MAD-based model	\$681.95/day
Optimal solution for CvaR-based model	\$20 800.66/day

Table 2. Computational statistics of GAMS implementation for determining optimal solutions of MAD- and CVaR-based mean–risk stochastic program

Solver	GAMS/CONOPT3
Number of continuous variables	281
Number of single equations	145
CPU time/resource usage	(trivial)
Number of iterations	20 (using MAD) 101 (using CVaR)

6. Conclusions

This work attempts to consider the use of the risk metrics of MAD and CVaR for the explicit handling of economic and operational risk management in refinery planning problems under uncertainty in prices, demands, and yields.

References

R. B. Webby, P. T. Adamson, J. Boland, P. G. Howlett, A. V. Metcalfe, and J. Piantadosi, 2007, The Mekong—applications of value at risk (VaR) and conditional value at risk (CVaR) simulation to the benefits, costs, and consequences of water resources development in a large river basin, Ecological Modeling, 201, 89–96.

C. S. Khor, A. Elkamel, K. Ponnambalam, and P. L. Douglas, 2008, Two-Stage Stochastic Programming with Fixed Recourse via Scenario Planning with Financial and Operational Risk Management for Petroleum Refinery Planning under Uncertainty, Chemical Engineering and Processing, 47, 9–10, 1744–1764.

H. Konno, and H. Yamazaki, 1991, Mean Absolute Deviation Portfolio Optimization Model and Its Application to Tokyo Stock Market. Management Science. 37, 519–531.

W. K, Mak, D. P. Morton, and R. K. Wood, 1999, Monte Carlo Bounding Techniques for Determining Solution Quality in Stochastic Programs, Operations Research Letters, 24, 47–56.

H. M. Markowitz, 1952, Portfolio selection, Journal of Finance, 7, 77–91.

R. T. Rockafellar and S. Uryasev, 2000, Optimization of conditional value-at-risk, Journal of Risk, 2, 3, 21–41.

R. T. Rockafellar and S. Uryasev, 2002, Conditional Value-at-Risk for general loss distributions, Journal of Banking and Finance, 26, 7: 1443–1471.

A. Shapiro, and T. A. Homem-de-Mello 1998, A simulation-based approach to two-stage stochastic programming with recourse, Mathematical Programming 81, 301–325.

A. Shapiro, 2000, Stochastic Programming by Monte Carlo Simulation Methods, Stochastic Programming E-Prints Series, Retrieved February 10, 2007 from the World Wide Web: http://www.isye.gatech.edu/%7Eashapiro/publications/SBOsurvey.pdf

F. You, J. M. Wassick, and I. E. Grossmann. Risk management for a global supply chain planning under uncertainty: models and algorithms [textfile]. Retrieved September 26, 2008 from the World Wide Web: http://egon.cheme.cmu.edu/Papers/RiskMgmtDow.pdf

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Oil products distribution systems: decomposition approach on pipeline and inventory scheduling

Susana Relvas,[a*] Ana Paula Barbosa-Póvoa,[a] Henrique A. Matos[b]

[a]*CEG-IST, UTL, Av. Rovisco Pais 1049–001 Lisboa, Portugal*
[b]*CPQ, IST, UTL, Av. Rovisco Pais 1049–001 Lisboa, Portugal*
[*]*corresponding author: susanaicr@ist.utl.pt*

Abstract

This paper addresses the problem of oil products scheduling in a system where the supply is provided by a pipeline and a tank farm that enables storage and market fulfillment. Multiproduct pipelines are complex equipments where the operational time window between product load to pipeline until product shipment for customers is wide, requiring longer scheduling horizons. However, problem complexity rises significantly with the time horizon extent and if the problem is approached through an exact optimization method such as mathematical programming the effort to obtain a solution is considerably high. This problem motivates the approach proposed within this paper consisting of an integrated framework that aims problem complexity reduction and lower computational effort. The approach combines two continuous time MILP models with different levels of detail with an iterative procedure that exchanges information between both levels. The approach is exemplified using a real world scenario of a Portuguese company that transports and distributes six oil products.

Keywords: Pipeline, Tank Farm, MILP, Decomposition.

1. Introduction

The oil supply chain is a complex dynamic system, where scheduling and planning are challenging problems to address. Besides the integration of several entities, the proper definition and representation of each activity requires problem-oriented approaches. One example is oil products' pipelines. The multiproduct nature, continuous operation and operational lag between product load and product consumption give origin to complex problems. The challenge is to solve the problem through an exact method to obtain the optimal solution with all associated constraints, in a reduced amount of time. Current published works have mainly focused on Mixed-Integer Linear Programming (MILP) models (Cafaro and Cerdá, 2008, Magatão *et al.*, 2008, Rejowski and Pinto, 2008). However, some integrated approaches were already proposed, such as the hybrid approach by Boschetto *et al.* (2008) that deals with pipeline networks. But, dealing with medium-term horizons has been a difficult task. Cafaro and Cerdá (2008) proposed a rolling horizon where some instances may be rescheduled, but each period still looks over a short-term horizon. In the present work we propose an iterative procedure between two MILP models that represent a multiproduct pipeline system through two different levels of detail in the tank farm management.

2. Problem Statement

Figure 1 resumes the operating system in study that comprises an oil products' pipeline that pumps from a refinery to a tank farm. The distribution centre is located in a strategic market position. Each tank has a fixed product service and the clients are supplied at the distribution centre with the respective products.

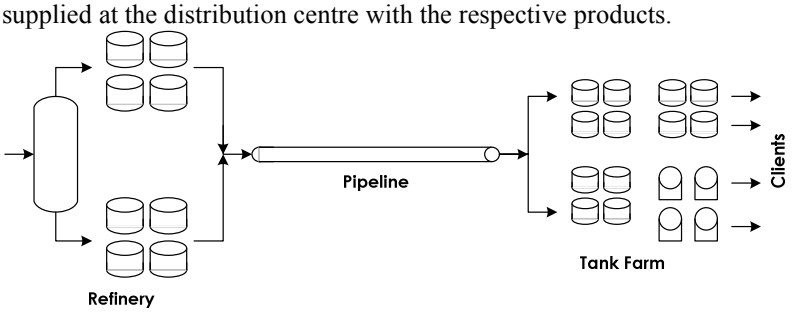


Figure 1. Multiproduct pipeline operating system

Given: a) the pipeline volume, b) maximum and minimum flowrates, c) the products to be pumped and matrix of possible sequences, d) the storage capacity by product, e) settling period by product, f) the time horizon extent, g) the maximum number of batches to be pumped, h) the initial inventory by product, i) the daily demand forecasts and k) the planned pipeline stoppages; **Determine**: 1) the pipeline schedule and 2) the inventory management. The pipeline schedule includes products' sequence, pumping flowrates, batches' volumes, timing issues and pipeline stoppages. The inventory management includes pipeline inputs, settling periods and outputs by product for customers. The problem is formulated through an MILP continuous time model and the main target is to obtain problem solutions for a time horizon compatible with the problem addressed. Short-term periods usually generate solutions where the pipeline schedule does not reflect customer requirements ahead due to the operational lag.

3. Proposed Approach

Figure 2 succinctly describes the proposed approach, which is divided in four steps:

1. sequencing heuristic;
2. sequence validation and selection through the lower detail model (ATFModel);
3. schedule generation through consecutive integer cuts (ATFModel);
4. inventory management schedule generation through a detailed model implementation (DTFModel).

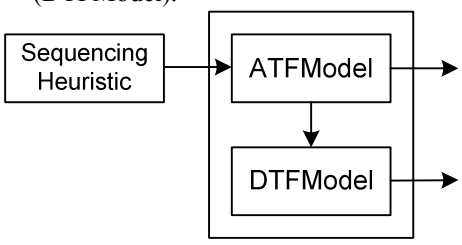


Figure 2. Proposed approach: integrated MILP models with a sequencing heuristic.

The sequencing heuristic builds valid product sequences to be pumped in the given time horizon. This heuristic uses as information the market requirements and initial inventory so as to establish product priorities (Relvas *et al.*, 2009). The proposed heuristic provides inputs for the MILP models. In a first level, an Aggregated Tanks' Farm

Model (ATFModel) is run. This model considers that the tank farm can be represented in an aggregated manner on capacity: each product is stored in a single tank with a total capacity given by the summation of real tanks' capacity (Relvas *et al.*, 2006). With this model it is possible first to validate the proposed sequences and select one (or more) that returns valid results for the given time horizon. The schedule generated in this second step can be used at a higher level (e.g. for refinery planning). However an integrated procedure can be carried out, since the system considered can be represented through a more detailed formulation. A Detailed Tanks' Farm Model (DTFModel) is derived from the ATFModel that considers tanks as individual model entities. Although the model complexity and size increases, a higher level of detail is achieved for the operational decisions such as schedules for product receiving vs. available tanks and delivery schedules vs. tanks with products ready for shipment. This model addresses the complex inventory management problem at liquid products' tank farm.
The main procedure developed is based on the model integration approach. The ATFModel generates a first pipeline schedule to feed the DTFModel for a further generation of a feasible inventory solution. Although, the ATFModel already takes into account a rough inventory management representation a valid solution from this model may return an infeasible solution at the DTFModel level. This gap is reduced developing a battery of schedules with the ATFModel to test within the DTFModel level. This set of solutions is developed through an iterative approach where at each new iteration an integer cut is added. The integer cut eliminates from the current solution space a batch volume sequence with a certain degree of similarity. This is explained below: given the binary variable $y_{i,p}$ that assigns a product *p* to a batch *i* and defined through the sequencing heuristic, it can be disaggregated in the related binary variable $ls_{i,p,lt}$ that assigns a valid volume *lt* to batch *i* of product *p*. Therefore, an integer cut in the form of equation (1) is added at each model iteration.

$$\sum_{i\in I}\sum_{p\in P}\sum_{lt\in LT} ls_{i,p,lt}\Big|_{\left(ls_{i,p,lt}\right)_{iter-1}=1} \le |I| - m, \quad \forall n \in N \tag{1}$$

where $n \in N$ represents the set of iterations, $|I|$ corresponds to the cardinality of the set of batches *I* and m represents a degree of diversity between consecutive batch volumes' sequences given by $ls_{i,p,lt}$. The model uses as stopping criteria either a limit on iterations reached or an infeasible solution found. The solution found in each iteration for $ls_{i,p,lt}$ is given as fixed data to the DTFModel. At each iteration, this model evaluates the feasibility of the fixed binary variables and the corresponding optimal solution with the complete pipeline schedule and inventory management.

4. Framework Implementation and Results

The proposed approach was implemented using as motivating example the real world scenario of Companhia Logística de Combustíveis (CLC), a Portuguese company that operates a multiproduct pipeline (single origin and single destination) and distributes six different oil products. At the ATFModel level, there will be six storage tanks, while at the DTFModel level there are 29 tanks, each with a fixed product service. Two case studies were considered: E1 with a one-week scenario and E2 with a two-week scenario with data information from September 2008.

The models were implemented in GAMS 22.6 and solved with CPLEX 11.0, on a Pentium D820 with 2 GHz RAM.
The extent of the set of iterations considered was 50 and the diversity parameter m was considered to be equal to 1 in E1 and 4 in E2. As stopping criteria it was used the optimal solution or, in case of E2/DTFModel, a computational time limit of 600 s.
The sequencing heuristic returned a set of 6 sequences per case study. E1 sequences had a number of batches between 6 and 11 while E2 had sequences with a number of batches between 11 and 16. Theses sequences were implemented in the ATFModel and compared through a series of operational and model performance criteria so as to select the final sequence. The E1 pumping sequence has 7 batches and for E2 has 14 batches.
Table 1 details the model size for each case study and for each model. The model size clearly increases with the time horizon extent, with the number of batches pumped and with the level of detail of the model.

Table 1. Model size by case study and model formulation

	E1		E2	
	ATFModel	DTFModel	ATFModel	DTFModel
Constraints	1867*	5049	6683*	11887
Continuous Variables	863	1240	2921	3254
Binary Variables	285	1027	964	2226

*the referred number of constraints corresponds to the problem before implementing the iterative procedure and, this means, before adding integer cuts.

Figure 2 represents the evolution of the objective function value with the respective iteration number. The used objective function optimizes operational variables such as minimization of the pumping flowrate and maximization (using a negative sign) of the total pipeline usage.

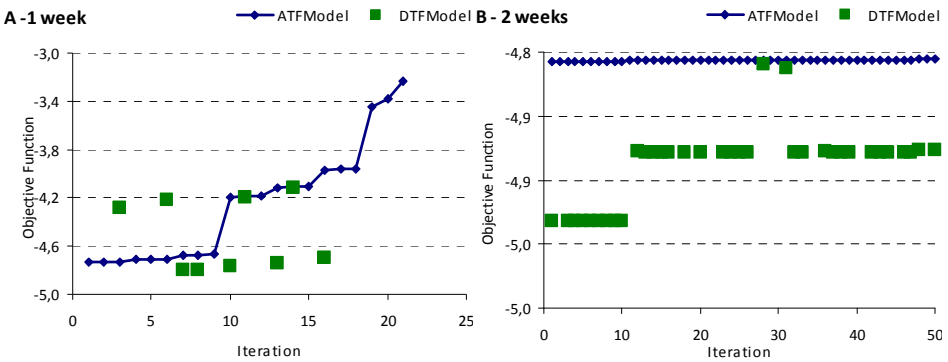


Figure 2 – Evolution of the objective function value with the iteration number in the integrated strategy

It can be observed that for E1 (Figure 2a) with only 21 iterations returned solutions whereas for case study E2 all iterations returned solutions. In terms of objective function degradation, E1 has a visible degradation trend while in E2 the trend is very smooth. It would be necessary to have either a higher diversity factor or a higher number of iterations to explore all the possible alternatives. This result emphasizes the combinatorial nature of the model. However, it can be seen that no continuous line is returned when the DTFModel is run. For a feasible ATFModel solution it is possible to have an infeasible solution at the DTFModel. This occurred for case study E1 while for

case study E2 the iterations with no solution correspond to instances where a feasible solution was not found within the maximum computational time of 600 s. In the case of E1, the solutions returned for the DTFModel are either similar, better or worse than the solutions returned from the ATFModel. On the contrary, the solutions found in the case of DTFModel in E2 are all better than for the ATFModel.

Table 2. Computational results

	E1		E2	
	ATFModel	DTFModel	ATFModel	DTFModel
Optimal Solutions	21	10	50	33
Total Number of Feasible Solutions	21	10	50	35
Total CPU effort (s)	7.8*	55.6/41.6*	4100.4	20350.6/10117.7*
Medium CPU effort (s)	0.37	4.62	82.01	316.18*
CPU effort (s) (no integration)	-	7.4	-	7225.2

*considering only optimal solutions

In terms of computational effort, table 2 presents the results summary. E1 used less than 10 seconds to run all iterations at the ATFModel level and less than 60 seconds to run all iterations at the DTFModel level. For comparison, it was run the DTFModel with no information provided from the ATFModel and the optimal solution was found in 7.4 s. When the time horizon expands to the double (from E1 to E2), the DTFModel without integration requires about 2 hours of computation time to prove optimality. For this scenario, however, the relative gap was 0.02%. The set of iterations was run in 4100 s for the ATFModel (in average each iteration takes 82 s) but at the DTFModel, 35 solutions were found after a total computational time of 20350 s (5.65 h).

Finally, table 3 presents the operational results obtained for each scenario and each model, as well as the correspondent schedule data developed by CLC's schedulers. It also has information about the iterations where the best objective function value was found for each model. While for E1 there is no correspondence between the iteration where the best solutions are found, in the case of E2 there is.

Table 3. Operational results

	E1			E2		
	ATFModel	DTFModel	CLC	ATFModel	DTFModel	CLC
Iteration with best FO	1-2	7-8	-	1-10	1;3-10	-
Medium flowrate (vu/h)*	469.9	470.4	519.0	465.2	466.2	494.6
ΔInventory (vu)*	+2350	+450	+4769	+265	+265	+11105
Pipeline usage (%)	95.9	93.4	89.6	93.8	93.7	94.8

*vu stands for volumetric units

Some operational indicators were selected for comparison. The medium flowrate is calculated for the period when the pipeline is pumping, taking into account the flowrate employed for each batch. The ΔInventory represents the difference between total pipeline inputs and total outputs for the local market. Finally, the pipeline usage is obtained by the ratio of total pumping time and total horizon time.

If the operational indicators are compared for each case study, there is a negligible degradation between the runs of the ATFModel and the DTFModel. However, when comparing the schedule developed by CLC's schedulers, it can be seen that the medium flowrate is higher and, at the end, it is transported a considerable amount of volume than it is consumed. However, when looking in depth over the flowrates by batch, the models propose flowrates no higher than 470 vu/h in E1 and no higher than 490 vu/h in E2, having as lower bound the value of 450 vu/h. In the schedule proposed by CLC's schedulers, flowrates vary between 450 and 700 vu/h. In terms of pipeline usage, the values are similar, except a lower pipeline usage proposed by CLC's schedulers in the first week.

5. Conclusions and Future Work

This paper proposes an approach to address short to medium term horizons in multiproduct pipeline scheduling problems with inventory management. The approach uses an iterative procedure combined with sequential valid integer cuts that explore the solution space at a low detail level model to provide initial data for a higher detail model. With this strategy it is possible to tackle longer pipeline time horizons and it is possible to explore two levels of detail and a set of solutions rather than having a single time consuming solution. This approach meets real world solutions requirements: lower computational effort and a valid set of solutions for comparison and selection. Additionally, when comparing with a real world scenario where schedulers use trial and error methods based in spreadsheets, the improvements are significant.

As future work, the authors propose to study the influence of the diversity degree between consecutive iterations, a selection of number of iterations criteria as well as the exchange of more information between models, such as pumping time windows per batch. The authors also propose to analyze how the approach behaves for higher time horizons, up to one month of time span, in order to provide pipeline requirements for the refinery monthly production planning.

6. Acknowledgments

The authors gratefully acknowledge the case study and the financial support provided by Companhia Logística de Combustíveis.

References

S.N. Boschetto, L.C. Felizari, L. Yamamoto, L. Magatão, S.L. Stebel, F. Neves Jr., L.V.R. Arruda, R. Lüders, P.C. Ribas, L.F.J. Bernardo, 2008, An Integrated Framework for Operational Scheduling of a Real-World Pipeline Network, In: Bertrand Braunschweig; Xavier Joulia. (Editors), Computer Aided Process Engineering, Elsevier , vol 25, 259;

D.C. Cafaro, J. Cerdá, 2008, Dynamic scheduling of multiproduct pipelines with multiple delivery due dates, Comp. & Chem. Eng., 32, 728;

L. Magatão, L.V.R. Arruda, F. Neves, Jr, 2004, A mixed integer programming approach for scheduling commodities in a pipeline, Comp. & Chem. Eng., 28, 171;

R. Rejowski Jr, J.M. Pinto, 2008, A continuous time representation for scheduling of pipeline systems with pumping yield rate constraints, Comp. & Chem. Eng., 32, 1042;

S. Relvas, H.A. Matos, A.P.F.D. Barbosa-Póvoa, J. Fialho, A.S. Pinheiro, 2006, Pipeline scheduling and inventory management of a multiproduct distribution oil system, Ind. Eng. Chem. Res., 45, 7841

S. Relvas, A.P.F.D. Barbosa-Póvoa, H.A. Matos, 2009, Heuristic batch sequencing on a multiproduct oil distribution system, Comp. & Chem. Eng., 33, 712.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Efficient Bulk Maritime Logistics for the Supply and Delivery of Multiple Chemicals

Jie Li,[a] I. A. Karimi,[a] Rajagopalan Srinivasan[a,b]

[a]*Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore, 117576*
[b]*Process Sciences and Modeling, Institute of Chemical and Engineering Science, 1 Pesek Road, Jurong Island, Singapore 627833*

Abstract

In this paper, we develop a novel unit-slot mixed integer linear programming (MILP) continuous-time model for a ship planning problem, in which a heterogeneous fleet of ships is used to supply and deliver multiple chemical products between production and consumption sites. We incorporate many real-life features such as multiple pickups and deliveries of multiple products using ships with multiple dedicated compartments, variable load/discharge quantities, multiple jetties, and fixed setup time for each product, etc. Furthermore, our model ensures product inventories at any time to meet their minimum and maximum capacities. Four examples are solved to illustrate the superiority and efficiency of our proposed formulation.
Keywords: IPDP, Mixed-integer linear programming (MILP), bulk product transportation, unit slot.

1. Introduction

Maritime transportation is very critical to the world economy, because approximately 90% by volume and 70% by value of all goods are transported worldwide by sea (Psaraftis, 1998). Moreover, approximately 80% of all maritime ton-miles is via bulk transportation, which offers the lowest cost per ton-mile. Therefore, operational efficiency of maritime transportation can reduce final product costs significantly.

In this paper, we address a ship planning problem, in which a heterogeneous fleet of ships with dedicated compartments is used to transport multiple chemicals between production and consumption sites. This is done while ensuring that adequate inventory levels of all chemicals are maintained at all consumption sites. This problem is termed as Inventory Pickup and Delivery Problem (IPDP), which has been addressed previously by Christiansen (1999), Al-Khayyal and Hwang (2007) and Li et al. (2008). However, they all have some limitations. For instance, the model of Christiansen (1999) lacked many realistic operation features such as multiple dedicated compartments, multiple chemicals, variable production/consumption rates, multiple jetties, setup times, etc. Al-Khayyal and Hwang (2007) did not allow several ships to load/discharge the same product at each site simultaneously. Moreover, they could not ensure that product inventories were maintained between allowable limits at any time. Lastly, Li et al. (2008) did not incorporate setup times and were unable to solve large-scale problems.

In this paper, we develop a novel unit-slot mixed integer linear programming (MILP) continuous-time formulation for this chemical logistics problem and incorporate many real-life features such as multiple pickups and deliveries of multiple products using ships with multiple dedicated compartments, variable load/discharge quantities, product load/discharge sequences, variable production/consumption rates, multiple jetties, and

fixed setup time for each product. Most importantly, our model ensures product inventories at any time to meet their minimum and maximum capacities and allows several ships to load/discharge the same product at each site simultaneously. Furthermore, one site can be visited several times by a ship. Four examples are solved to illustrate the efficiency of our proposed formulation.

2. Problem Statement

Most multinational companies have plants distributed all over the world. Some of these may produce chemical products that are raw materials to other plants, while others may simply consume products. Thus, a plant can be a production site, consumption site, or both. If a site does not have sufficient inventory of a product, then it may have to purchase from external suppliers in the local market. Thus, we define two types of sites. Internal site belongs to a company, while an external site does not. We consider inventory at internal sites only. Consider I (i = 1, 2, 3, …, I) sites, B (b = 1, 2, 3, …, B) jetties, and M (m = 1, 2, …, M) products in exactly M dedicated product tanks. A heterogeneous fleet of ships V (v = 1, 2, …, V) with multiple dedicated compartments is used to load a product from a production site and discharge it at a consumption site to ensure enough inventory at each site. Figure 1 shows an example transport scenario. With this, the problem can be stated as follows.

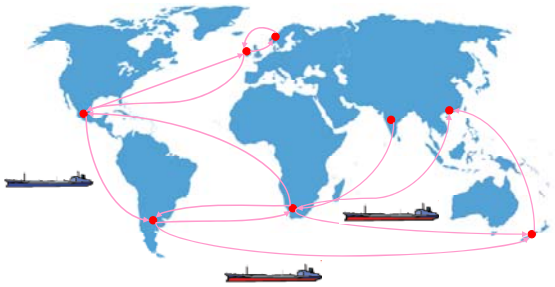


Figure 1 An example transport scenario

Given:
1. V ships, products that they can carry, minimum and maximum capacities for these products, initial position, initial products and their amounts, load and discharge rates.
2. I Sites, products that they can produce or consume, jetties, initial holdups, minimum and maximum capacities of products, limits on production or consumption rates.
3. Ship travel times between sites, and setup times for switching from a product to another.
4. Ship travel costs between sites, fixed cost for product load or discharge at each site.

Decide:
1. Which ship should load or discharge which product at which site, at which time, and in what amounts
2. The inventory profile of each product at each site.

3. MILP Formulation

We divide the scheduling horizon [0, H] into K (k = 1, 2, …, K) contiguous slots (Figure 2) and denote the time before the horizon start by slot zero (k = 0). We define ships, jetties, and product tanks as various units. Each unit has K slots, which are not synchronized (Lim & Karimi, 2003) across the units. In other words, they are unit slots (Liu & Karimi, 2008), where the start/end times and slot lengths of a given slot k need

not be the same across all units. We let T_{qk} (k = 0, 1, 2, …, K; $T_{q0} \geq 0$, $T_{qK} \leq H$) denote the end time of slot k on unit q, where q becomes v for a ship, becomes b for a jetty, and becomes m for a product tank. Slot k on unit q starts at $T_{q(k-1)}$ and ends at T_{qk}. Since the slots are asynchronous, T_{vk}, T_{bk}, and T_{imk} may vary with units. Thus,

$$T_{qk} \geq T_{q(k-1)} \qquad 1 \leq k \leq K \qquad (1)$$

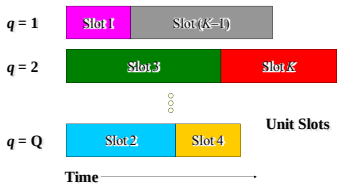


Figure 2 Schematic of unit slots design

In the following, $\mathbf{B}_i$ is the set of jetties that belong to site i, $\mathbf{M}_i$ is the set of products that can be produced or consumed at site i, $\mathbf{M}_v$ is the set of products that can be carried by ship v, $\mathbf{V}_i$ is the set of ships that can visit site i, $\mathbf{I}_v$ is the set of sites that ship v can visit, and $\mathbf{IT}_i$ is the set of internal sites.

3.1. Ship Load or Discharge Operations

We use two dummy sites (i = 0, I + 1) to denote the initial and final positions of a ship respectively. Thus, we have I+2 sites including I real site (i = 1, 2, …, I) and two dummy (i = 0, I + 1) sites. Suppose a ship v is loading or discharging products in its own slot k, whereas a product tank m is receiving from or feeding this v in its own slot k'. Three scenarios are possible: $k' < k$, $k' = k$, and $k' > k$. For $k' < k$, we can simply introduce additional slots on product tank i to make $k' = k$. For $k' > k$, we can do the same on ship v. In other words, with no loss of generality, we can demand that if a ship v is loading or discharging products from or to a product tank m at any time, then the unit slots corresponding to that time on both ship v and product tank m must have the same index. The same holds true for a ship and a jetty. Thus, we now define binary variables X_{ikv}, Y_{mkv} and y_{bkv} and 0-1 continuous variables ye_{mkv} and Z_{ijkv}.

$$X_{ikv} = \begin{cases} 1 & \text{If ship } v \text{ visits site } i \text{ during slot } k \\ 0 & \text{Otherwise} \end{cases} \qquad 0 \leq i \leq I+1,\ 0 \leq k \leq K$$

$$Y_{mkv} = \begin{cases} 1 & \text{If ship } v \text{ loads/discharges product } m \text{ during slot } k \\ 0 & \text{Otherwise} \end{cases} \qquad 0 < k \leq K$$

$$y_{bkv} = \begin{cases} 1 & \text{If ship } v \text{ loads/discharges via jetty } b \text{ in slot } k \\ 0 & \text{Otherwise} \end{cases} \qquad 0 < k \leq K$$

$$ye_{mkv} = \begin{cases} 1 & \text{If ship } v \text{ ends load/discharge product } m \text{ at the end of slot } k \\ 0 & \text{Otherwise} \end{cases} \qquad 0 < k < K$$

$$Z_{ijkv} = \begin{cases} 1 & \text{If ship } v \text{ visits site } i \text{ in slot } k \text{ and } j \text{ in } k+1 \\ 0 & \text{Otherwise} \end{cases} \qquad 0 \leq i, j \leq I+1,\ 0 \leq k < K$$

We treat X_{ikv} (i = 0, 1, 2, 3, …, I) as binary and $X_{(I+1)kv}$ as 0-1 continuous variable. Thus, at any time, a ship v visits exactly one site,

$$\sum_{i \in \mathbf{I}_v} X_{ikv} = 1 \qquad 0 < k \leq K \qquad (2)$$

At most one ship can load or discharge through one jetty b in site i at a time. Then,

$$\sum_{v \in \mathbf{V}_i} y_{bkv} \le 1 \qquad b \in \mathbf{B}_i,\ 1 \le i \le I,\ 0 < k \le K \qquad (3)$$

3.2. Slot Timings on Product Tanks

When using unit slots in the presence of shared resources such as inventories, the main challenge is to relate the timings of different units that share the same resource. The flow in/out of a resource must be ordered chronologically, so that we can get a correct resource profile. If a ship v is loading or discharging product m during slot k, then the start (end) of a slot k on a product tank m must precede (succeed) the start (end) of slot k on the ship.

$$T_{im(k-1)} \le T_{v(k-1)} + H(2 - X_{ikv} - Y_{mkv})$$

$$i \in \mathbf{IT}_i,\ v \in \mathbf{V}_i,\ m \in \mathbf{M}_v \ni \mathbf{M}_i,\ 1 \le i \le I,\ 0 < k \le K \qquad (4a)$$

$$T_{imk} \ge T_{v(k-1)} + RU_{imkv} - H(2 - X_{ikv} - Y_{mkv})$$

$$i \in \mathbf{IT}_i,\ v \in \mathbf{V}_i,\ m \in \mathbf{M}_v \ni \mathbf{M}_i,\ 1 \le i \le I,\ 0 < k \le K \qquad (4b)$$

For checking the inventory at the end of each slot, we define an intermediate point t_{imk} between $T_{im(k-1)}$ and T_{imk}. Then, we demand that the end of each loading or discharging in slot k must match with this point on the tank by using the following.

$$T_{v(k-1)} + RU_{imkv} \le t_{imk} + H(2 - X_{ikv} - Y_{mkv})$$

$$i \in \mathbf{IT}_i,\ v \in \mathbf{V}_i,\ m \in \mathbf{M}_v \ni \mathbf{M}_i,\ 1 \le i \le I,\ 0 < k \le K \qquad (5a)$$

$$T_{v(k-1)} + RU_{imkv} \ge t_{imk} - H(2 - X_{ikv} - Y_{mkv})$$

$$i \in \mathbf{IT}_i,\ v \in \mathbf{V}_i,\ m \in \mathbf{M}_v \ni \mathbf{M}_i,\ 1 \le i \le I,\ 0 < k \le K \qquad (5b)$$

3.3. Site Inventory

The timing constraints of section 3.2 enable us to write the following inventory balances for product tanks.

$$Iv_{imk} = Iv_{im(k-1)} - JJ_{im} \sum_{v \in \mathbf{V}_i} qv_{imkv} + JJ_{im} SQ1_{imk} \qquad i \in \mathbf{IT}_i,\ m \in \mathbf{M}_i,\ 1 \le i \le I,\ 0 < k \le K \qquad (6a)$$

$$Iv_{im}^L \le Iv_{im(k-1)} - JJ_{im} \sum_{v \in \mathbf{V}_i} qv_{imkv} + JJ_{im} SQ2_{imk} \le Iv_{im}^U$$

$$i \in \mathbf{IT}_i,\ m \in \mathbf{M}_i,\ 1 \le i \le I,\ 0 < k \le K \qquad (6b)$$

where, Iv_{imk} ($Iv_{im}^L \le Iv_{imk} \le Iv_{im}^U$) is the inventory of product m at site i at the end of slot k. $SQ1_{imk}$ is the amount of product m produced/consumed at site i in slot k, and $SQ2_{imk}$ is the amount of product m at site i produced/consumed from $T_{im(k-1)}$ to t_{imk}. JJ_{im} is 1 if site i produces m, otherwise −1. The expressions for $SQ1_{imk}$ and $SQ2_{imk}$ are not given here.

3.4. Objective Function

Our objective is to minimize total operating cost over the planning horizon, which consists of traveling and load/discharge costs.

$$\text{Min } TC = \sum_{v=1}^{V} \sum_{i \in \mathbf{I}_v} \sum_{j \in \mathbf{I}_v \ni j \ne i} \sum_{k=0}^{K} C_{ijv} Z_{ijkv} + \sum_{v=1}^{V} \sum_{m \in \mathbf{M}_v} \sum_{k=1}^{K} C_m ye_{mkv} + \sum_{v=1}^{V} \sum_{m \in \mathbf{M}_v} C_m Y_{mKv} \qquad (7)$$

where, C_{ijv} is the travelling cost of ship v from sites i to j. C_m is the load/discharge cost of product m.
So far, our complete model comprises eqs. 1-7 and other equations such as slot timings on jetties, inventory balance at the end of the planning horizon, etc. that are given above.

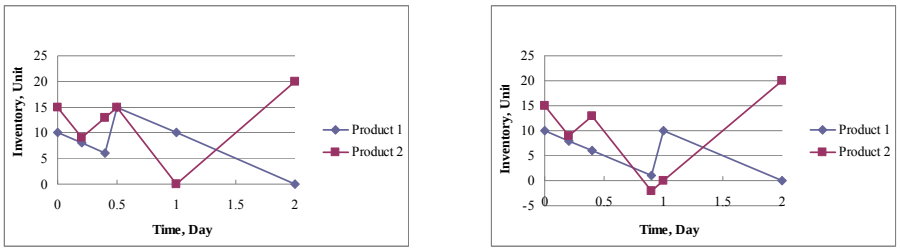
4. Results and Discussions

We solve four examples to evaluate our proposed formulation. These examples vary from ships, sites, products, jetties, etc. and are solved using CPLEX 10.0.1/GAMS 22.2 on a Dell workstation PWS690 (Intel® Xeon™ CPU 3.00 GHZ, 16 GB memory) running Windows XP. Table 1 shows solution statistics for Examples 1-4. Note that the models of Christiansen (1999) and Al-Khayyal and Hwang (2007) need the estimated arrival times of each site. We use the same estimated arrival times for all sites. To choose the best arrival times, we start from a small number of arrival times, and then increase by 1 until the objective function does not change. This approach is the same as that of choosing the best slot.

Table 1 Solution statistics of various models for Examples 1-4

Example	Model	Discrete Variables	Continuous Variables	Constraints	CPU Time (Seconds)	MILP Solution	Relative Gap (%)
1	Ours	40	359	592	0.73	7	0
	Li et al. (2008)	120	831	1050	1.77	7	0
	Al-Khayyal and Hwang (2007)	158	161	719	N/A	N/A	N/A
2	Ours	35	485	642	0.78	32.5	0
	Li et al. (2008)	106	967	905	1.84	32.5	0
	Christiansen (1999)	336	177	1256	3.58	32.5	0
3	Ours	415	2659	3369	706.8	22	0
	Li et al. (2008)	597	4861	4301	86400	28	14.03
4	Ours	105	1092	1101	136.3	50.4	0
	Christiansen (1999)	952	341	3317	354.3	50.4	0

N/A: The obtained solution is infeasible



(a) (b)

Figure 3 Inventory profiles of products 1 and 2 at site 1 for Example 1 with different product load/discharge sequences: (a) first load/discharge product 1 (b) first load/discharge product 2

Example 1

This example involves two ships with two products, and three sites, each of which has two jetties. While site 1 consumes product 1 and produces product 2, sites 2 and 3 both consume product 2 and produce product 1. The capacities of products 1 and 2 at sites 1 are [0, 12 unit] and [0, 21 unit]. Figure 3 gives the optimal inventory profiles of products 1 and 2 at site 1 from the model of Al-Khayyal and Hwang (2007). While the inventory of product 1 at 0.5 day is 15 unit in case (a), which violates its maximum capacity (12 unit), that of product 2 at 0.9 day is -2 unit in case (b), which also violates its minimum capacity (0). Hence, the model of Al-Khayyal and Hwang (2007) cannot guarantee the inventories of products to meet their minimum and maximum capacities

throughout the entire planning horizon. The model of Christiansen (1999) cannot be used to solve this example because it is restricted to single product. The model of Li et al. (2008) and our model obtain the optimal solution of 7. While the model of Li et al. (2008) needs 1.77 CPU s, our model needs only 0.73 CPU s. No violation on inventory capacities occurs through the entire horizon with our model.

Example 2

This Example is taken from Li et al. (2008) involving two ships, five sites and one single product with 16-day planning horizon. We obtain the optimal solution of 32.5 within 0.78 CPU Seconds, while the model of Li et al. (2008) needs 1.84 CPU seconds and the model of Christiansen (1999) needs 3.58 CPU s. When only one product is involved, the model of Al-Khayyal and Hwang (2007) is the same as that of Christiansen (1999).

Examples 3-4

Example 3 consists of five ships, eight sites, and two products. The planning horizon is 80 days. Example 4 involves five ships, five sites, and one product with 60-day planning horizon. For Example 3, our model obtains the optimal solution of 22 within 706.8 CPU s. However, the model of Li et al. (2008) gets suboptimal solution of 28 with 14.03% relative gap after 24 hours. Since the model of Christiansen (1999) is only for single product and the model of Al-Khayyal and Hwang (2007) may violate inventory capacities when two or more products are involved, they are not used to solve this example. For Example 4, our model obtains the optimal solution of 50.4 within 136.3 CPU s. However, the model of Christiansen (1999) needs 354.3 CPU s to obtain the optimal solution.

5. Conclusion

In this paper, we developed an efficient unit-slot MILP formulation for a ship planning problem. We incorporated many more real-life features compared to the models of Christiansen (1999) and Al-Khayyal (2007). Furthermore, our model ensured product inventories to meet their minimum and maximum capacities at any time during the planning horizon. On 4 test problems of varying sizes, our model performed better than those of Li et al. (2008) and Christiansen (1999).

Acknowledgment

The authors would like to acknowledge financial support for this work from The Agency for Science, Technology, and Research (A*Star) under grant 052 116 0074.

References

F. Al-Khayyal, S. J. Hwang, 2007, Inventory constrained maritime routing and scheduling for multi-commodity liquid bulk, Part I: Applications and model, *European J. of Operational Research*, 176, 106-130.

M. Christiansen, 1999, Decomposition of a combined inventory and time constrained ship routing problem, *Transportation Science*, 33, 1, 3-16.

J. Li, I. A. Karimi, R. Srinivasan, 2008, Supply and distribution of multiple products via bulk maritime logistics, *5th Foundations of Computer-Aided Process Operations- FOCAPO*, Cambridge, Massachusetts, USA, June 29- July 2.

M. F. Lim, I. A. Karimi, 2003, Resource-constrained scheduling of parallel production lines using asynchronous slots, *Ind. Eng. Chem. Res.*, 42, 6832-6842.

Y. Liu, I. A. Karimi, 2008, Scheduling multistage batch plants with parallel units and no interstage storage, *Comput Chem Eng.*, 32, 671-693.

H. N., Psaraftis, 1999, Forward to focused issue on maritime transportation, *Transportation Science*, 33, 1-2.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

MINLP Model and Algorithms for Optimal Design of Large-Scale Supply Chain with Multi-Echelon Inventory and Risk Pooling under Demand Uncertainty

Fengqi You, Ignacio E. Grossmann

Department of Chemical Engineeirng, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract

We address the optimal design of a multi-echelon supply chain and the associated inventory systems in the presence of uncertain customer demands. By using the guaranteed service approach to model the multi-echelon stochastic inventory system, we develop an optimization model for simultaneously optimizing the transportation, inventory and network structure of a multi-echelon supply chain. We formulate this problem as an MINLP with a nonconvex objective function including bilinear, trilinear and square root terms. By exploiting the properties of the basic model, we reformulate the problem as a separable concave minimization program. A spatial decomposition algorithm based on Lagrangean relaxation and piecewise linear approximation is proposed to obtain near global optimal solutions with reasonable computational expense. Examples for industrial gas supply chains with up to 5 plants, 50 potential distribution centers and 100 markets are presented.

Keywords: Supply Chain, Safety Stock, Risk-pooling, Uncertainty, MINLP.

1. Introduction

Due to the increasing pressure for remaining competitive in the global market place, optimizing inventories across the supply chain has become a major challenge for the process industries to reduce costs and to improve the customer service (Grossmann, 2005). This challenge requires integrating inventory management with supply chain network design, so that decisions on the locations to stock the inventory and the associated amount of inventory in each stocking location can be determined simultaneously for lower costs and higher customer service level. However, the integration is usually nontrivial for multi-echelon supply chains and their associated inventory systems in the presence of uncertain customer demands (Zipkin, 2000).

The objective of this work is to develop optimization models and solution algorithms to address the problem of joint multi-echelon supply chain network design and inventory management. By using the guaranteed service approach to model the multi-echelon inventory system (Graves & Willems, 2005) we capture the stochastic nature of the problem, and develop an equivalent deterministic optimization model. The model determines the supply chain design decisions such as the locations of distribution centers (DCs), assignments of markets to DCs, assignments of DCs to plants, shipment levels from plants to the DCs and from DCs to customers, and inventory decisions such as pipeline inventory and safety stock in each node of the supply chain network. The

model also captures risk-pooling effects (Eppen, 1979) by consolidating the safety stock inventory of downstream nodes to the upstream nodes in the multi-echelon supply chain. The model is first formulated as a mixed-integer nonlinear program (MINLP) with a nonconvex objective function, and then reformulated as a separable concave minimization program. To solve the problem efficiently, a decomposition algorithm based on Lagrangean relaxation and piece-wise linear approximation is developed to obtain near global optimal solutions within 1% optimality gap with modest CPU times. Examples are presented to illustrate the application of the model and its performance.

2. Problem Statement

We are given a potential supply chain consisting of a set of plants (or suppliers), a number of candidate sites for distribution centers, and a set of customer demand zones whose inventory costs should be taken into account. The market can represent a local distributor, a regional warehouse, a dealer, a retailer, or a wholesaler. Alternatively, one might view the customer demand as the aggregation of a group of customers operated with vendor managed inventory, which is a common business model in the industrial gases industry and some chemical companies.

In the given potential supply chain, the locations of the plants, potential distribution centers and markets are known and the distances between them are given. The investment costs for installing DCs are expressed by a cost function with fixed charges. Each market k has an uncorrelated normally distributed demand with mean μ and variance σ in each unit of time. Single sourcing restriction, which is common in the industries of specialty chemical and industrial gases, is employed for the distribution from plants to DCs and from DCs to markets. That is, each DC is only served by one plant, and each market is only assigned to one DC to satisfy the demand. Linear transportation costs are incurred for shipments from plant i to distribution center j with unit cost, and from distribution center j to market k with unit cost. The corresponding deterministic order processing times of DCs and market that includes the material handling time, transportation time and inventory review period. The service time of each plant, and the maximum service time of each markets are known. We are also given the safety stock factor for DCs and markets, and, which correspond to the standard normal deviate of the maximum amount of demand that the node will satisfy from its safety stock. A common review period is used for the control of inventory in each node. Inventory costs are incurred at DCs and markets, and consist of pipeline inventory and safety stock, of which the unit costs are given. The objective is to determine how many distribution centers (DCs) to install, where to locate them, which plants to serve each DC and which DCs to serve each market, how long should each DC quote its service time, and what level of safety stock to maintain at each DC and market so as to minimize the total installation, transportation, and inventory costs.

3. Model Formulation

The joint multi-echelon supply chain design and inventory management model is a mixed-integer nonlinear program (MINLP) that deals with the supply chain network design for a given product, and considers its multi-echelon inventory management.

3.1. Objective Function

The objective function of this model (the total supply chain design cost) is given by,

$$\begin{aligned}\min: \quad & \sum_{j\in J} f_j Y_j + \sum_{j\in J}\left(g_j \sum_{k\in K} \chi Z_{jk}\mu_k \right) + \sum_{i\in I}\sum_{j\in J}\left(c1_{ij} X_{ij} \sum_{k\in K} \chi Z_{jk}\mu_k \right) + \sum_{j\in J}\sum_{k\in K} c2_{jk}\chi Z_{jk}\mu_k \\ & + \sum_{i\in I}\sum_{j\in J}\left(\theta 1_j t1_{ij} X_{ij} \sum_{k\in K} Z_{jk}\mu_k \right) + \sum_{j\in J}\sum_{k\in K} \theta 2_k t2_{jk} Z_{jk}\mu_k + \sum_{j\in J} \lambda 1_j h1_j \sqrt{N_j} \cdot \sqrt{\sum_{k\in K} \sigma_k^2 Z_{jk}} + \sum_{k\in K} \lambda 2_k h2_k \cdot \sigma_k \sqrt{L_k}\end{aligned} \tag{1}$$

which includes the following items:

DC Installation Cost: The cost of installing a DC in candidate location *j* is expressed by a fixed-charge cost model that captures the economies of scale in the investment. The annual expected demand of DC *j* is ($\sum_{k\in K} \chi Z_{jk}\mu_k$), which equals to the annual mean demand of all the markets served by DC *j*. Hence, the cost of installing DC *j* consists of fixed cost f_j and variable cost ($g_j \sum_{k\in K} \chi Z_{jk}\mu_k$), which is the product of variable cost coefficient and the expected demand of this DC in one year. Thus, the total installation cost of all the DCs is given by $\sum_{j\in J} f_j Y_j + \sum_{j\in J}\left(g_j \sum_{k\in K} \chi Z_{jk}\mu_k \right)$.

Transportation costs from plants to DCs and from DCs to markets: The product of the annual mean demand of DC *j* and the unit transportation cost ($\sum_{i\in I} c1_{ij} X_{ij}$) between DC *j* and the plant that serves it yields the annual plant to DC transportation cost, $\sum_{i\in I}\sum_{j\in J}\left(c1_{ij} X_{ij} \sum_{k\in K} \chi Z_{jk}\mu_k \right)$. Similarly, the product of yearly expected mean demand of market *k* ($\chi\mu_k$) and the unit transportation cost ($\sum_{j\in J} c2_{jk} Z_{jk}$) between market *k* and the DC that serves it, yields the corresponding transportation cost, $\sum_{j\in J}\sum_{k\in K} c2_{jk}\chi Z_{jk}\mu_k$.

Pipeline inventory costs in DCs and markets: Based on Little's law, the pipeline inventory PI_j of DC *j* equals to the product of its daily mean demand ($\sum_{k\in K} Z_{jk}\mu_k$) and its order processing time ($\sum_{i\in I} t1_{ij} X_{ij}$), which is in terms of days. Thus, the annual total pipeline inventory cost of all the DCs is given by, $\sum_{i\in I}\sum_{j\in J}\left(\theta 1_j t1_{ij} X_{ij} \sum_{k\in K} Z_{jk}\mu_k \right)$, where $\theta 1_j$ is the annual unit pipeline inventory cost of DC *j*. Similarly, the total annual pipeline inventory cost of all the markets is given by, $\sum_{j\in J}\sum_{k\in K} \theta 2_k t2_{jk} Z_{jk}\mu_k$, where $\theta 2_k$ is the annual unit pipeline inventory cost of market *k*.

Safety stock costs in DCs and markets: The demand at market *k* follows a given normal distribution with mean μ_k and variance σ_k^2. Due to the risk-pooling effect (Eppen, 1979), the demand over the net lead time (N_j) at DC *j* is also normally distributed with a mean of $N_j \sum_{k\in J_k} \mu_k$ and a variance of $N_j \sum_{k\in J_k} \sigma_k^2$, where J_k is the set of markets *k* assigned to DC *j*. Thus, the safety stock required in the DC at candidate location *j* with a safety stock factor $\lambda 1_j$ is $\lambda 1_j \sqrt{N_j} \cdot \sqrt{\sum_{k\in K} \sigma_k^2 Z_{jk}}$. Considering the annual inventory holding cost at DC *j* is $h1_j$, we have the annual total safety stock cost at all the DCs equals to, $\sum_{j\in J} \lambda 1_j h1_j \sqrt{N_j} \cdot \sqrt{\sum_{k\in K} \sigma_k^2 Z_{jk}}$. Similarly, the demand over the net lead time of markets *k* (L_k) is normally distributed with a mean of $L_k\mu_k$ and a variance of $L_k\sigma_k^2$. Thus, the annual safety stock cost at all the markets is given by, $\sum_{k\in K} \lambda 2_k h2_k \cdot \sigma_k \sqrt{L_k}$.

3.2. Constraints

Three constraints are used to define the network structure. The first one is that if DC *j* is installed, it should be served by only one plant. If it is not installed, it is not assigned to any plant. This can be modelled by,

$$\sum_{i\in I} X_{ij} = Y_j, \ \forall j \tag{2}$$

The second constraint states that each market *k* is served by only one DC,

$$\sum_{j\in J} Z_{jk} = 1, \ \forall k \tag{3}$$

The third constraint states that if a market *k* is served by the DC in candidate location j, the DC must exist,

$$Z_{jk} \le Y_j, \ \forall j,k \tag{4}$$

Two constraints are used to define the net lead time of DCs and markets. The replenishment lead time of DC *j* should be equal to the guaranteed service time (SI_i) of plant *i*, which serves DC *j*, plus the order processing time (t_{ij}). Since each DC is served by only one plant, the replenishment lead time of DC *j* is given by $\sum_{i\in I}(SI_i + t_{ij})\cdot X_{ij}$. Thus, the net lead time of DC *j* should be greater than its replenishment lead time minus its guaranteed service time to its successor markets, given by the linear inequality,

$$N_j \ge \sum_{i\in I}(SI_i + t1_{ij})\cdot X_{ij} - S_j, \ \forall j \tag{5}$$

Similarly, the net lead time of a market *k* is greater than its replenishment lead time minus its maximum guaranteed service time, R_k, is given by the nonlinear inequalities,

$$L_k \ge \sum_{j\in J}(S_j + t2_{jk})\cdot Z_{jk} - R_k, \ \forall k \tag{6}$$

Finally, all the decision variables for network structure are binary variables, and the variables for guaranteed service time and net lead time are non-negative variables.

$$X_{ij}, Y_j, Z_{jk} \in \{0,1\}, \ \forall i,j,k \tag{7}$$

$$S_j \ge 0, \ N_j \ge 0, \ \forall j \tag{8}$$

$$L_k \ge 0, \ \forall k \tag{9}$$

3.3. Reformulation

The original model is a nonconvex MINLP. To reduce the computational efforts, we use linearization techniques to reformulate the model as a separable concave minimization problem. The reformulation model (AP) has a new objective function given as follows.

$$\min \ \sum_{j\in J} f_j Y_j + \sum_{i\in I}\sum_{j\in J}\sum_{k\in K} A_{ijk} XZ_{ijk} + \sum_{j\in J}\sum_{k\in K} B_{jk} Z_{jk} + \sum_{j\in J} q1_j \sqrt{NZV_j} + \sum_{k\in K} q2_k \sqrt{L_k} \tag{10}$$

In addition to linear constraints (2), (3), (4), (7), (8), (9), the reformulated model includes the following linear constraints for linearization the product of a binary variable and a continuous variable (Glover, 1975):

$$N_j \ge \sum_{i\in I} \overline{S}_{ij}\cdot X_{ij} - S_j, \ \forall j \tag{11}$$

$$XZ_{ijk} \le X_{ij}, \ \forall i,j,k \tag{12}$$

$$XZ_{ijk} \le Z_{jk}, \ \forall i,j,k \tag{13}$$

$$XZ_{ijk} \ge X_{ij} + Z_{jk} - 1, \ \forall i,j,k \tag{14}$$

$$SZ_{jk} + SZ1_{jk} = S_j, \ \forall j,k \tag{15}$$

$$SZ_{jk} \le Z_{jk}\cdot S_j^U, \ \forall j,k \tag{16}$$

$$SZ1_{jk} \le (1 - Z_{jk})\cdot S_j^U, \ \forall j,k \tag{17}$$

$$NZ_{jk} + NZ1_{jk} = N_j, \quad \forall j,k \tag{18}$$

$$NZ_{jk} \le Z_{jk} \cdot N_j^U, \quad \forall j,k \tag{19}$$

$$NZ1_{jk} \le (1 - Z_{jk}) \cdot N_j^U, \quad \forall j,k \tag{20}$$

$$NZV_j = \sum_{k \in K} \sigma_k^2 \cdot NZ_{jk}, \quad \forall j \tag{21}$$

$$L_k \ge \sum_{j \in J} SZ_{jk} + \sum_{j \in J} t2_{jk} \cdot Z_{jk} - R_k, \quad \forall k \tag{22}$$

$$XZ_{ijk} \ge 0, \; SZ_{jk} \ge 0, \; SZ1_{jk} \ge 0, \; NZ_{jk} \ge 0, \; NZ1_{jk} \ge 0, \; NZV_j \ge 0, \; \forall i,j,k \tag{23}$$

4. Solution Algorithm

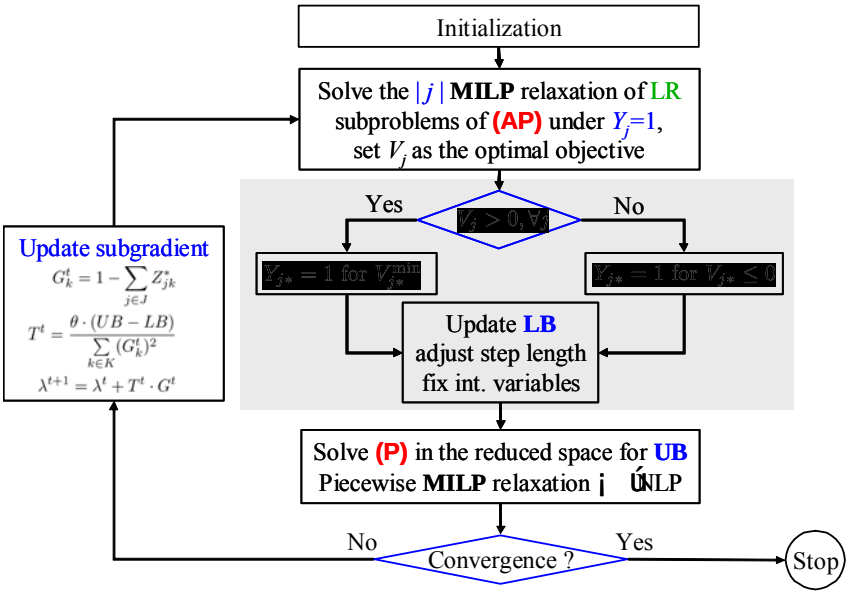


Figure 1. Lagrangean Relaxation Algorithm

To effectively solve the proposed MINLP model, a global optimization algorithm based on Lagrangean relaxation is developed. The algorithm flowchart is given in Figure 1. The basic idea of this algorithm is to consider the alternative formulation (AP) of the model. Next, we dualize the assignment constraint (10) to allow decomposing the entire problem based on DC *j*. To solve each subproblem, we used piece-wise linear approximation to underestimate the square root terms. We should note that the entire solution algorithm requires at least an MILP solver; the NLP solver is not required. Due to the duality gap, this algorithm stops after a finite number of iterations. As will be shown in the computational results, the dual gaps are quite small.

5. Results

To illustrate the application of the proposed model, we consider an industrial gas supply chain (liquid oxygen-LOX) with two plants, three potential DCs and six customers. The associated superstructure, as well as the optimal network structures with and without considering inventory cost, is given in Figure 2. The results show that although inventory cost only make up less than 20% of the total cost, it is necessary to take into account in the supply chain design. Comparison of the performance of the proposed

algorithms and commercial MINLP and global optimizers for medium and large scale instances (5 plants, 50 DCs, 150 markets) are given in Table 1. The advantage of using the Lagrangean relaxation algorithm for solving the proposed model can be clearly seen.

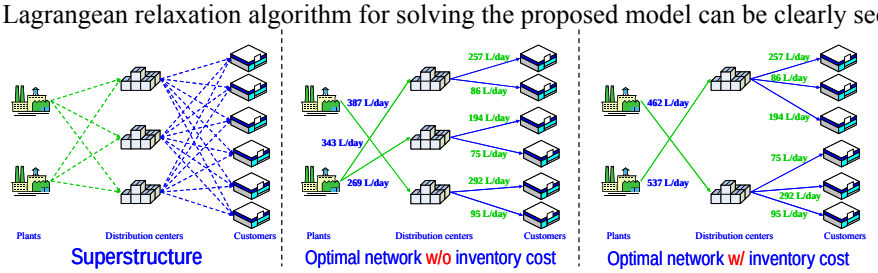


Figure 2. Optimal network structure for the LOX supply chain

Table 1. Comparison of the performance of the algorithms for medium and large scale instances

i	*j*	*k*	Solve (P0) directly with BARON				Solve (P2) with CPLEX for at most 1 hour, then solve (P1) with DICOPT or SBB				Lagrangean Relaxation Algorithm				
							DICOPT		SBB						
			Solution	LB	Gap	Time (s)	Solution	Time(s)	Solution	Time(s)	Solution	Global LB	Global Gap	Time (s)	Iter.
2	20	20	1,889,577	1,159,841	62.92%	36,000	1,820,174	140.3	1,813,541	163.6	1,776,969	1,775,957	0.06%	175.0	11
5	30	50	---*	---*	---*	36,000	---**	36,000	---**	36,000	4,417,353	4,403,582	0.31%	3,279	24
10	50	100	---*	---*	---*	36,000	---**	36,000	---**	36,000	7,512,609	7,477,584	0.47%	27,719	42
20	50	100	---*	---*	---*	36,000	---**	36,000	---**	36,000	5,620,045	5,576,126	0.79%	27,748	53
3	50	150	---*	---*	---*	36,000	---**	36,000	---**	36,000	12,291,296	12,276,483	0.12%	16,112	32

* No solution or bounds were returned due to solver failure.; ** No solution was returned after 10 hours

6. Conclusion

In this paper, we present an MINLP model that determines the optimal network structure, transportation and inventory levels of a multi-echelon supply chain with the presence of customer demand uncertainty. The well-known guaranteed service approach is used to model the multi-echelon inventory system. The risk pooling effect is also taken into account in the model by consolidating the demands in the downstream nodes to their upstream nodes. To solve the resulting MINLP problem effectively for large scale instances, a decomposition algorithm, based on Lagrangean relaxation and piecewise linear approximation was proposed. Computational experiments on large scale problems show that the proposed algorithm can obtain global or near-global optimal solutions (typically within 1% of the global optimum) in modest computational expense without the need of a global optimizer.

References

G. Eppen, 1979, Effects of centralization on expected costs in a multi-echelon newsboy problem. Management Science, 25, (5), 498.

M. L. Fisher, 1985, An application oriented guide to Lagrangian relaxation. Interfaces 1985, 15, (2), 2.

F. Glover, 1975, Improved Linear Integer Programming Formulations of Nonlinear Integer Problems. Management Science, 22, (4), 455.

S. C. Graves, S. P. Willems, 2005, Optimizing the supply chain configuration for new products. Management Science, 51, (8), 1165.

I. E. Grossmann, I. E., 2005, Enterprise-wide Optimization: A New Frontier in Process Systems Engineering. AIChE Journal, 51, 1846.

P. H. Zipkin, 2000, Foundations of Inventory Management. McGraw-Hill: Boston, MA, 2000.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Unit slots based short-term scheduling for multipurpose batch plants

Naresh Susarla, Jie Li, I. A. Karimi

Department of Chemical & Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117576

Abstract

In chemical process industries, scheduling of multipurpose batch plants is challenging and has received tremendous attention so far. Recently, Susarla et al. (2009) explored the concept of unit- (asynchronized) slots for scheduling multipurpose batch processes and successfully reduced number of slots required for the optimal solution of a given problem. In this paper, we extend the work of Susarla et al. (2009) and utilize unit-slots to formulate a continuous-time mixed integer linear programming (MILP) model for the short-term scheduling of multipurpose batch processes. In addition to Susarla et al. (2009), our model is capable of handling various utility resources (other than the processing units and material resources) and sequence dependent changeover/setup times. Also, our model is capable of handling various storage configurations explicitly (Classes: UIS, LIS, and FIS with policies: UW, LW, and NW, Liu & Karimi, 2007). We demonstrate the performance of our model through an extensive numerical evaluation with some of the best known models from the literature. This rigorous comparison further elucidates that our model uses fewer binary variables, continuous variables and constraints.

Keywords: Scheduling, multipurpose batch plants, slot based formulations, unit slots, MILP.

1. Introduction

Short term scheduling is a regular task in any batch-wise manufacturing facility. In particular, multipurpose batch plants (MBPs) produce a number of products with the limited and shared resources. Inherent complexity of handling shared resources induces both opportunities and challenges for the manufacturer. Attempts to address these challenges have started way back in 1980's. Mendez et al. (2006) present excellent review highlighting the current approaches and associated challenges.

The scheduling problem primarily revolves around the allocation of the equipments and other resources and then aligning various tasks over time. Apart from the raw materials, resources in MBPs generally include other utilities such as: steam, electricity, cooling water, manpower (operators), etc. However, very limited work is reported in the literature that considers utilities other than equipment and raw materials in their mathematical modeling. All the previous attempts show that different types of time representation chosen have a lasting impact on the MILP models. Three popular approaches for modeling time in the literature are slot based, event based, and sequence

based (precedence based). Recently, Susarla et al. (2009) presented a concise representation of the prevalent continuous time modeling techniques.

For slot based representation, the scheduling horizon is modeled in terms of sequential time blocks or time slots of unknown lengths. In literature, slot based approach appears in two different styles of implementation that are process slots (Sundaramoorthy & Karimi, 2005) and unit slots (Susarla et al., 2009). Events based approach defines random points over time to which different tasks are associated. Two different ways of implementing this approach, global event points (Maravelias & Grossmann, 2003; Castro et al., 2004) and unit-specific event points (Janak et al., 2004) are popular in the literature. Similar to the process slots, global event points define time points that are common across all the units. In unit-specific event points, the time points are not common and are defined independently on each unit. Li et al. (2009) show with the help of several examples the fallibility of unit-specific event based models in accounting for shared resources. The sequence-based approach mainly uses either direct or indirect sequencing of task-pairs on various units.

Susarla et al. (2009) proposed a novel continuous-time MILP model for the short-term scheduling of MBPs employing unit slots. They introduced a set of task sequencing constraints to know the relative timings of the tasks in order to handle sharing of resources and thus eliminate the need of an extra binary variable. However, they consider the only resources as the equipments and the raw materials. Also, they assumed the changeover/setup times involved (if any) to be lumped up with the batch processing times. In this work, we extend their model considering other utilities apart from equipment and raw materials. We also explicitly model the sequence dependent changeover/setup times over each unit. Consequently, this enhances the generality of the model for its application over a broader variety of problems. As we show later, even the introduction of more constraints our model preserves its superiority over other reported models in the literature.

2. Problem Statement

A multipurpose batch facility produces a number of products by performing I tasks (i = 1, 2, …, I) sharing a set of J batch processing units (j = 1, 2, …, J) that involves S material states (s = 1, 2, …, S) each with a dedicated storage s. Production in the plant follows a set of processes that is described by the given recipe diagram. In most of the cases, resources used at each step for processing the material also involve other utilities such as HP/LP steam, cooling water, etc. A generalized recipe diagram clearly depicts the sequence of processes, specific utilization of resources (materials, equipments, and utilities) at each step, and their utilization coefficients, UC (σ_{sij}). Each material state s in a multipurpose batch plant has a specified storage capacity (UIS/LIS/NIS) and wait policy (UW/LW/NW). The scheduling problem in MBPs can be described as follows. Given the (1) generalized recipe diagram (GRD), (2) processing units (J), their suitable tasks (I), and processing capacity limits (3) storage capacities and wait policies, (4) required utilities and their availabilities, (5) sequence-dependent changeover/setup times, and (6) market price of each material state, we need to determine (1) the optimal sequence of the tasks, (2) their batch sizes, and (3) inventory profiles. For this, we assume (1) deterministic scenario i.e., no disruptions whatsoever, (2) batch size dependent processing times, and (3) negligible unit to unit transfer times. The objective for this scheduling is mainly to either maximize the revenue for a given scheduling horizon [0, H] or minimize the makespan of the products to meet a given product demand.

3. MILP Formulation

We consider time to be continuous on each unit j and storage S and model it in terms of K (k = 1, 2, …, K) contiguous slots of unknown and arbitrary lengths within the given horizon [0, H]. We represent the end time of slot k on processing unit j as T_{jk} [k = 0, 1, 2, …, K; $T_{j0} \geq 0$; $T_{jK} \leq H$; $T_{jk} \geq T_{j(k-1)}$, $1 \leq k \leq K$]. Similarly, T_{sk} denotes the end time of slot k on storage S. The time before the slot 1 starts is slot 0 (k = 0). So, a slot k on a unit (processing, j, and storage, S) starts at time $T_{j(k-1)}$, ends at time T_{jk}, and has a length [$T_{jk} - T_{j(k-1)}$]. Also, we assume that the real task begins at the start and may end at any time during the slot. Thus, the idle time (if any) is always towards the end of a slot. Moreover, a unit performs at most one task in each slot. However, it is quite possible that a unit performs no task during a slot or remains idle. To model this idling of units we define a zero (or an idle) task, i = 0. To schedule various tasks on different units, we now define one binary variable and two 0-1 continuous variables that denote respectively the start, end and continuations of a task on a unit slot as follows,

$$ys_{ijk} = \begin{cases} 1 & \text{if unit } j \text{ begins a new batch of task } i \text{ at } T_{jk} \\ 0 & \text{Otherwise} \end{cases} \qquad 1 \leq j \leq J,\ i \in \mathbf{I}_j,\ 0 \leq k < K$$

$$ye_{ijk} = \begin{cases} 1 & \text{if unit } j \text{ ends a batch of task } i \text{ within slot } k \\ 0 & \text{Otherwise} \end{cases} \qquad 1 \leq j \leq J,\ i \in \mathbf{I}_j,\ 1 \leq k \leq K$$

$$y_{ijk} = \begin{cases} 1 & \text{if unit } j \text{ continues a batch of task } i \text{ at } T_{jk} \\ 0 & \text{Otherwise} \end{cases} \qquad 1 \leq j \leq J,\ i \in \mathbf{I}_j,\ 0 \leq k < K$$

y_{ijK} is undefined, y_{ij0} is known and fixed. If a batch of task $i > 0$ is unfinished at time zero, and must continue, then y_{ij0} = 1, otherwise y_{ij0} = 0. This also sheds light on the fact that we do not force that each batch must start and end during a single slot, but it can span several slots.

Using the above, Susarla et al. (2009) developed several constraints for task & batch allocation, batch sizing, time sequencing, and material balancing. Some of these constraints are given below. For the batch to unit allocation, the following unit status balance equation is used

$$ye_{ijk} = [y_{ij(k-1)} + ys_{ij(k-1)}] - y_{ijk} \qquad 1 \leq j \leq J,\ 1 \leq k < K \qquad (1)$$

Similarly, for determining batch sizes for each batch the following balance on the amount of task i over slot k give

$$BO_{ijk} = b_{ij(k-1)} + [B_{ij}^{L}\, ys_{ij(k-1)} + \Delta BI_{ij(k-1)}] - b_{ijk} \qquad 1 \leq j \leq J,\ i \in \mathbf{I}_j,\ i > 0,\ 1 \leq k \leq K \qquad (2)$$

where, BO_{ijk} is amount of material that is released at the end of the ongoing task i during slot k, ΔBI_{ijk} is differential amount of material that is consumed at the beginning of slot k by task i, and b_{ijk}. Again, a balance on the remaining batch processing time (t_{jk}) yields the following equation

$$t_{j(k+1)} \geq t_{jk} + \sum_{i \in \mathbf{I}_j} [(\alpha_{ij} + \beta_{ij} B_{ij}^{L}) ys_{ijk} + \beta_{ij} \Delta BI_{ijk}] - (T_{j(k+1)} - T_{jk}) \qquad 1 \leq j \leq J,\ 0 \leq k < K \qquad (3)$$

In addition to the aforementioned constraints, the following sequencing constraints are required to ensure proper mass balance

$$T_{sk} \le T_{jk} + H[1 - \sum_{i \in \mathbf{I}_j, \sigma_{sij} < 0} ys_{ijk}] \quad 1 \le j \le J, \sum_{i \in \mathbf{I}_j, \sigma_{sij} < 0} \sigma_{sij} < 0, 0 \le k < K, s \ni I_s^U \text{ is LIS} \qquad (4a)$$

$$T_{sk} \ge T_{jk} - H[1 - \sum_{i \in \mathbf{I}_j, \sigma_{sij} > 0} ye_{ijk}] \quad 1 \le j \le J, \sum_{i \in \mathbf{I}_j, \sigma_{sij} > 0} \sigma_{sij} > 0, 1 \le k < K, s \ni I_s^U \text{ is LIS} \qquad (4b)$$

The above constraints, along with few more complete the model proposed by Susarla et al. (2009), but is incapable of handling utilities/resources other than raw materials and the processing equipments. However, chemical process industries in general and multipurpose batch plants in particular, require several other utilities/resources like- HP / LP steam for heating purposes, cooling water / ethylene glycol for cooling purposes, etc. While monitoring the consumption and a balance on these utilities are necessary, they are not addressed by Susarla et al. (2009). Therefore, in this paper, we extend that formulation by incorporating constraints for utility balance. For this, we first define $T_{u,k}$ as the end time of the slot on the utility usage, u. Also, we define $U_{u,k}$ as the rate of utility u in use during slot k. Now, as each utility is associated with certain processing task, the timings for the slots on these utilities have to be same as that of the processing units. So,

$$T_{uk} = T_{jk} \qquad 1 \le k \le K \qquad (5)$$

After associating the timings of utility usage, we now ensure the resource usage balance for each utility u. We model the utility usage based on its rate of supply. As we allow an unfinished batch to continue at time zero, the utility usage at the beginning of the horizon (U_{u0}^{in}) is known for every unit j. The rate of utility usage is determined by the size of a batch. The following equations make a balance on the rate of utility available at the end of any slot k.

$$U_{u1} = U_{u0}^{in} + \sum_{i \in \mathbf{I}_r, i>0} \sum_{j \in \mathbf{I}_j} (\gamma_{ij} + \delta_{ij} B_{ij}^L) ys_{ij1} + \delta_{ij} \Delta BI_{ij1} \qquad (6a)$$

$$U_{uk} = U_{u(k-1)} - \sum_{i \in \mathbf{I}_r, i>0} \sum_{j \in \mathbf{I}_j} (\gamma_{ij} ye_{ijk} + \delta_{ij} BO_{ijk}) + \sum_{i \in \mathbf{I}_r, i>0} \sum_{j \in \mathbf{I}_j} ((\gamma_{ij} + \delta_{ij} B_{ij}^L) ys_{ijk} + \delta_{ij} \Delta BI_{ijk})$$

$$1 < k < K \qquad (6b)$$

$$U_{uK} = U_{u(K-1)} - \sum_{i \in \mathbf{I}_r, i>0} \sum_{j \in \mathbf{I}_j} \gamma_{ij} ye_{ijK} + \delta_{ij} BO_{ijK} \qquad (6c)$$

The formulation so far assumes the changeover/setup times to be lumped up with batch processing times. However, with a small extension this formulation can handle sequence dependent changeover/setup times.

We define a maximum idling parameter, Φ, for each unit j that denotes the maximum number of consecutive slots for which a unit can remain idle. We start with $\Phi = 1$ and increase its value by 1 until there is no change in the objective function. Now, a sequence dependent changeover/setup time ($\tau_{i'ij}$) is to be incorporated between two consecutive real tasks. For this, we demand the start time of a real task to be greater than or equal to the sum of processing time of the immediately preceding real task and the associated changeover/setup time.

$$T_{jk} \ge T_{jk'} + (\alpha_{i'j} + \beta_{i'j} B_{i'j}^L) ys_{i'jk'} + \beta_{i'j} \Delta BI_{i'jk'} + \tau_{i'ij} ys_{ijk} - H(1 - ys_{i'jk'}) - H \sum_{\substack{i'' \in \mathbf{I}_j, i''>0 \\ i'' \ne i, i'' \ne i'}} \sum_{k' < k'' < k} ys_{i''jk''}$$

$$i, i' > 0,\ i \neq i',\ i, i' \in \mathbf{I}_j,\ 1 < k < K,\ k > k',\ (k - k') \leq \Phi,\ \tau_{i'ij} > 0 \qquad (7)$$

Note that eq. 7 imposes the required changeover/setup time ($\tau_{i'ij}$) between two consecutive tasks and is relaxed otherwise.
Now, considering the suitable wait policy (UW/LW/NW) for each material state s, we implement the following constraints.

$$t_{j(k+1)} \leq t_{jk} + \sum_{i \in \mathbf{I}_j} [(\alpha_{ij} + \beta_{ij} B_{ij}^L) ys_{ijk} + \beta_{ij} \Delta BI_{ijk}] - (T_{j(k+1)} - T_{jk}) + H[1 - \sum_{i \in \mathbf{I}_j, \sigma_{sij} > 0} yr_{ij(k+1)}]$$

$$1 \leq j \leq J,\ 0 \leq k < K \qquad (8a)$$

$$T_{sk} \leq T_{j(k-1)} + t_{j(k-1)} + \sum_{i \in \mathbf{I}_j} [(\alpha_{ij} + \beta_{ij} B_{ij}^L) ys_{ij(k-1)} + \beta_{ij} \Delta BI_{ij(k-1)}] + \sum_{i \in \mathbf{I}_j, \sigma_{sij} > 0} [w_{ij} ye_{ijk}] + H[1 - \sum_{i \in \mathbf{I}_j, \sigma_{sij} > 0} ye_{ijk}]$$

$$1 \leq j \leq J,\ \sum_{i \in \mathbf{I}_j, \sigma_{sij} > 0} \sigma_{sij} > 0\ ,\ 1 \leq k < K,\ s \ni I_s^U \text{ is LIS} \qquad (8b)$$

With the aforementioned equations in association with the model from Susarla et al. (2009) completes our model for the problems with utility constraints and sequence dependent changeover times.

4. Model Evaluation and Results

We present two case studies from Maravelias & Grossmann (2003), to evaluate the performance of our model. Fair and unbiased evaluations of different models require (Karimi et al., 2004; Susarla et al., 2008) a careful attention to the factors such as hardware, software, and the operating system. For our evaluation, we used CPLEX 11/GAMS 22.8 on a Dell precision PWS690 workstation with Intel® Xeon® 3 GHz CPU, 16 GB RAM, running windows XP Professional x64 Edition. Also, we solve various scenarios of each problem, demonstrating the robustness of our model.

4.1. Case study-1

Maravelias & Grossmann (2003) first used this example to illustrate the handling of utility constraints. The relevant data for this example can be obtained from the original paper. It involves 3 units, 4 tasks, 7 material states, and 2 utilities (cooling water and high pressure steam). We solve this example for different scenarios based on the rate of availability of utilities. In scenario A, we assume that both the utilities are available at a rate of 40Kg/min and scenario B we assume it to be 30 Kg/min. Also, we solve this example for two scheduling objectives: maximizing revenue and minimizing makespan. Table 1 consolidates the model statistics for both (ours and M&G) models. Clearly, our model takes fewer binary variables, continuous variables, constraints, and non-zeros in all the scenarios. This explains the simplicity of our model.

4.2. Case study-2

This example is from Maravelias & Grossmann (2003) (Example 3). While, the relevant data can be found in the original paper, it involves 6 units, 10 tasks, 14 material states, and 3 utilities. This example demonstrates the capability of these models in handling zero wait policy for some of the intermediate material states and various storage capacities for others. Again, we consider two different scheduling horizons to solve two different problems. Our model performs better and again requires fewer binaries, continuous variables, constraints, and non- zeros.

Table 1 Model and solution statistics for Case studies 1 & 2

Model	K	CPU (s)	Nodes	Objective	Binary Var	Continuos Var	Constraints	Nonzeros
Case Study 1: Scenario A - Revenue Maximisation								
Ours	7	1.72	1416	$5,904	54	365	452	1676
M & G	7	1.70	975	$5,904	84	484	1052	4117
Case Study 1: Scenario B - Revenue Maximisation								
Ours	6	0.26	126	$5,227.78	45	310	378	1396
M & G	6	0.15	67	$5,227.78	72	415	901	3356
Case Study 1: Scenario A - Makespan Minimisation								
Ours	8	6.67	6053	8.5 h	63	452	570	2057
M & G	8	4.30	1463	8.5 h	96	610	1212	5073
Case Study 1: Scenario B - Makespan Minimisation								
Ours	7	0.53	405	9.025 h	54	393	491	1765
M & G	7	0.69	292	9.025 h	84	534	1061	4241
Case Study 2: Scenario A (H = 12h) - Revenue Maximisation								
Ours	9	20.30	8037	$13,000	128	883	1366	4667
M & G	9	14.80	5315	$13,000	180	1009	2252	8033
Case Study 2: Scenario B (H = 14h) - Revenue Maximisation								
Ours	10	143.00	41933	$16,350	144	985	1535	5251
M & G	10	262.03	31906	$16,350	200	1121	2503	9240

M & G = Maravelias & Grossmann (2003)

5. Conclusions

This paper successfully utilizes the concept of unit slots, originally proposed by Susarla et al., (2009) and extends their formulation to consider utility handling constraints, sequence dependent changeover/setup times. While additional constraints are added for the said extension, the model preserves its superiority among the equals. Another critical point in this paper is that it does not add any extra binary variable to know the relative positions of the tasks as commonly done in slot based formulations.

References

Susarla, N., Li, J., Karimi, I. A., A Novel Approach for Scheduling Multipurpose Batch Plants using Unit Slots, *Submitted to AIChE Journal*, 2009.

Liu, Y., Karimi, I. A., Scheduling multistage, multiproduct batch plants with nonidentical parallel units and unlimited intermediate storage, *Chemical Engineering Science*, 2007, 62, 1549-1566.

Mendez, C.A., J. Cerdá , I. E. Grossmann, I. Harjunkoski, and M. Fahl, State-Of-The-Art Review of Optimization Methods for Short-Term Scheduling of Batch Processes, *Computers & Chemical Engineering*, 2006, 30, 913-946.

Sundaramoorthy, A., Karimi, I. A., A simpler better slot-based continuous-time formulation for short-term scheduling in multipurpose batch plants, *Chemical Engineering Science*, 2005, 60, 2679-2702.

Maravelias, C. T., Grossmann, I. E., New general continuous-time state-task network formulation for short-term scheduling of multipurpose batch plants, *Industrial & Engineering Chemistry Research*, 2003, 42, 3056-3074.

Janak, S. L., Lin, X., Floudas, C. A., Enhanced continuous-time unit-specific event-based formulation for short-term scheduling of multipurpose batch processes: Resource constraints and mixed storage policies, *Industrial & Engineering Chemistry Research*, 2004, 43, 2516-2533.

Li, J., Susarla, N., Karimi, I. A., The Fallibility of Some Unit-Specific Event-based Models for The Short-term Scheduling of Noncontinuous processes, *Submitted to Industrial & Engineering Chemistry Research*, 2009.

Castro, P. M.; Barbosa-Povoa, A. P.; Matos, H. A.; Novais, A. Q., Simple Continuous-Time Formulation for Short-Term Scheduling of Batch and Continuous Processes. *Industrial & Engineering Chemistry Research* 2004, *43* (1), 105-118.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Linking Marketing and Supply Chain Models for Improved Business Strategic Decision Support

José Miguel Laínez,[a] Gintaras V. Reklaitis,[b] Luis Puigjaner[a]

[a]*Department of Chemical Engineering, Universitat Politècnica de Catalunya, Barcelona, Spain, Luis.Puigjaner@upc.edu*
[b]*School of Chemical Engineering, Purdue University, West Lafayette, USA*

Abstract

Nowadays, a Supply Chain (SC) management model incorporating business strategic components is becoming of paramount importance to gain a competitive edge in the market place. To be successful, the enterprise model has to contemplate not only the supply chain, but also the demand chain. Understanding the market and customer behavior is extremely crucial for developing a good business policy. Marketing is a boundary-spanning activity, linking selling entities with buyers and intermediate channels. To operate most effectively, marketing activities must be coordinated with other functional areas of the firm. The marketing - SC management interface is an issue that deserves further research. Business managers should evaluate the existing trade-off between marketing and SC planning decisions in order to enhance the performance of the overall business metric: the shareholders value. Recently, there is a significant trend driving business managers to implement marketing science models for reaching more rational and holistic decisions. In this work, a mathematical model is presented for the enterprise that accounts for the three main business functionalities (i.e. operations, finances and marketing). The posed problem is tackled by developing a holistic MINLP model which optimizes in tandem the SC and marketing strategic decisions. Moreover, a financial model that allows to quantitatively assessing the enterprise value is also incorporated. Finally, main advantages of coordinated decisions are discussed through an illustrative example.
Keywords: enterprise wide optimization, marketing models, business decision support systems

1. Introduction

The business strategy is most often modeled as a hierarchical process in which functional strategies, such as operations, logistics, marketing, and finance are driven by a higher level strategy (Fine and Hax, 1985). A key element of the strategic framework involves coordinating functional-level plans to work in concert so as to achieve the overall business strategy rather than to locally optimize outcomes for individual functions, business units, plants, or stores. One of the primary challenges in implementing an effective strategy involves achieving consensus within the business organization. Unfortunately, while this concept is clearly sound on a conceptual level, actual implementation is typically very difficult (Bozarth and Berry, 1997).

One important strategic issue that needs more research efforts is the integration of supply chain (SC) production-distribution operations and marketing activities. Although several authors have highlighted the conflicting goals of supply chain (SC) and marketing managers (Eliashberg and Lilien, 1993; Shapiro, 2007), it is still typically

assumed that under a decentralized decision making scheme, marketing decisions are made first; determining demand forecasts which are later considered by the SC model to support production-distribution related decisions. By deploying this sequential procedure, the firm may be significantly under-estimating its overall performance. Integrating marketing and operations is a challenge in any business, since there is a natural tension between these two functional areas (Bozarth and Berry, 1997). At best, the tension between these two functions results in a dampening of marketing's tendency to over-promise to lure customers and a push on operations to move beyond an internal focus on reducing costs without a clear vision of end-consumer needs.

Usually, the primary objective of marketing function is maximizing revenues creation by satisfying customers through the products and services offered. On the other hand, SCM's focus is on the synchronization of production and distribution activities along the different entities comprising the SC network. The main objective is typically to minimize the total SC cost. In general, conflicts arise between marketing and SC because of these contrasting performance indicators which eventually are used to develop incentive structures for managers and their corresponding employees. For instance, one classical conflict between these two functions is the one associated to the inventory management. SC managers strive to keep low stock levels, while marketing managers long for high stock levels to guarantee that most of customer orders are met, thus improving revenue generation. Nevertheless, the enterprise main goal is to create and maximize shareholders value which actually is a function of revenues, cost and other economic factors. Consequently, business managers are in need of an integrated analytical decision support tool that is capable of appraising the trade-off between operations and marketing while evaluating and maximizing shareholders value.

Nowadays, there are more and more companies that are continuously searching for competitive advantages in order to get a better position in markets. One way of doing so may be by aligning functionalities strategic/tactical decisions towards the optimization of the overall business performance metric. In this work, we present a novel approach to address this challenge. We develop a MINLP model that tackles SC network design and marketing strategic decisions in tandem. Then, such model is coupled with a capital budgeting formulation which allows calculating shareholders value by means of the discounted-free-cash-flow (DFCF) method.

2. The marketing strategic model

Today firms have access to more market and customer data than they can use. Having too much data without the models and systems for analyzing what is important and what can be discarded can be as bad as or even worse than having too little data. Increasingly marketing managers are being asked to clear the same budget-justification hurdles imposed on other types of investment firms make. It is not surprising, therefore, that more managers are seeking help in turning their data and knowledge into improved decision making (Lilien and Rangaswamy, 2002). The systematic translation of data and knowledge into a normative and/or descriptive model that can be used for marketing decision support is what it is recently regarded as marketing engineering.

Here, BRANDAID, a marketing engineering contribution, is used. This flexible strategic marketing mix model is basically comprised of three sub-models, namely: advertising sub-model, pricing sub-model and sales force sub-model. It is noteworthy

that the model is (i) modular, so managers determine which sub-models to take into account according to their needs and (ii) customizable, since managers may introduce into the sub-models those functions that best describe their own marketing processes. The model is briefly presented next following the work of Little (1975). Two sub-models are considered in this paper; however other marketing influences/activities can be handled in the same manner.

2.1. The main model structure.

The model is based on the two following expressions. Eq. (1) expresses firm's target market demand (dem_{it}) as a function of market share (ms_{it}) and total market demand (D_t), while Eq. (2) computes market share as a reference value (S_o) modified by the effect of marketing activities (pricing, advertising, promotions, sales force) and other considered influences (e_{it}) such as competitors actions.

$$dem_{it} = ms_{it} D_{it} \qquad \forall i,t \tag{1}$$

$$ms_{it} = S_o \prod_{ma} e_{i,ma,t} \qquad \forall i,t \tag{2}$$

2.2. Advertising sub-model

This module assumes that there exists an advertising rate that maintains sales at target level. When the advertising rate is above reference, sales are assumed to increase; otherwise they decrease. The next equations describe this procedure. We assumed *ma=1* to describe advertising activities.

$$e_{i1t} = \varpi_1 e_{i1t-1} + (1-\varpi_1) A_{it} \qquad \forall i,t \tag{3}$$

$$A_{it} = f(h_t, k_t, AdvExp_t) \qquad \forall i,t \tag{4}$$

$$A_{it} = \frac{h_{it} k_{it} AdvExp_{it}}{h_{io} k_{io} AdvExp_{io}} \qquad \forall i,t \tag{5}$$

Here, ϖ_1 represents the carryover effect of advertising per period. The advertising rate (A_{it}) is the rate of messages delivered to customers by exposure in media. This rate is usually considered as a function of media efficiency (h_{it}), copy-effectiveness (k_{it}) and advertising expenditures ($AdvExp_{it}$). As previously mentioned, this function can be estimated by each particular case by examining the data available regarding this marketing activity. Hence, particular and current market-media trends can be taken into account by constructing an adequate descriptive model. For our illustrative example a function similar to Eq. (5) is used. In that equation, h_{io}, k_{io}, and $AdvExp_{io}$ are reference values of the above-mentioned parameters.

2.3. Pricing sub-model

Price elasticity of demand is considered by this sub-model. Again, more complicated descriptive models can be used to describe the specific product behavior of price-demand. We assumed *ma=2* to describe pricing activities. Here the following linear function is used. ε_t is the elasticity coefficient.

$$price_o(1-e_{i2t}) = \varepsilon_t(price_{it} - price_o) \quad \forall i,t \tag{6}$$

3. The SC design – planning model.

The design-planning approach is based on the work developed by Laínez et al. (2007). In this formulation, a four echelon SC is considered as shown in Figure 1. In this work, the equation that expresses that part of the demand can be left unsatisfied because of limited production capacity becomes one of the integrating equations as stated in section 5.

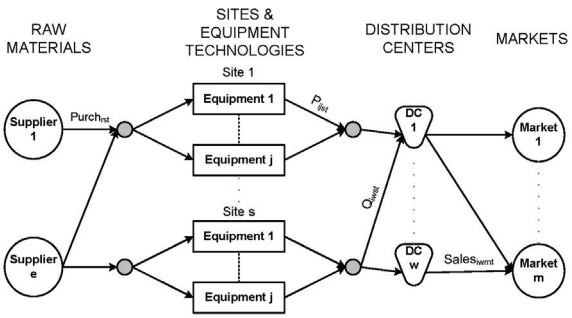


Figure 1: Supply chain model structure

4. The financial formulation

The financial component of the problem is tackled through the inclusion of a set of constraints that characterize economic issues, such as, payments to providers, loans, pledging decisions, etc. Also liquidity is controlled by this model. Furthermore, the Corporate Value (CV), is the objective function used. The CV of a company is a function of four factors: (i) investment, (ii) cash flows, (iii) economic life, and (iv) capital cost. Specifically, our work applies the DFCF method to compute the CV. Such method rates an entire company by determining the present value of its future free cash flows and discounting them, taking into account the appropriate capital cost during the evaluating time horizon. Capital cost is calculated using the weighted average method (WACC) which considers the firm's overall equity and debt.

Free cash flow at every period t (FCF_t) is defined by a function that depends on net operating profit after taxes, change in net working capital (ΔNWC_t), net change in investments ($NetInvest_t$) and advertising expenditures ($AdvExp_t$). Notice that Eq. (7) is already integrating marketing with finances through advertising expenditures. Eq. (8) is to compute the CV. For financial formulation details the reader is referred to Laínez et al. (2007)

$$FCF_t = Profit_t(1-trate) - (\Delta NWC_t + NetInvest_t + AdvExp_t) \quad \forall t \tag{7}$$

$$CV = \sum_{t=0}^{T} \frac{FCF_t}{(1+WACC_t)^t} - NetDebt_0 \tag{8}$$

5. Integration among models

Additionally to Eq. (7), integration among models is carried out by demand and revenue ($ESales_t$). As it can be observed a better balance between the demand generated by means of advertising and the available capacity proposed by the SC model can be achieved by utilizing Eq. (9).

$$\sum_w Sales_{iwt} \leq dem_{it} \qquad \forall i,t \tag{9}$$

$$ESales_t = \sum_i \sum_w Sales_{iwt}\, price_{it} \qquad \forall t \tag{10}$$

Notice that price and demand are problem variables instead of parameters and products *i* should be transported to final market from distribution centers *w*.

6. Illustrative example

The advantages of the proposed approach are demonstrated by solving an illustrative SC design-planning problem; which contains three processing sites (*S1*-*S3*), three distribution centers (*W1*-*W3*) and three market locations (*M1*-*M3*). A set of potential equipment technologies are assumed to be available for the processing sites. Three product families (*P1*-*P3*) can be manufactured on three different equipments types (*TA to TC*). A time horizon of 5 annual periods is considered. For all markets and products, the elasticity coefficient and the reference advertising expenditures are assumed to be equal to 0.85 and $3x10^6$ c.u. respectively. It takes 4 CPU seconds to reach the optimal solution for the integrated problem on an Intel Core 2 Duo at 2.0 GHz computer using DICOPT.

For comparison purposes, the problem has been also solved using the traditional sequential approach (SA) described in section 1. First marketing decisions are taken and then the SC and financial decisions are determined. The marketing decisions are typically obtained by optimizing net revenues (subtracting advertising expenditures from total revenues), while the financial and SC decisions are computed by maximizing the CV. Numerical results show that the solution calculated by the integrated approach (IA) offers improved performance over SA. Certainly, the optimal expected CV from IA is 47% higher than the one obtained by utilizing SA. The IA obtains such performance by a 2% decrease in net revenues (Table 1). The two approaches also yield different SC design decisions: SA proposes to install all equipment technologies in sites *S1* and *S3,* and *TA* in *S2*, while the IA proposes to install merely *TA* and *TC* in location *S1* and all technologies in site *S3*. Figure 2 shows advertising expenditures and the product demand induced by marketing activities for both approaches.

Table 1. Performance comparison between approaches

Approach	Revenues	Advertising Expenditures	Net Revenues	Corporate value	Investment in capacity
Sequential	$522.32x10^6$	$86.11x10^6$	$436.21 x10^6$	$84.35x10^6$	$79.77 x10^6$
Integrated	$512.94x10^6$	$85.15x10^6$	$427.79 x10^6$	$123.93x10^6$	$43.42 x10^6$

*All figures in c.u.

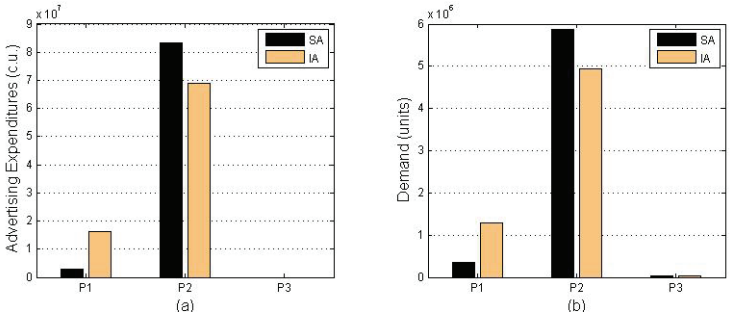


Figure 2. (a) Advertising expenditures and (b) Induced demand for both approaches

7. Conclusions

The intent of this paper is to motivate and draw attention to the need of further research in this kind of decision problems which are interfacing SC operations and marketing functions. A MINLP model integrating SC design/retrofitting, capital budgeting and marketing decisions is presented. Performance comparison with the traditional sequential decision approach is made, which demonstrates the significant economic benefits that the holistic approach provide. Our aim is to emphasize the relevance of a correct appraisal of the trade-off existing between the demand (which can be induced by marketing expenditures) and the SC capacity investments required to meet such demand. Finally, we would like to point out the great potential of response surface and data mining techniques in this field. Further work is focused on applying decomposition techniques so that (i) medium industrial size cases may be tackled using the proposed model and (ii) more complicated marketing models explored may be computed in acceptable time.

8. Acknowledgements

Financial support received from the "Generalitat de Catalunya" (FI grants; Comissionat per a Universitats i Recerca del Departament d'Innovació, Universitats i Empresa), and from the e-Enterprise Center, Purdue University is gratefully acknowledged. Besides, financial support from Xartap (I0898) and ToleranT (DPI2006-05673) projects is fully appreciated. A valuable literature survey was carried out by Ms. Carolina Quezada and co-workers at UCA for which the authors express their gratitude.

References

Bozart, C. C. and W. L. Berry, 1997. *Decision Sciences*, 28 (1) 121.

Eliashberg, J. and R. Steinberg, 1993. Marketing-Production Joint Decision-Making. In *Handbooks in Operation Research and Management Science*, Vol. 5, Marketing. Elsevier Science Publishers B.V.

Grant, J. L., 2003. *Foundations of Economic Value Added*; John Wiley and Sons

Laínez J.M., G. Guillén-González, M. Badell, A. Espuña and L. Puigjaner, 2007. *Ind. Eng. Chem Res.,* 46 (23), 7739.

Lilien, G. and A. Rangaswamy, 2001. *Interfaces*, 31(3) S1

Little, J.D., 1975. *Operations Research*, 23: 628.

Shankar, V., 2008. Strategic Marketing Decision Models for the Pharmaceutical Industry. In: *Handbook of Marketing Decision Models*. Springer: New York.

Shapiro, J. F., 2007, *Modeling the Supply Chain*. Duxbury.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Operation of the Argentinean Interconnected Electricity Network

Pablo Martínez, Ana María Eliceche

Departamento de Ingeniería Química, Universidad Nacional del Sur, PLAPIQUI-CONICET, Camino La Carrindanga km 7, 8000 Bahía Blanca, Argentina.

Abstract

A methodology to select optimally the electricity generation plants of the Argentinian interconnected network is developed. The electricity generation grid has different fossil fuel, hydroelectric and nuclear power stations. The power plants configuration connected to the grid and their operating loads are selected minimizing life cycle greenhouse gas emissions and operating cost simultaneously. A life cycle approach to estimate greenhouse gas emissions of thermoelectric, hydroelectric and nuclear power plants is followed. Binary operating variables represent discrete decisions to select which power plant is connected to the grid and the type of fossil fuel used. Continuous operating variables are introduced to select the optimal load for each power plant. A mixed integer linear programming problem is formulated and solved in GAMS. Significant reductions in green house emissions and operating cost are achieved simultaneously in the operation of the electricity network. Thus, a useful tool to support a decision-making process in the operation of a key energy sector is presented.

Keywords: Interconnected electricity network, operation, greenhouse emissions, cost.

1. Introduction

Environmental concerns have reached a high societal interest in a few years. Special interest has received greenhouse emissions of energy generation, claiming for environmental responsible energy policies. The main source of greenhouse emissions is the combustion of fossil fuels although greenhouse emissions are also present in the entire life cycle of many products or services, and electricity is not an exception. The upstream processes include raw material extraction, processing and distribution consuming energy, which have associated greenhouse emissions due to both fossil fuel consumption and fugitive emissions. Weisser (2007) presented an exhaustive work on life cycle greenhouse emissions in energy generation paying special attention to fossil fuel, nuclear and renewable energy technologies in the European Union and Japan. Hashim et al. (2005) studied the Ontario energy system minimizing CO_2 emissions. However, the authors have not considered the life cycle CO_2 emissions occurred in the upstream processes of each electricity generation option. The international price of CO_2 emissions is included in the objective function of the mixed integer linear programming problem formulated.

In the present work, life cycle green house emissions are estimated in the electricity generation and the limits of each power plant are extended, from raw material extraction to waste disposal. The Argentinean electricity grid has coal, fuel oil, gas oil and natural

gas driven thermoelectric plants, nuclear and hydroelectric plants. The Objective function minimized is a combined function that includes the operating cost and the benefits in the emission trade market for reducing the greenhouse life cycle emissions. The greenhouse gases international price was used to translate an environmental burden into an economic value. The network operating cost includes the costs of fuels and maintenance of each plant. Eliceche and Martínez (2007) successfully used a similar approach to select the operating conditions of a steam and power plant minimizing life cycle greenhouse emission of imported electricity and natural gas feedstock.
The methodology presented for the selection of the operating conditions of the Argentinian electricity network leads to significant reductions in operating cost and green house emissions simultaneously.

2. Electricity Network Modeling

2.1. Electricity Generation
The modeling of the interconnected system includes continuous and binary variables. The generation of each power plant is modeled as a fraction of its installed capacity. Binary variables represent discrete decisions and are introduced to select the type of fossil fuel used in a given thermoelectric plant, and to select which power plant (fossil-fuel based or not) is connected to the grid for a fixed yearly demand. The mix of all the electricity generated is injected into the electricity grid. The model considers the electricity generated by a certain power plant as a fraction of its maximum installed capacity, $\mathbf{G^{Max}}$:

$$\mathbf{G(q,f) = G^{Max}(q,f) \times d(q,f)} \qquad \forall \mathbf{q} \in \mathbf{F} \qquad (1)$$

$$\mathbf{G(q) = G^{Max}(q) \times d(q)} \qquad \forall \mathbf{q} \in \mathbf{NF} \qquad (2)$$

Where **G(q,f)** is the electricity generated, in *Gwh/yr*, by power plant **q** burning fossil fuel **f**. All the fossil fuel driven power plants are included in the group **F**. **G(q)** is the electricity generated by the power plant **q** which does not consume fossil fuels and belong to the group **NF** where nuclear and hydroelectric power plants are included. The variables **d(q,f)** or **d(q)** are the availability factor of each power plant. It express the ratio between the energy produced by a power plant in a certain period of time and the energy that it would be generated by the power plant working at its maximum installed capacity during the same period of time.
In order to select the fossil fuel used in a certain fossil fuel power plant, it is necessary to include binary variables in the mathematical model. The binary variable $\mathbf{y_{q,f}}$ takes the value 1 if the power plant **q** is burning the fossil fuel ***f*** and it is equal to 0 otherwise. The fact that a certain thermoelectric power plant could only work with an alternative fossil fuel, in a given time period, is modeled with the following equation:

$$\sum_{\mathbf{f}} \mathbf{y_{q,f}} \leq 1 \qquad (3)$$

Binary variables y_q are defined for the group of non-fossil fuel plants NF, to select which hydroelectric or nuclear power plants are on or off during the operation. The electricity generated for any power plant could not be greater than its installed capacity and it cannot be lesser than a certain value imposed by the interconnected system:

$$\mathbf{G(q,f) \le G^{Max}(q,f) \times \sum_{f} y_{q,f}} \qquad \forall \mathbf{q} \in \mathbf{F} \qquad (4)$$

$$\mathbf{G(q) \le G^{Max}(q) \times y_{q}} \qquad \forall \mathbf{q} \in \mathbf{NF} \qquad (5)$$

$$\mathbf{d(q,f) \ge d^{LB}_{q,f} \times y_{q,f}} \qquad \forall \mathbf{q} \in \mathbf{F} \qquad (6)$$

$$\mathbf{d(q) \ge d^{LB}_{q} \times y_{q}} \qquad \forall \mathbf{q} \in \mathbf{NF} \qquad (7)$$

Equations 4 and 5 represent upper bounds on energy production from each plant **q**. The Eq. 4 ensures that electricity generation from power plant **q** is zero when no fossil fuel is assigned to the power plant and the plant is shut down. The Eq. 5 indicates that electricity production in non fossil fuel (**NF**) plant **q** is smaller or equal to its maximum capacity. Equations 6 and 7 set up the lower limits in the availability factors of each group of power plants. These lower limits establish the minimum quantity of electricity generated by a certain power plant **q**.
An upper limit on the availability factor is set up in the Equations 8 and 9.

$$\mathbf{d(q,f) \le (1+\beta_{q}) \times d^{a}(q,f)} \qquad \forall \mathbf{q} \in \mathbf{F} \qquad (8)$$

$$\mathbf{d(q) \le (1+\beta_{q}) \times d^{a}(q)} \qquad \forall \mathbf{q} \in \mathbf{NF} \qquad (9)$$

The superscript "**a**" indicates the current value of the availability factor for each power plant. The parameter $\boldsymbol{\beta_q}$ is the maximum increment allowed for the availability factor for each power plant in the time period considered.
A demand satisfaction constraint is shown in Eq. 10, where **D** is the entire network electricity demand for the time period considered:

$$\mathbf{\sum_{q \in F} \sum_{f} G(q,f) + \sum_{q \in NF} G(q) \ge D} \qquad (10)$$

The operating cost equation for the entire network follows, where $\mathbf{C_F(q,f)}$ and $\mathbf{C_{NF}(q)}$ are the operating cost of each power plant in *US$/Gwh*, including fuels and maintenance costs.

$$\mathbf{C_G = \sum_{q \in F} \sum_{f} G(q,f) \times C_F(q,f) + \sum_{q \in NF} G(q) \times C_{NF}(q)} \qquad (11)$$

2.2. Greenhouse Emissions Quantification
A life cycle approach estimates greenhouse gas emissions of each thermoelectric, hydroelectric and nuclear power plant. The life cycle approach considers emissions during the life cycle of each electricity generation plant, from raw material extraction to waste disposal including the generation step itself. The quantification of greenhouse gases emissions (**GHG**) is estimated using emission factors, which express the mass of a certain greenhouse gas **k** emitted by unit of electricity generated. Greenhouse gases include CO_2, N_2O, CH_4, SF_6 and CFCs, each of them having different heat-trapping properties. To compare their effects on the atmosphere the Global Warming Potential, the **gwp** factors are used. Global Worming Potential express the ability of a greenhouse

gas to trap heat in the atmosphere relative to an equal amount of carbon dioxide, thus **gwp** factor is expressed in *mass of CO_2 equivalent/mass of GHG* k. Hence, to obtain the amount of greenhouse emissions CO_{2e} (mass of carbon dioxide equivalent), the mass of greenhouse gas **k** (CO_2, CH_4, N_2O) is multiplied by its corresponding $\mathbf{gwp_k}$ factor (1, 21, 310, respectively), Guinée et al (2002). The emissions of SF_6 and CFCs are negligible in fossil fuel combustion and during electricity life cycle (Dones et al, 2004), thus they were not considered in the present work. Therefore, the life cycle greenhouse emissions for the entire network are calculated as follows:

$$\mathbf{F_{GHG}^{LC}} = \sum_{\mathbf{q}\in\mathbf{F}}\sum_{\mathbf{f}}\sum_{\mathbf{l}} \mathbf{G(q,f)} \times \mathbf{E_{GHG}^{l}(q,f)} + \sum_{\mathbf{q}\in\mathbf{NF}}\sum_{\mathbf{l}} \mathbf{G(q)} \times \mathbf{E_{GHG}^{l}(q)} \qquad \mathbf{l} = 1,...,\mathbf{l_q} \qquad (12)$$

$$\mathbf{E_{GHG}^{l}(q,f)} = \sum_{\mathbf{k}} \mathbf{E_k^l(q,f)} \times \mathbf{gwp_k} \qquad (13)$$

$$\mathbf{E_{GHG}^{l}(q)} = \sum_{\mathbf{k}} \mathbf{E_k^l(q)} \times \mathbf{gwp_k} \qquad (14)$$

Where $\mathbf{E_k^l}$ is the emission factor of greenhouse gas **k** in the life cycle stage **l** in *ton of CO_{2e}/Gwh*. The subscript **GHG** indicates the summation over the three greenhouse gases considered in each life cycle stage **l**. The total number of life cycle stages, $\mathbf{l_q}$ considered are:

i. *Thermoelectric power generation*: exploration, extraction, refining, transport and generation step for four different fuels: natural gas, fuel oil, gas oil and coal.
ii. *Hydroelectric power generation*: material transport in construction and submerged biomass decay in the operation.
iii. *Nuclear power generation*: exploration, extraction, refining, fuel assembly and transport of uranium, waste treatment and disposal of spent fuel, transport in construction phase of the nuclear plant.

A detailed analysis of each life cycle stage considered as well as the literature sources was presented in Eliceche and Martinez (2007).

3. Optimization Problem Formulation

The objective is to determine the optimal configuration and load distribution for all power plants, to provide the electricity demand to the grid minimizing life cycle green house emissions and operating cost simultaneously. A combined objective function includes the operating cost and the benefit for reducing green house emissions. The greenhouse gases international market price $\mathbf{pr_{GHG}}$ (*US$/ton of CO_{2e}*) was considered to translate an environmental burden into an economic value.

$$\mathbf{EnvEco(x)} = \mathbf{C_G(x,y)} - \left[\mathbf{F_{GHG}^{LC}(x,y)}\Big|_{\mathbf{in}} - \mathbf{F_{GHG}^{LC}(x,y)}\Big|_{\mathbf{o}} \right] \times \mathbf{pr_{GHG}} \qquad (15)$$

The benefit (term in brackets) is proportional to the reduction in green house emissions from the initial point (sub index **in**) and the optimal solution (sub index **o**). The constant term corresponding to the initial point life cycle greenhouse emissions is removed from Eq. 15 leading to the objective function named "**EnvEco**" in the following Mixed Integer Linear optimization problem P1:

$$\min_{\mathbf{x},\mathbf{y}} \quad \mathbf{EnvEco(x,y)} = \mathbf{C_G(x,y)} + \mathbf{F_{GHG}^{LC}(x,y)} \times \mathbf{pr_{GHG}}$$

$$\begin{aligned} &\textbf{s.t.} \\ &\mathbf{A.x} = 0 \\ &\mathbf{B.x} + \mathbf{C.y} \le 0 \\ &\mathbf{x} \in \mathbf{R^n} \\ &\mathbf{y} \in \{0,1\} \end{aligned} \qquad P1$$

Where **x** and **y** are the continuous and discrete optimization variables respectively; equality constraints represent the generation model given in Equations 1 and 2 and the emission model given in Eq. 12 with life cycle emission factors given in Eq. 13 and Eq. 14. Inequality linear constraints are included to represent minimum and maximum plant capacity constraints, demand satisfaction constraint and bounds on continuous variables, Equations 3 to 11. The continuous optimization variables of the problem are the availability factors for all the power generation plant. Fuel resource constraints and mandatory fuel usage can be included as constraints to represent temporal resource availability due to natural gas pipeline capacity and national policies regarding for example the use of coal. It is also possible to include an inequality constraint with a minimum required reduction in total greenhouse emissions to comply with Kyoto protocol (IPCC, 2001) targets stated for a given country or industrial branch. This constraint was not included, because Argentina has no greenhouse emissions reduction targets up to now.

4. Numerical Results

The configuration and loads of the power plants of the Argentinean interconnected electricity system were selected optimally minimizing operating cost and green house emissions as formulated in problem P1. There are 51 thermoelectric power plants consuming 4 alternative different fossil fuels, 41 hydroelectric power stations and 2 nuclear power plants. The 51 thermoelectric power plants have 170 thermoelectric machines including gas and steam turbines and combined cycle units. Data of system operation and installed capacity, availability factors bounds, power plant efficiencies and generation costs were taken from the electricity national company, CAMMESA (2004). The electricity demand for year 2004 was formulated as an inequality constraint, being active at the solution point. The initial point of the optimization problem corresponds to the operating conditions of year 2004. The optimization problem P1 is a mixed integer linear programming problem (MILP) solved in GAMS (Brooke, 2003), using CPLEX as a MILP solver. The main results comparing the operation of the electricity network in year 2004 (the initial point) and the optimal operating conditions (solution point minimizing EnvEco) are shown in Table 1. An international price of U$S 20/ton CO_{2e} was considered (Point Carbon, 2004).

Significant reductions of 48 % and 44 % in green house emissions and operating cost are reported by properly selecting the configuration and loads of the electricity generator plants. More constraints reinforcing national regulations related to the use of certain fuels like coal in order to keep employment in mining areas, might reduce the improvement reported in this case study.

Thermoelectric generation is reduced, while hydroelectric and nuclear generation is increased. Thermoelectric units that were kept switch on, were those more efficient and less pollutant as it is the case of combined cycle thermoelectric units burning natural

gas. The shut down thermoelectric units were burning fuel oil, gas oil and coal, which are the most expensive and pollutant fossil fuels. Hydroelectric, nuclear and combined cycle burning natural gas are also the cheapest options.
The methodology presented can be extended to different regional applications and time periods.

Table 1. Improvements achieved selecting the power plants in operation.

		Year 2004	Min EnvEco	% Reduction
GHG Emissions	CO_{2e} 10^3 ton	42.00	21.75	48.21
Operating Cost	10^3 US$	2407.67	1353.62	43.77
EnvEco	10^3 US$	3247.69	1788.71	44.92
Thermoelectric	Gwh / plants	42766.00 / 170	38871.16 / 23	9.11
Hydroelectric	Gwh / plants	31827.00 / 41	35019.58 / 40	-10.03
Nuclear	Gwh / plants	7312.90 / 2	8015.16 / 2	- 9.60
Total Electricity Generation	Gwh	81905.90	81905.90	0.00

5. Conclusions

The methodology developed for the selection of the operating conditions of the Argentinian electricity network leads to significant reductions of more than 40 % in operating cost and green house emissions simultaneously. Hydroelectric, nuclear and the most efficient and less pollutant thermoelectric units, as the combined cycle burning natural gas, are in operation while the less efficient thermoelectric power plants burning coal, fuel oil and gas oil are shut down. Hydroelectric, nuclear and combined cycle burning natural gas plants are also the cheapest options. Thus, a useful tool to support a decision-making process in a key energy sector has been presented.

References

A. Brooke, D. Kendrick, A. Meeraus, R. Raman (eds.), 2003, GAMS, A user guide, GAMS Development Corporation, Washington DC.

A. Eliceche and P. Martinez, 2007, Minimization of life cycle CO2 emissions in the operation of a steam and power plant, ESCAPE17, V. Plesu and P.S. Agachi (Editors).

CAMMESA, 2004, Compañía Administradora del Mercado Eléctrico Mayorista Sociedad Anonima, Informe Anual 2004. Buenos Aires.

D. Weisser, 2007, A guide to life-cycle greenhouse gas (GHG) emissions from electric supply technologies, Energy, 32, 1553-1559.

H. Hashim, P. Douglas, A. Elkamel and E. Croiset, 2005, Optimization Model for Energy planing with CO2 Emission Consideration, Ind. Eng. Chem. Res., 44, 879-890.

IPCC. Climate Change 2001: the scientific basis. Contribution of working group I to the third assessment report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University. www.ipcc.ch

J. Guinée, R. Heijungs, G. Huppes , R. Kleijn, A. Koning, L. van Oers, A. Sleeswijk, S. Suh, H. Udo de Haes (eds.), 2002, Handbook on Life Cycle Assessment. Operational Guide to the ISO Standards. Kluwer Academic Publishers, Dordrecht.

Point Carbon, 2004. Carbon Market Consulting Services. www.pointcarbon.com

R. Dones, T. Heck and S. Hirschberg, 2004, Greenhouse gas emissions from energy systems, comparison and overview, Encyclopedia Energy, 3, 77-95.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Systematic Framework to Calculate Economic Value and Environmental Impact of Biorefining Technology

Norman Sammons Jr.[a], Wei Yuan[a], Susilpa Bommareddy[a], Mario R. Eden [a], Burak Aksoy[b], Harry Cullinan[b]

[a]*Auburn University, Department of Chemical Engineering, 222 Ross Hall, Auburn University, AL 36849-5127, USA, edenmar@auburn.edu*
[b]*Auburn University, Alabama Center for Paper and Bioresource Engineering, 242 Ross Hall, Auburn University, AL 36849-5127, USA, culliht@auburn.edu*

Abstract

The integrated biorefinery has the potential to provide a strong, self-dependent, and sustainable alternative to the use of fossil fuels for the production of chemicals and energy, but difficulties arise in measuring the potential economic benefit and environmental impact of current and emerging biorefinery technology. A plethora of products and production pathways are possible in this growing field of biorefining, and the production path with maximum value and minimum environmental impact cannot be determined by heuristics alone. A framework is needed to determine the most optimal route based on measures of economic and environmental performance. Gross profit is used as a short-term economic metric requiring few details, while net present value is used as a long-term metric that requires information about rate of return, regulatory and market possibilities, and hedging options. Environmental impact is evaluated using the US-EPA WAR algorithm developed by Young and Cabezas (1999). The best candidates in economic and environmental performance are then subject to process integration techniques in order to maximize resource usage, and these integrated biorefineries are once again analyzed for optimal performance.

Keywords: Environmental impact, Economic analysis, Biorefinery, Product allocation, Optimization

1. Introduction

The integrated biorefinery provides a unique opportunity for reinvigorating an entire manufacturing sector by utilizing a renewable resource to create valuable product streams in the forms of chemicals, fuels, pharmaceuticals, and energy (Bridgwater, 2003). Through the optimal use of renewable feedstocks such as forest residues, forest byproducts, and crop residues, long-term economic and environmental sustainability can be achieved. The vast amount of possibilities in feedstocks, chemical processes, and final product routes results in a need for a process systems engineering (PSE) approach to ensure maximum economic and societal benefit through minimizing the usage of raw material and energy resources as well as the cost involved in supply chain operations intrinsic to biorefining. A unique partnership has been established consisting of researchers, government entities, equipment vendors and industry stakeholders to maximize the applicability of such systematic methods and to integrate experimental

and modeling work. This partnership has played an integral role in the procurement of the wide range of information necessary such as data needed for process models, capacity constraints, financial data, and optimization techniques. The overall goal of this work is to develop a system that will enable decision makers to evaluate novel production pathways in biorefining in order to optimize profitability while determining and minimizing environmental impact.

2. Development, Integration, and Evaluation of Process Designs

In biorefining, the large number of possible process configurations and products results in a highly complex problem that cannot be solved using simple heuristics or rules of thumb. Thus, it is necessary to develop a framework which incorporates profitability measures as well as environmental metrics and other techno-economic metrics. This framework should enable policy and business decision makers to answer a number of important questions like:

- For a given set of product prices, what should the process configuration be, i.e. what products should be produced in what amounts?
- For a given product portfolio, how can process integration methods be utilized to optimize the production routes leading to the lowest environmental impact?
- What are the discrete product prices that result in switching between different production schemes, i.e. what market developments or legislative strategies are required to make a certain product or pathway attractive?
- What are the ramifications of changes in supply chain conditions on the optimal process configuration?

The introduction of process system engineering (PSE) methods into biorefining research provides a systematic framework capable of seamlessly interfacing results generated in both experimental work and simulation studies. Such a framework is imperative when attempting to combine the vast array of knowledge and information from many research areas and disciplines. Figure 1 illustrates the data generation and flow necessary for this framework. While the overall strategy of the framework will be reviewed here, complete details of the strategy are presented in earlier work (Sammons et al., 2007).

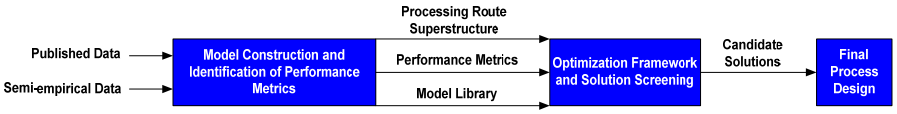


Figure 1 – Overall data generation and flow necessary for product allocation framework.

An initial superstructure lists feasible technologies for a given feedstock, and basic simulation models are then constructed for those corresponding processes based on published and semi-empirical data. If a given model is solvent-based, computer-aided molecular design (CAMD) and property clustering techniques are employed to identify alternative solvents that minimize safety and environmental concerns (Eden et al., 2004; Harper and Gani, 2000). Process integration techniques are then used to optimize the simulation models. This is an integral step in the model development as it ensures optimal utilization of biomass and energy resources, which will inevitably result in reduced environmental impact. The optimized models are used to generate data for the economic as well as environmental performance metrics. The estimation of

environmental performance is achieved through the use of the US-EPA Waste Reduction (WAR) algorithm (Young and Cabezas, 1999). The end result of this strategic portion is a superstructure of biorefining processing routes, a library of simulation models for these routes, and a database of corresponding economic and environmental performance metrics.

As stated in previous work, this superstructure and database is then exported into a mathematical optimization program whose objective is to identify candidate solutions based on maximum profitability (Sammons et al., 2007). The environmental impact of these candidate solutions is also measured for ranking purposes. If a candidate solution satisfies both economic objectives and environmental constraints, then the optimal production pathways have been identified. However, if none of the candidate solutions satisfy the environmental impact requirements, then the framework is modified by relaxing economic performance constraints until one or more solutions will satisfy both criteria. In doing this, economic and environmental performance is decoupled in order to avoid the pitfall of identifying the zero production facility that will identify the minimum environmental impact (Sammons et al., 2007). It should be noted, that only the economic and environmental performance metrics are incorporated in the solution framework described below, thus decoupling the complex models from the decision-making process. This approach allows for updating the model parameters as new data and/or technology becomes available without having to change the selection methodology, thus making it robust and flexible.

3. Generalized Biorefinery Model and Economic Metric Optimization

Figure 2 illustrates the generalized biorefinery model that has been used to develop the structure of the optimization framework. The model structure was formulated to include a variety of basic complexities encountered in the decision making process (Sammons et al., 2007).The objective function maximizing the overall gross profit of the biorefinery over a given time period *t* is illustrated in Equation 1, but this may be replaced with a more robust objective function maximizing net present value as shown in Equation 2.

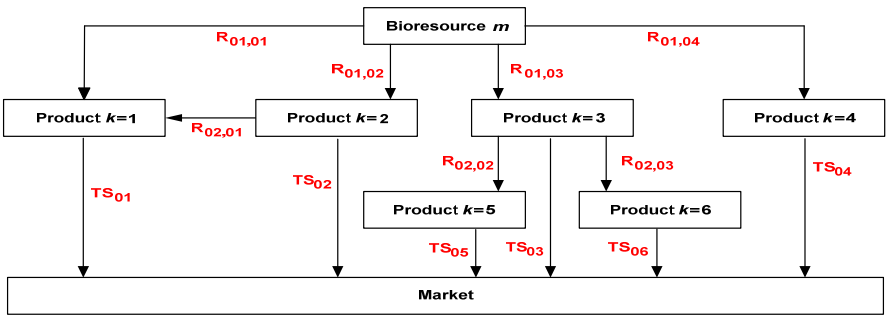


Figure 2 – Generalized model to illustrate possibilities in decision-making tree.

$$GP_t = \sum_m \left(\sum_k TS_{mkt} C_{kt}^s - \sum_i \sum_j R_{mijt} C_{mijt}^P - C_{mt}^{BM} \sum_j R_{m1jt} \right) \tag{1}$$

$$\max NPV = \sum_t \frac{GP_t (1 - Tax_t) + Dep_t Tax_t - Hedge_t + Gov_t}{(1+R)^t} \tag{2}$$

Using this nomenclature, the first set of terms in Eq. (1) represents the sales revenue from the products made from each bioresource *m*. TS_{mkt} is a variable that denotes the production rate of product *k* from bioresource *m* that is sold to the market during time period *t*. C_{kt}^{S} is the sales price of product *k*, which is a scalar and is determined through a market survey of published prices and vendor quotes. The second set of terms represents the total processing cost incurred by the pathways pursued in production. R_{mijt} is a variable that represents the processing rate of route *ij* while C_{mijt}^{P} is a scalar that represents the processing cost of that route and is determined through simulation models and process economics. The third set of terms represents the total cost of the biomass resource *m*, and this is broken down into the scalar purchase price of bioresource *m* in C_{mt}^{BM} and the combined rate of biomass processed by the plant in R_{m1jt}. Although both TS_{mkt} and R_{mijt} are variables in the optimization program, they are not independent since the variables are related via mass balance constraints around the product points and process conversion factors.

This gross profit term may be maximized by itself or incorporated into a net present value objective function that takes many more issues into account. Tax_t represents the marginal tax rate, $Hedge_t$ represents the expenses associated with hedging against catastrophic market actions, Gov_t is the net benefit realized through government incentives or penalties, and *R* is the expected rate of return, or cost of capital involved with the time value of money. The gross profit version of the model may be utilized for short-term decision making and scheduling with existing equipment and supply chain arrangements, while the net present value version is better suited for long-term planning that will involve extensive capital expenditures, anticipated government actions, and supply chain considerations.

4. Results and Case Studies

Without including any constraints on capacity of the process points, the solution is a single-product configuration in which all available biomass is converted into the most profitable product, or the product with the greatest contribution margin. However, if constraints are imposed on the most profitable route, the framework identifies the additional products and processing routes with the next highest contribution margins in order to maximize profitability, thus leading to a polygeneration facility (Sahinidis et al., 1989). In order to effectively address the strategic planning objectives of business decision makers, it is necessary to incorporate the total capital investment as a constraint in the formulation. Inclusion of capital cost constraints is crucial for practical application of the results, i.e. enabling consistent evaluation of the potential benefits to be obtained for a given maximum investment.

A case study was performed on a potential grassroots biorefinery involving the conversion of chicken litter to syngas, hydrogen, and electricity. Conversion rates were obtained from experimental work performed by the university and affiliated agencies as well as simulations constructed in ASPEN. Figure 3a shows the possible pathways for production and sale of these products on the common market. The execution of the framework verified the results obtained from manual calculation; producing syngas from chicken litter and selling it on the market via pipeline would maximize profit due to the high costs intrinsic with converting the syngas to hydrogen or electricity. Figure 3b highlights the active pathway chosen by the optimization program.

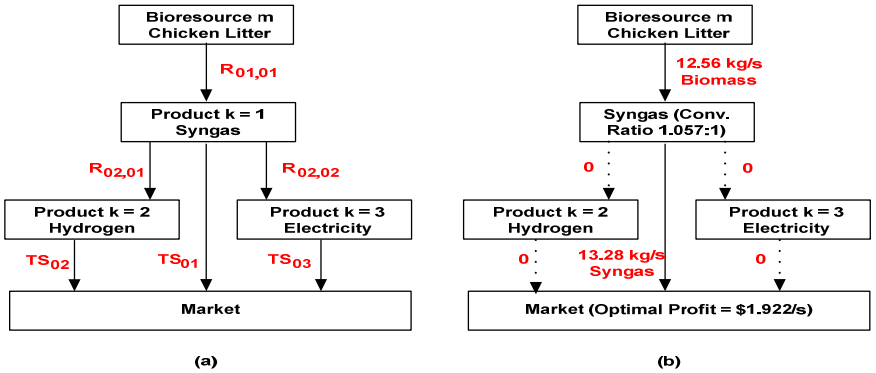


Figure 3 – (a) Unsolved decision making tree. (b) Solved decision making tree with flowrate values and objective function value.

A second case study was performed on the potential of gasification of black liquor byproduct from pulp and paper mills and its subsequent conversion into electricity, steam, and liquid fuels. These processes were compared with the use of a replacement state-of-the-art Tomlinson boiler, which represents the status quo of burning the organic components of black liquor to generate steam in a relatively inefficient manner. Data from an extensive study performed by Larson et al. (2006) was used to construct simulation models that provided the data necessary for the framework to be utilized. Figure 4 illustrates the unsolved decision making tree for the black liquor case study, which shows the process models simulated and observed in the Larson study. From execution of the optimization code, the framework suggested that all biomass should be gasified and processed through the FTc pathway, which represents one of the ways in which Fischer-Tropsch liquid fuels may be produced through the gasification of black liquor along with gasifying supplemental biomass. This result, shown in Figure 5, is qualitatively confirmed by the observation made by Larson et al. (2006) that this process is the most profitable since it has the highest net present value over time.

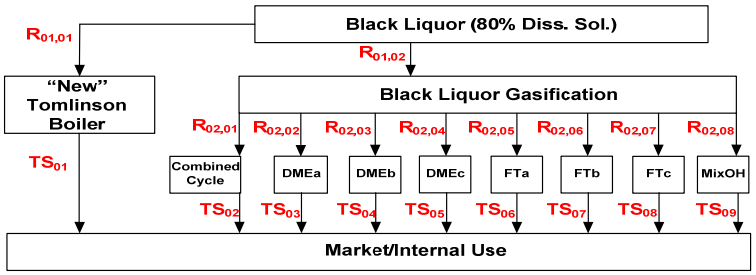


Figure 4 – Unsolved decision making tree for black liquor gasification study based on data from Larson et al (2006).

5. Conclusions and Future Work

A general systematic framework for optimizing product portfolio and process configuration in integrated biorefineries has been presented. Decoupling the process models from the decision-making framework reduces problem complexity and increases robustness. The next phase of this work involves development of additional process

models for the generation of performance metrics, specifically information on conversion, yield, and production cost for economic metrics.

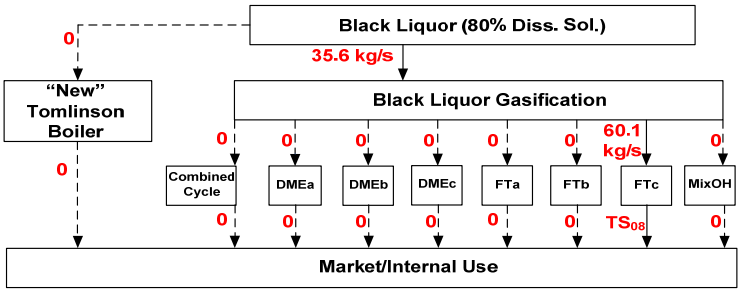


Figure 5 – Solved decision making tree for black liquor gasification study. TS_{08} contains multiple products of FT fuels, electricity, and process steam.

The EPA WAR algorithm will be incorporated into the mathematical optimization software, and data from the additional process models will be used to generate numerous time-based and mass-based measures of environmental impact. From there, process integration will be utilized to optimize these process models by reducing energy usage, material consumption, and waste streams. The framework will also become a stronger financial tool through the incorporation of various economic ideas and analyses. The further development of qualitative predictive models for capital investment and inclusion of capital amortization into the objective function will also increase the strength of the framework. Incorporation of options theory into the framework will allow management to develop financial strategies in response to events in the market or legislative environment. Optimization under uncertainty and superstructure generation techniques will be studied to increase the robustness of the framework (Banerjee and Ierapetritou, 2003; Chakraborty and Linninger, 2003).

References

A.V. Bridgwater, (2003). Renewable fuels and chemicals by thermal processing of biomass. Chemical Engineering Journal, 91, 87-102.

M. R. Eden, S. B. Jørgensen, R. Gani and M. M. El-Halwagi, (2004). A novel framework for simultaneous separation process and product design. Chemical Engineering and Processing, 43, 595-608.

P. M. Harper and R. Gani, (2000). A multi-step and multi-level approach for computer aided molecular design. Computers and Chemical Engineering, 24, 677-683.

D. M. Young and H. Cabezas, (1999). Designing sustainable processes with simulation: the waste reduction (WAR) algorithm. Computers and Chemical Engineering, 23, 1477-1491.

N. E. Sammons, W. Yuan, M. R. Eden, H. T. Cullinan, B. Aksoy, (2007). A flexible framework for optimal biorefinery product allocation. Journal of Environmental Progress, 26(4), 349-354.

N. V. Sahinidis, I. E. Grossmann, R. E. Fornari and M. Chathrathi, (1989). Optimization model for long range planning in the chemical industry. Computers and Chemical Engineering, 13, 1049-1063.

E. D. Larson, S. Consonni, R. E. Katofsky, K. Iisa, and W. F. Frederick, (2006). A Cost-Benefit Assessment of Gasification-Based Biorefining in the Kraft Pulp and Paper Industry, Volume 1, Princeton University.

I. Banerjee and M. G. Ierapetritou, (2003). Parametric process synthesis for general nonlinear models. Computers and Chemical Engineering, 27, 1499-1512.

A. Chakraborty and A.A. Linninger, (2003). Plant-wide waste management 2: Decision making under uncertainty. Industrial and Engineering Chemical Research, 42, 357-369.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Development of a computer support system for the management of regulatory compliance of pharmaceutical processes

M. Berkan Sesen[a], Pradeep Suresh[b], René Bañares-Alcántara[a,]*, Venkat Venkatasubramanian[b]

[a]*Department of Engineering Science. University of Oxford, Oxford OX1 3JP, UK*
[b] *School of Chemical Engineering. Purdue University, W. Lafayette, IN 47907, USA*

Abstract

The pharmaceutical sector is one of the most tightly regulated industries today and is constantly being challenged to meet rising standards of quality. However, it still uses paper documents, spreadsheets and conventional databases for the storage and manipulation of the manufacturing and regulatory process knowledge. Furthermore, in the current industrial approach, the interpretation of the regulations (which are written at a very abstract level) into operating procedures is done manually and as an afterthought to pharmaceutical process development. This approach is error-prone, time consuming and very effort intensive as it does not take advantage of recent advances in the field of knowledge management.

We have been working in the development of a computer-based support system to assist in the identification of regulatory compliance of a drug manufacturing process. OntoReg, the current prototype, encapsulates pharmaceutical process and regulation knowledge in two complementing representations:

1. OWL ontologies (a knowledge representation consisting of taxonomies of concepts and logical axioms allowing to structure those concepts and to detect inconsistencies in the resulting structure), and
2. SWRL rules (which act as constraints able to enforce values inside the concepts or create relations between them).

These two components, ontologies and rules, are integrated through a Java user interface which is able to identify when a pharmaceutical process does not comply with a regulation and to suggest remedial action.

The resulting OWL ontology is structured in terms of three types of concepts belonging to a Regulatory, a Process or an Abstract domain (such as Time and Parameter). OntoReg has been tested with a case study for the aspirin production process in the context of its compliance with some equipment maintenance and cleaning regulations taken from the European Union Guidance on Good Manufacturing Practice.

Our approach has the potential to substantially decrease pharmaceutical process validation time (up to several weeks) and effort and thus reduce development costs and commercialization prices (a well known estimate is $1 million per day profit for a blockbuster drug). Furthermore, this novel approach has the potential to be extended to other regulatory applications in the future, e.g. environmental compliance.

Keywords: Decision support systems, Pharmaceutical processes; Regulatory Compliance; Ontologies.

1. Introduction

During the preparation for pharmaceutical process validation, data must be systematically gathered, analyzed and presented to give reasonable documented evidence that a pharmaceutical process, when operating within specified parameters, will "consistently" produce a product that meets its pre-specified quality attributes. Closely related to both are Compliance Management Systems (CMSs), i.e. systems that keep track of the tasks needed to comply with the validation requirements. Conventionally, compliance management is a manual and highly labor-intensive task that creates additional overheads to many businesses [1]. Pharmaceutical manufacture, in particular, is one of the most tightly regulated industries today and it is constantly being challenged to meet rising standards of quality and to comply with rigorous regulatory requirements. Ensuring regulatory compliance and managing a myriad of validation documents constitutes a major engineering challenge because manual handling of voluminous and heterogeneous data is not only time-consuming but also very expensive and error-prone.

In the context of the pharmaceutical industry, while front end applications have been designed to automate process controls for drug manufacture, back end data collection and storage applications have significantly lagged behind, still relying on manual and paper based archiving methods. Motivated by this need, various compliance management systems have been offered in the past. In 2000, when electronic record keeping was still new to the Food and Drug Administration, Conley [2] looked at automating regulatory compliance systems, focusing mainly on the NuGenesis Scientific Data Management System (SDMS). In 2006, Boland et al. [3] compared the key features of the main commercial CMSs within the context of environmental management. Yip et al. [1] used semantic modeling to develop a compliance audit mechanism, namely XISSF, which enforces information security standards and regulations. These important initial attempts, while recognizing the need for an informatics framework to automate CMSs, are limited in their conceptualization of the raw regulatory data, parsing of regulations into tasks and in their application to the pharmaceutical domain. In terms of compliance systems, establishing the correlations and relationships among raw data in order to acquire useful business information is vital. However, a systematic way to convert the raw/implicit regulatory data to human-interpretable and machine-readable explicit information is still lacking.

The rest of the paper is organized as follows: Section 2 presents the methodology followed to develop the OntoReg platform, its architecture and structure. Results are presented in Section 3, and Section 4 concludes with the contributions and future work.

2. Methodology

Motivated by the lack of an existing tool, this work aims to prove that an ontological knowledge-base infrastructure can be applied to the pharmaceutical process regulation domain in order to facilitate the management of the validation requirements. In a nutshell, ontologies capture and describe the semantics of data sources and make their contents explicit and platform-independent. They are not only useful for unifying database, data-warehouse, and knowledge-base vocabularies but also for maintaining consistency through inferential reasoning tools. Compared to a database schema (which targets physical data independence) and to an XML schema (which targets document structure), an ontology targets shared and explicit semantics of information [4]. In our application, this explicit semantic knowledge is embedded on a domain specific ontology, named OntoReg, and processed by the combination of an ontological reasoner

(Pellet 1.5.2) and a rule engine (Jess 7.1) through the use of a Java integrated development environment (IDE). The ontology aims to capture pharmaceutical regulatory and process domain knowledge in a generic way and to provide a common representation of the domain, which may be reused, shared, and standardized across applications and groups [5]. The contribution of this work is the use of this domain specific ontology, i.e. OntoReg, to form an integrative platform, which amalgamates rule enforcement and knowledge inference functionalities through a Java application.
OntoReg was initially developed on paper and then formalized in the OWL language with Protégé–OWL graphical editor. The key objective behind its design was to formalize the tacit human concepts in a way that would adequately represent physical (which is subdivided into regulatory and process domains) and abstract concepts and allow explicitly formal reasoning, which is both understandable by humans and processable by software, see Figure 1.

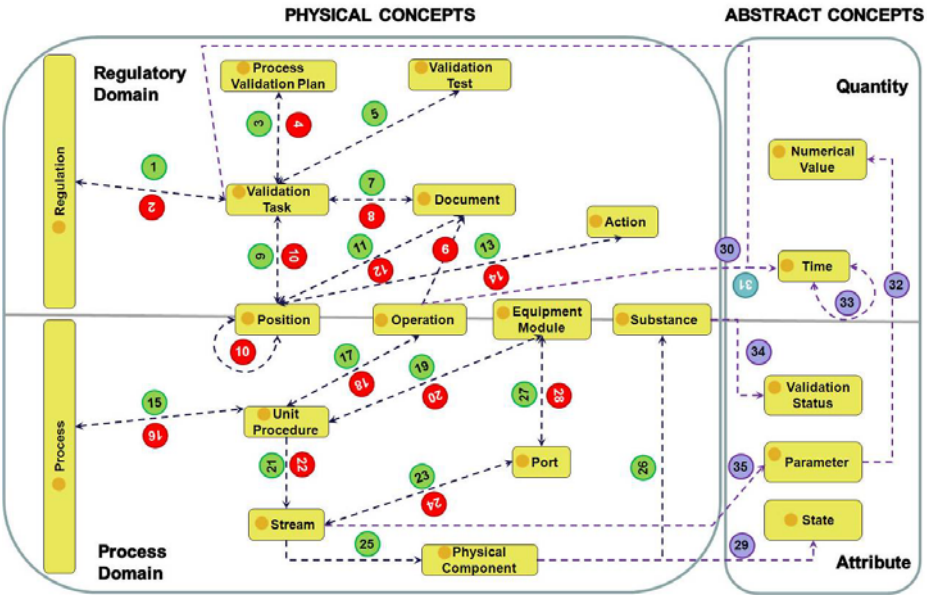


Figure 1. The main classes and interrelations (T-Box) in OntoReg.

A few process engineering ontologies, such as OntoCAPE [6] and POPE [7], were studied to capture relevant concepts that would help in the development of the Process Domain of OntoReg. Additionally one upper class ontology, namely SUMO [8], was examined but eventually deemed to be incompatible with description logics (DL) reasoning that is inherent to the OWL-DL language.
Figure 2 depicts how the ontological knowledge base OntoReg, lying at the heart of the whole infrastructure, is utilized within the application. The application is organized in three layers. The Knowledge Layer, where the regulatory and process concepts are encoded into the ontology in the form of classes, axioms and rules (constituting the Terminological Box or T-Box). The Inference Layer, where the rule engine is used to interpret and enforce the rules imposed on the ontological knowledge base. In addition, a Description Logics reasoner is also present in this layer and is used to classify the taxonomy and compute inferred individual types that are asserted by axioms. The third layer is the Java Integration Layer, which integrates the first two layers along with

presenting them through a front end application for user interaction. The Java engine and GUI (Graphical User Interface) of the OntoReg platform are designed to aid the user in the process of resolving logical inconsistencies and furnishing critical missing data in the knowledge base. All these tools interact with each other to provide an integrated environment that allows deduction of new facts while constraining existing ones in order to ensure regulation compliance of a pharmaceutical process.

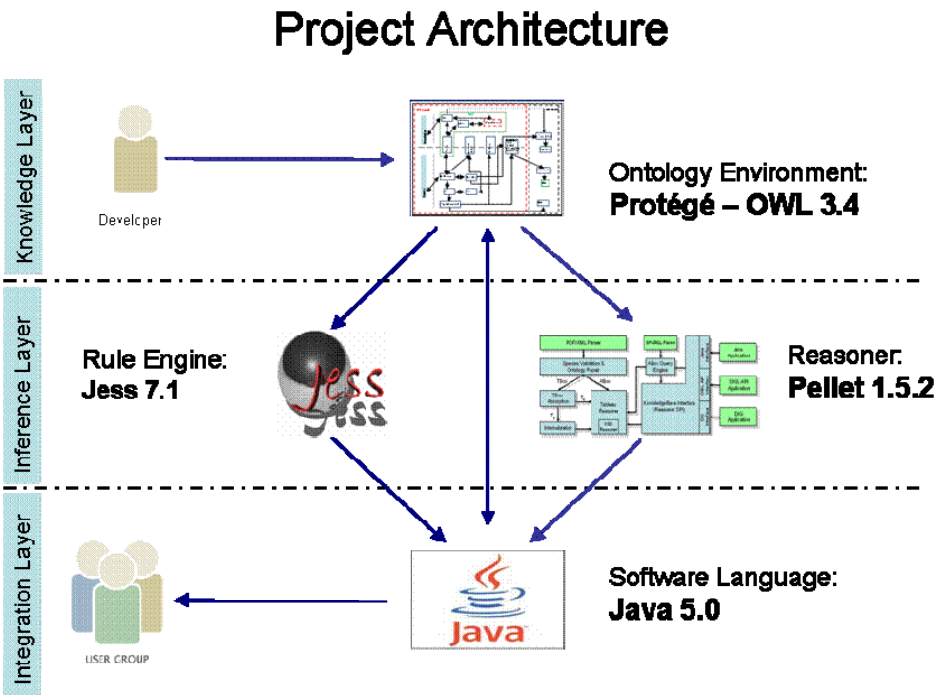


Figure 2. Architecture of the OntoReg platform.

3. Results

The structure of the regulatory domain in OntoReg has been derived after carefully going through the European Union Drug Regulatory Network Legislation (Eudralex) Guide 2007 [9] and Haider's Pharmaceutical Validation Master Plan [10]. After the conceptualization of the ontology was complete, a few small portions of the Eudralex guidelines were selected (varying from Cleaning Validation to Process Validation and Substance Qualification) for the creation of these case studies that demonstrate the capabilities of the current platform in the enforcement of the given regulations on an Aspirin production process [11] introduced to the knowledge base. Discussed below is one simple case study that shows the use of the OntoReg platform and the value added towards automated regulatory compliance. This particular example shows how OWL axioms can be used to enforce some basic requirements such as the need that a process material is to be supplied by a vendor who is approved by the organization (the pharmaceutical company). The axioms involved in enforcing this requirement are

1 `StartingMaterial isSuppliedBy some (PharmaSupplier and (hasSupplierAssessment has Approved))`

2 `StartingMaterial isSuppliedBy exactly 1`

3 `PharmaSupplierAsssess isResponsibilityOf some (PurchasingOfficer and (performs only Assessing))`
4 `PharmaSupplierAsssess hasReport some SupplierAssessmentReport`
5 `PharmaSupplierAsssess hasReport exactly 1`
6 `ApprovedSupplier hasSupplierAssessment has Approved`
7 `DisapprovedSupplier hasSupplierAssessment has Disapproved`
8 `ApprovedSupplier isPatientOf some PharmaSupplierAssessTask`

Now, to start with, assume that salicylic acid (which is used in the Aspirin production process) has no associated vendor or supplier. This would lead to the warning shown below indicating that StartingMaterial *Salicylic_Acid* does not have an "isSuppliedBy" relation with any Suppliers (violates Axiom2), see Figure 3.

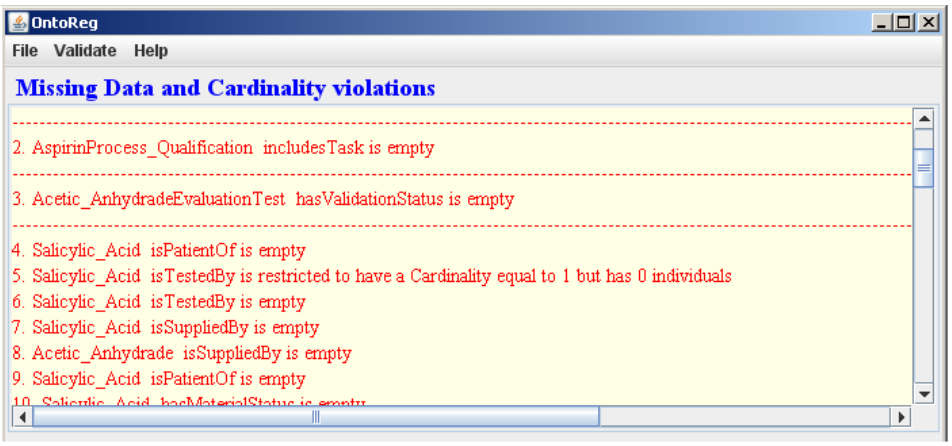


Figure 3. The Missing Data warning: a supplier has to be assigned for Salicylic_Acid.

In the next step the user sees the cardinality violation and chooses *CBAChemicals* as the supplier of *Salicylic_Acid*. Assume that *CBAChemicals* is an individual of the DisapprovedSupplier class (violates Axiom1). After selecting this supplier, and on running a consistency check, the reasoner displays an error message stating that *CBAChemicals* can not be the supplier of a StartingMaterial (not shown).
The user realizes his/her mistake and finds that the supplier of *Salicylic_Acid* was indeed *ABCChemicals*, an ApprovedSupplier, and not *CBAChemicals*; which fixes this validation problem.
Further case studies, which are more lengthy and complicated and that make use of SWRL rules, can be found at [12] and show the real power of an automated system and the amount of regulatory burden it can circumvent. It should be noted though that these case studies were chosen to be merely proofs of concept studies to showcase the capabilities of a semantic approach towards an automated process validation framework

4. Conclusions and Future Work

The majority of the pharmaceutical industry still uses paper documents, spreadsheets and conventional databases for process and regulatory knowledge storage and validation. This approach is error prone and does not take advantage of the advances in the field of information technology and knowledge management. This project proposes an ontological knowledge base, OntoReg, which encapsulates the process and regulation knowledge in OWL ontologies, and uses axioms and rules as means to enforce

pharmaceutical regulatory guidelines on the drug manufacturing process. This approach treats regulatory compliance as a parallel activity to the development of drug manufacturing process rather than as an afterthought, ensuring rapid elimination of gross oversights and inconsistencies that could prove costly once the drug making process is under review from the regulators. By partly automating the match between regulatory guidelines and a pharmaceutical process, and reporting any inconsistencies, this approach also avoids human errors. Furthermore, the knowledge-based assistance enables a more structured approach to regulatory compliance, and makes assumptions explicit. Valuable know-how, previously residing only in an expert's mind, can be preserved. Overall, a knowledge-based approach may catalyze the whole effort, decreasing compliance time and costs.

Aside from the obvious step of extending the scope of the OntoReg ontology to account for a larger set of principles from the EU Guidance on GMP [9], a number of extensions are currently being implemented in our group:

- Natural language processing to parse regulations and extract knowledge in the form of ontologies and rules.
- Provision of the OntoReg functionality as a set of web services.
- Tailoring OntoReg's functionality in terms of two sets of users: engineers from a pharmaceutical company and validation experts working for a regulatory agency.

Other interesting pieces of future work, such as the checking of the internal consistency of a set of regulations and/or between different sets of regulations (e.g. FDA vs. MHRA) have been pointed out to be important by the pharmaceutical and other regulated industries. In order for OntoReg to evolve, it is crucial for it to be applied to real industrial processes.

References

[1] F. Yip, P. Ray, N. Paramesh, 2006, "Enforcing Business Rules and Information Security Policies through Compliance Audits", in BDIM '06 International Workshop, pp. 81- 90.

[2] J. Conley, 2000, "Automating Regulatory Compliance", Sci. Computing & Instrumentation.

[3] R.F. Boland, 2006, "Reduce Business Risk with a CMS", Chemical Engineering Prog, 102, 10.

[4] V. Venkatasubramanian, C. Zhao, G. Joglekar, A. Jain, L. Hailemariam, P. Suresh, P. Akkisetty, K. Morris, G. V. Reklaitis, 2006, "Ontological informatics infrastructure for pharmaceutical product development and manufacturing", Comps&ChemEng, pp. 1482–1496.

[5] Y. Sure, J. Angele, S. Staab, 2002, "OntoEdit: Guiding Ontology Development by Methodology and Inferencing", in Proceedings of ODBASE.

[6] A. Wiesner, J. Morbach, W. Marquardt, 2007, "An overview on OntoCAPE and its latest applications" AIChE Annual Meeting.

[7] L. Hailemariam, P. Suresh, P. Akkisetty, G. Joglekar, S.H. Hsu, A. Jain, K.R. Morris, G.V. Rekaitis, P. Basu, V. Venkatasubramanian, 2008, "Excipient Interaction Prediction: Application of the Purdue Ontology for Pharmaceutical Engineering (POPE)", in ESCAPE-18 proceedings, Lyon, France.

[8] Suggested Upper Merged Ontology (SUMO), 2008, http://www.ontologyportal.org/.

[9] Medicines and Healthcare products Regulatory Agency (MHRA), 2007, "Rules and Guidance for Pharmaceutical Manufacturers and Distributors".

[10] S. I. Haider, 2002, "Pharmaceutical Master Validation Plan: The Ultimate Guide to FDA, GMP, and GLP Compliance".

[11] E. E.A. Handel-Vega, 2001, "Synthetic Procedure for the Manufacture of Aspirin." vol. US 6278014 B1, U. S. P. Organization.

[12] M.B. Sesen "OntoReg: A Support System for the Management of Regulatory Compliance of Biopharmaceutical Processes", 2008, MSc dissertation, Dept. of Eng. Sci., Univ. of Oxford.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Multi-objective Game Models for Chemical Industrial Park

Jia Xiao-Ping[a,c], Wang Fang[b], Xiang Shu-Guang[a]

[a]*Center for Chemical Process Engineering, Qingdao University of Science and Technology, Qingdao, 266042, China. E-mail address: cnjiaxp@yahoo.com.cn*
[b]*College of Environmental and Safety Engineering, Qingdao University of Science and Technology,Qingdao, 266042, China.*
[c]*The Guangdong Provincial Laboratory of Green Chemical Technology, School of Chemistry and Chemical Engineering, South China University of Technology, Guangzhou, 510640*

Abstract

Chemical industrial park is a practical form of industrial symbiosis network to emerge to achieve the goal of resource sharing among the participating chemical plants. The aim of this work is to present a systematic framework for industrial symbiosis networks (ISN) evaluation and modeling. A mathematical multi-objective game model is proposed to enhance understanding of ISN issues and motivate improvements of economic, resources and environmental sustainability. Results suggest that the effectiveness of policies designed to improve chemical industrial park management may be enhanced by this proposed framework. It allows the system designers to consider simultaneously the synthesis of ISN according to the existing data to establish synergies between each two participating plants.

Keywords: Sustainability, Game theory, Multi-objective programming, Industrial symbiosis

1. Introduction

Sustainability of industrial systems is becoming increasingly important. To improve economic and environmental performance of industrial system, the members of this system discover the opportunities for internal recycle and external mass or energy exchange and exploring market opportunities for waste [1]. Industrial symbiosis, an important concept and practical application in the field of industrial ecology, is about identifying and using such synergies and linkages between members. It engages traditionally separate industries in a collective approach to competitive advantage involving physical exchange of materials, energy, water, and/or by products [2]. From the theoretical point of view, ISN are also called joint production processes or multi-plant chains (webs). Tools for systematic designing and evaluating such industrial ecology networks are still in their infancy [3]. Therefore, the goal of this work is to present a systematic framework for IS evaluation and modeling.

2. Multi-objective modeling

2.1. Conceptual structure

Each plant considers its resource needs and waste aspects separately in the traditional industrial systems. However, separate plants could be linked through resource sharing

from a practical point of view. ISN provides highly inter-dependent relationship between two plants, exchanging materials and energy in a mutually advantageous manner [4]. Fig.1 shows the conceptual structure of ISN. It should be noted that this structure is addressed by considering a large redundancy of plants and connections. It allows the designers to consider simultaneously the synthesis of ISN according to the existing data collected by the designers seeking to establish synergies with other participating plants.

ISN would provide one or more of the following benefits across inter-plants: 1) reduction in the use of raw materials as resource inputs; 2) reduction in pollution; 3) increased systemic energy efficiency leading to reduced systemic energy use; 4) reduction in the volume of waste products requiring disposal; 5) increase in the amount and types of process outputs that have market value [5].

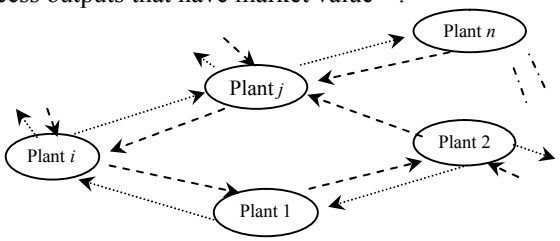


Fig. 1. The conceptual structure of industrial symbiosis network

2.2. Mathematical model

A multi-objective model is proposed mathematically in this subsection. The objectives involve economic, resources and environmental sustainability. Economic objective relates to indicators including investment, net profit, net present value, operation and maintenance cost, and environmental cost. Resource sustainability includes energy, water, and raw materials input, mass of recycled materials, energy and water, and mass of by-products exchanged. Environment objective includes potential environmental impacts associated with releases (including global warming potential, ozone depletion potential, eutrophication potential, photochemical oxidation potential, acidification potential, and human toxicity potential), mass of air and water pollutant released, and mass of solid waste disposed. Each objective is obtained from the combination of indicators using analytic hierarchy process respectively. Thus, the multi-objectives model problem is formulated as shown in the following.

$$\max \quad (f_{\mathrm{ECO}} - f_{\mathrm{RES}} - f_{\mathrm{ENV}})$$

$$s.t. \quad P\ (m_{\mathrm{I}} - m_{\mathrm{O}}) = 0 \tag{1}$$

$$E\ \ (e_{\mathrm{I}} - e_{\mathrm{O}}) = 0 \tag{2}$$

$$R^k = A^k W^k \tag{3}$$

$$\mathrm{A}^k = (a_{ij}{}^k)_{n\times n} \qquad k,i,j = 1,2,....,n \tag{4}$$

where f_{ECO}, f_{RES}, f_{ENV} represent economic, resources, environmental sustainability objectives respectively; *P, E* represent mass and energy transformation matrices respectively; *m, e* represent mass and energy matrices of ISN respectively; Subscripts I, O mean inputs and outputs of the members respectively. The system of equations (1~2) represents the mass and energy balances of ISN. In equations (3~4), *R* means by-products or energy accepted from plant *i* to plant *j*. *W* means by-product or energy

supplied from from plant *j* to plant *i*. *A* means transformation matrix of plant *k* between feed and by-product or energy discharged.

3. Multi-objective game model

3.1. multi-objective game problem

The members in the ISN are independent because of market economics. The interest and production strategy of an individual member are also independent. Therefore, there are two or more decision makers in this multi-objective situation described in section 2. The strategies of the members are the responses to other members, possibly conflicting each other. It increases the complexity of ISN. Game theory is introduced for solving this problem [6]. Pareto Nash-Equilibrium solution will be obtained. The following subsection will provide a procedure for this multi-objective game problem under discussion.

3.2. game theory with goal programming

A hybrid multi-objective model approach is proposed to solve such problem. It consists of ISN (outer) and separate plant (inner) levels. Game theory is used at the outer level. Game theory with goal programming is used at the inner level. This approach provides an environment between multi-objective problems and decision-makers to incorporate different types of interaction. The stepwise approach is presented (see Fig. 2).

Step 1: Construct ISN problem to find possible alternatives;

Step 2: Construct game theory model *G* using the following procedure. There are *n* players Y_i (i=1, 2, …, n). Each plant has its own production strategies S_s (s=1, 2, … , m); the corresponding payoffs are F_i (i=1,2, …, n), for example, F_1 (f_{ECO}, f_{RES}, f_{ENV}) is the payoffs of plant 1.

Step 3: Find the Nash-equilibrium using linear programming.

Step 4: Construct goal programming problem in the inner level such as

$$G(Y', S', F', g, Q_i, w) \tag{5}$$

$$\min \quad Z_i = Q_{il} \sum_{j=1}^{k_i} (w_{ilj}^- d_{ij}^- + w_{ilj}^+ d_{ij}^+), \quad l = 1, 2, \mathrm{L} \ L$$

$$F'_{ij}(s_1, s_2, \mathrm{L} \ s_n) + d_{ij}^- + d_{ij}^+ = g_i^j,$$

$$d_{ij}^-, d_{ij}^+, s_n \geq 0$$

$$s_i \in S'_i \quad i = 1, 2, \mathrm{L} \ m$$

Where *g, Q, and w* represent goal constraints, the preference of objective, and goal deviations respectively.

Step 5: Implement goal programming for each objective of each plant and obtain the intersection set *C* of solution C_i.

Step 6: If *C* is null set, then minimum Q_i are deleted and return. Else if *C* exists, then *C* is determined as Pareto Nash-equilibrium solution.

Step 7: Optimize the objectives of each plant and calculate the objectives of ISN.

Step 8: Decision-making of the production strategies of each member in ISN.

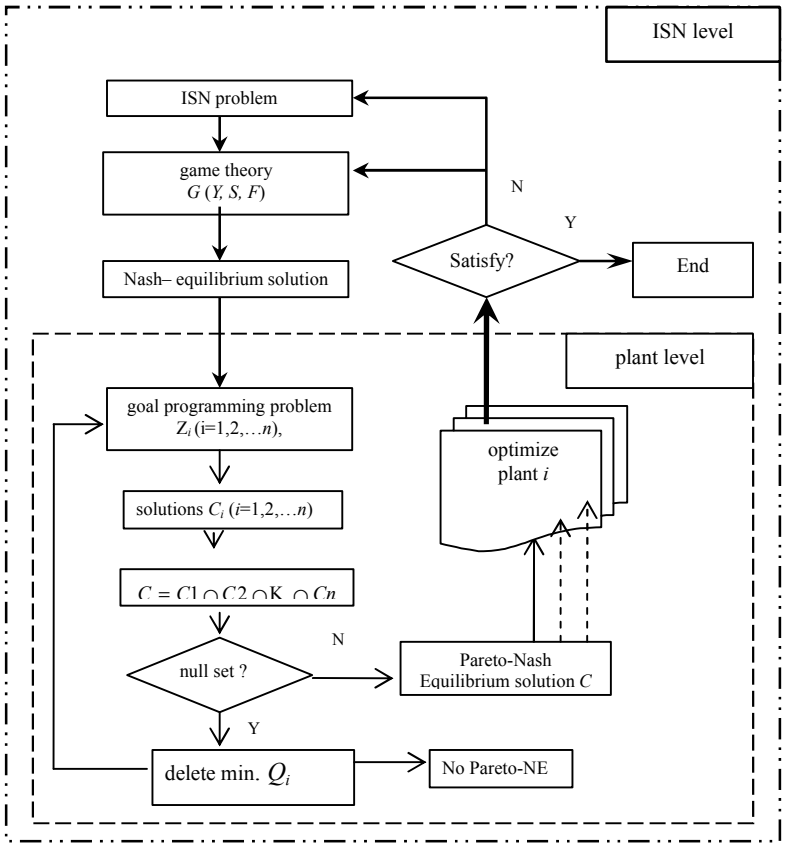


Fig.2. Framework for a multi-objective game problem solving procedure

4. Case Study

4.1. Network Description

This section discusses the symbiotic relationship among four independent companies in a chemical park. Their collaboration represents inter-company exchange of sulphur by-products.

Chemical company A is a complex of basic chemicals and chemical fertilizer. Because of using low level equipments and technologies, large quantities of waste are generated and emitted into air, water and soil. SO_2 emission from boiler flue gas and off-gases of sulfuric acid plant is the most important pollutant. In chemical company C, large quantities of low concentration SO_2 in off-gases are produced from metal sulphide roasters, sulfuric acid plant and lead sintering machines. A large quantity of high concentration sulfur slag is also generated. It has two pyrite-based sulfuric acid plants using double-conversion and double-absorption process. Chemical company D is the off-gas centralized treatment plant.

4.2. Optimal network

Due to high environmental awareness and business opportunities, these four companies will generate new revenues and cost savings for the companies involved and reduced SO_2 pollution to air, water, and land. The new revenues cover SO_2 emission reduction

fee and high value added sulfur products. Based on the materials exchanges among companies, inter-companies sulfur-based flow optimal diagram is shown in Figure 3. Thanks to a limited extent, we present initial results for multi-objective game models only rather than the detailed problem solving procedure in this paper.

4.2.1. PA: Chemical company A

To recover sulphur by-product produced from PC, PA will build a 200,000 tonnes per year sulphur based sulfuric acid plant using double-conversion and double-absorption process to replace the existing pyrite-base one. It will also reduce the off-gas emission, which is sent to off-gas centralized treatment plant. A part of sulfuric acid product is sent to company B to produce sodium hydrosulfite. The SO_2 from the boiler flue gas and off-gases of sulfuric acid plant is recovered, and then used as the raw materials in company B. A part of sulfuric acid is sent to sulfuric acid-based titanium dioxide production process.

4.2.2. PB: Chemical company B

SO_2 recovered from off-gas centralized treatment plant is used as the raw material in PB. SO_2 from the boiler flue gas off-gas in PB is recovered as the raw material. Sulphur from PC is purified and used as the raw material to produce sodium hydrosulfite.

4.2.3. PC: Chemical company C

To fulfill PC's requirements for compliance with sulphur emission standards and profits, low concentration SO_2 is separated into three parts, as shown in Figure 3. In the PC Company itself, Dynawave scrubber gas cleaning technology and Topsoe's wet gas sulfuric acid technology (WSA) are used to recover sulphur and produce commercial-grade concentrated sulfuric acid. The off-gas from WSA plant is sent to a final centralize treatment plant. The second part is sent to PA for producing sulfuric acid. The rest of low concentration SO_2 is sent to PB for producing sodium hydrosulfite. High concentration sulfur slag is used to produce sulphur, which then is sent to PB for producing sodium hydrosulfite.

4.3. Results

The total SO_2 emission reduction is up to 15,000 tonnes per year. The estimated output of industrial symbiosis network in the park is: 0.17 million tonnes of sulfuric acid, 60,000 tons of sodium hydrosulfite, 56,000 tonnes of sulphur.

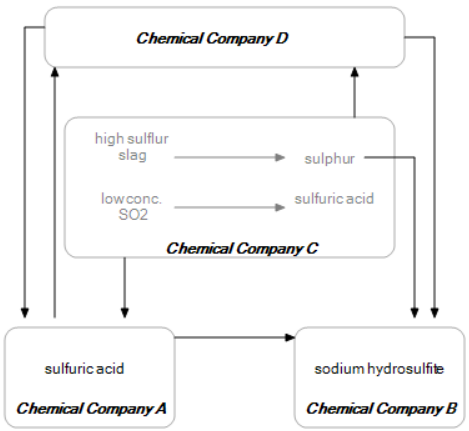


Fig.3 Final network for sulfur-based industrial symbiosis

5. Conclusions

This paper introduced the concept and structure industrial symbiosis network (ISN) to improve the sustainability of industrial systems. The framework for ISN evaluation and modeling is presented. To solve the multi-objective game model problem, a two-level, hybrid multi-objective approach is proposed. The development of ISN offers a potential regional platform to contribute to ensuring efficient utilization resources while minimizing the negative impacts on the environment. To illustrate the concept and potential application, this paper introduced sulfur oriented ISN as the case study. It consists of interlinked production of sulfur related products. These production networks utilize nearly all sulfur by-products to generate new revenues and reduce environmental emissions and disposal costs. It is expected that this approach provides an interactive environment between problems and decision-makers for modeling of industrial systems in the future research.

Acknowledgements

The authors gratefully acknowledge the research funding from Natural Science Foundation of China (40601037), Guangdong provincial laboratory of green chemical technology (GCT200803) and Doctoral Funds of Qingdao University of Science and Technology.

References

Lou Helen, M. A. Kulkarni, et al., A game theory based approach for emergy analysis of industrial ecosystem under uncertainty, Clean Technologies and Environmental Policy, 6(2004)156

Marian R. Chertow, Industrial symbiosis: Literature and Taxonomy, Annu. Rev. Energy Environ., 25(2000)313

Allen, D. T., R. S. Butner, Industrial Ecology: A Chemical Engineering Challenge, Chem. Eng. Prog., 98(11)(2002) 40

Manahan, Stanley, Industrial Ecology: Environmental Chemistry and Hazardous Waste, Lewis Publishers, Washington D.C., 1999

Gertler,N., Industrial Ecosystems: Developing sustainable industrial structures. Master thesis, Department of Civil & Environmental Engineering, MIT, USA, 1995

Fudenberg D., J. Tirole, Game theory, MIT press, Cambridge, Mass., USA, 1991

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A semantic information model for data integration across the chemical process design process

Andreas Wiesner, Jan Morbach, Wolfgang Marquardt

AVT-Process Systems Engineering, RWTH Aachen University, 52056 Aachen, Germany

Abstract

Information integration during the design process of chemical plants is a long-standing and not sufficiently solved problem in industrial practice to date. The major challenge identified is the capturing of the information's semantics. Within an ongoing research project an ontology-based approach for information integration in process engineering design projects is developed. This contribution sketches the semantic information model applied in the integration software, which is based on the formal ontology OntoCAPE.

Keywords: Ontology, CAPE, information modeling, information integration, design process

1. Introduction

The design of a chemical plant is characterized by a sequence of design phases, which involves different departments, disciplines, and contractors. In order to handle the design information, many software tools are used to support the project teams in the different phases, and a multitude of documents in different electronic formats are produced. Some of these software tools are domain-specific (i.e. process simulators, CAD or CAE systems, etc.) and others are generic and independent of the requirements of chemical engineering design processes (i.e. word processors, project management systems, etc.). However, the contents of these documents and data files typically are heavily intertwined and overlapping due to strong dependencies between the information generated in the different phases. Unfortunately, the information models of domain-specific software tools suffer from the lack of a well-structured, standardized information representation [1,2]. Almost all of them lack capabilities for interoperability and are thus mainly applied in a stand-alone manner. In order to cover design data over the entire design process, many information modeling activities, e.g., STEP [3], have been carried out, not only but also in the process engineering domain. ISO 15926 [4] is a prominent example. These formats are intended to be used for data exchange in order to integrate software tools during the engineering design process.

However, the most prevalent problems of information integration have not been successfully addressed by any of the current approaches: (i) consistency checking, (ii) the determination of dependencies between the contents of the documents, or (iii) the homogenization of different proprietary data formats. One reason for that seems to be an insufficient capturing of the semantics of the information modeled. Hence, we are exploring the use of more expressive modeling approaches to facilitate the capturing of the detailed semantics of the modeled information. Within an ongoing research project an ontology-based software prototype for the integration (consolidation and reconciliation) of distributed design data stemming from different design phases is developed. This contribution focuses on the prototype's semantic information model

which is based on the domain ontology OntoCAPE [5,6]. For details on the conceptual design of the software prototype and its implementation we refer to [7,8].
The remainder of the paper is organized as follows: Section 2 gives a brief overview on the typical design phases and the resulting documents. Section 3 introduces the principles applied to formally represent design aspects in OntoCAPE. In Section 4, the connection between the design aspects in OntoCAPE and the data stored in the software tools required for compliance checking is shown. Finally, Section 5 concludes the contribution.

2. Design phases and its corresponding documents

Like in any other engineering domain, the chemical plant design process can be subdivided in a number of distinct phases [9]. Typically in the first phase, the requirements for the material products to be produced by the chemical plant are fixed. Besides various business objectives (e.g. the aspects of market demand and the current and expected supply by competitors for the estimation of production capacity), the material product properties identified by the chemists are of major importance for the actual process design. These requirements form the basis for a first description of the design problem. However, the requirements cannot be completely captured and expressed by common CAE systems such that formal representations of requirements are rarely available in real-world projects. Typically, they are stored as text files, data sheets or simply presentation slides.
The design problem formulation is further refined in the conceptual design, where the major conceptual decisions on raw materials, on the chemical synthesis route, on the process structure, and even on the strategy for plant operation are taken. These conceptual considerations are refined in front-end engineering and detailed and completed during basic engineering [9,10]. At the end of basic engineering all major design data of the plant are fixed. The important documents at the end of front-end engineering are the PFD and so-called process data sheets for the most important pieces of equipment as well as a skeleton P&ID for basic engineering.
Basic engineering is followed by detail engineering, where the design data fixed during basic engineering are used to specify all pieces of equipment of the plant including all the instrumentation and control systems in full detail. The design process is completed at the end of detail engineering. The result of the design process is a set of complete specifications which are on a level of detail to allow procurement of all parts and construction of the plant on site in the subsequent plant lifecycle phases. This phase is usually documented by means of full-fledged P&ID's (which represents the major equipment items and their main dimension, but no geometrical details) and on a more detailed information level attained by isometric drawings and 3D models.

3. Representation of design aspects in OntoCAPE

OntoCAPE is a formal ontology for the domain of Computer-Aided Process Engineering (CAPE). An ontology is an explicit specification of a conceptualization, typically involving classes, their relations, and axioms for clarifying the intended semantics [11]. A formal ontology, in particular, captures the consensual knowledge of an application domain in such a way that it can be reused and shared across software systems. OntoCAPE formally represents domain knowledge about CAPE, with the intention of enabling the construction of software for different tasks such as knowledge management or plant design. Terms corresponding to classes defined in OntoCAPE are written in *italic* font in the following.

OntoCAPE is organized according to the principles of general systems theory[1] and systems engineering[2]. It introduces important systems-theoretical and physicochemical primitives such as *system*, *property*, *value*, etc., and specifies their mutual relations (e.g., a *system* is characterized by its *properties*, each of which can take numerous *values*). It also establishes the notion of so-called *aspect systems*, which represent a system considered from a particular viewpoint[3] which are used to partition complex systems into manageable parts. For a comprehensive description of *technical systems,* i.e. a system that has been developed through an engineering design process, five designated viewpoints are of major importance [1]: the *requirements* and the *function* of the *system*, as well as its *realization*, its *behavior*, and its *performance*. However, within this contribution we focus on the first three viewpoints as they are mainly related to the major design documents generated in a typical design process. For a comprehensive description we refer to [12].

First of all, the precise meaning of the *requirements* and *function* viewpoint is clarified. The ontological representation of function in design is a long-standing research issue. Various definitions of the function concept have been proposed in the literature; for a review of those, see for example [13]-[15]. For our purposes, we adopt the definition of [13], who defines function as desired behavior. According to these authors, two interpretations of the function concept must be distinguished for a *technical system*: function seen from an environment-centric viewpoint and function seen from a device-centric viewpoint. In OntoCAPE, a system's *requirements* specify the desired effect of the *technical system* on the environment. From the perspective of *requirements*, the *technical system* is viewed as a black box: Its structure and the underling physical and technical principles are not considered; only the effect on the environment is specified. However, a system's *function* describes the intended behavior of a *technical system* from a device-centric perspective. To indicate the system's *function* of a *technical system*, the conceptual design of the *technical system* must be specified in terms of the underlying physicochemical and/or technical principles. As an example, consider the design of a process unit. The system's *requirements* can be stated by describing the effect that the process unit shall exert on the processed materials (e.g., "separate dispersed particles from a liquid"). Yet to specify the system's *function*, one needs to consider the physical or technical principles based on which the desired effect is going to be achieved (e.g., decide whether the separation is realized by means of sedimentation, centrifugation, or filtration).

The third *aspect system* to be described here is the *realization* viewpoint. It reflects the physical (or virtual) constitution of the *technical system*. In case of a chemical process system, the *realization* describes the system's physical structure with respect to its geometrical and mechanical properties, e.g. comprising the equipment and machinery required for materials processing. Generally, the *realization* gives a static description of the technical system which is comparable to a technical specification, as it is typically created in an engineering design project to specify the technical system that is to be built. In this context, it is important to remember that a system realization holds only

[1] General systems theory is an interdisciplinary field that studies the structure and properties of systems [16].

[2] Systems engineering can be viewed as the application of engineering techniques to the engineering of systems, as well as the application of a systems approach to engineering efforts [17].

[3] The adoption of a viewpoint is a technique for complexity reduction that is widely used in systems engineering. A viewpoint is an abstraction that yields a specification of the whole system restricted to a particular set of concerns [18].

information pertaining to the system itself; information that specifies how to realize a technical system (e.g., assembly instructions) does not form part of the *realization*.
Finally, important relations and dependencies exist between the *aspect systems* of a *technical system*: In the course of a design project, the *requirements* built the cornerstones of the design project. Thus, the system's *function* as well its *realization* has to comply with them. This is usually guaranteed by the fact that the system's *requirements* are directly transformed into system's *functions*, which specify the conceptual design. Similarly, the system's *function* is detailed into the system's *realization* at the stage of basic design.

4. Realization for compliance checking for design data

Due to the lack of software vendor solutions to remedy interoperability, most of the major chemical and engineering companies have established in-house solutions for data exchange. Typically, these solutions are based on XML such that tailored converters allow the import and export of the respective design data. However, the major problem in today's engineering departments is to achieve consistency for the data set within the course of an entire design project. Regardless of the project phase, it must be ensured that the suggested design still complies with requirements as envisioned in the first phase of the design process, e.g., each refinement has to meet the constraints formulated in an earlier design phase. Not being compliant with predefined requirements often leads to unnecessary revision of the design data which usually causes considerable extra cost. Therefore, the novel software prototype, which is currently under development, aims at two goals [8]: Firstly, it allows for data exchange of XML files between tools which support object-oriented data export/ import; secondly and most importantly it allows for comprehensive consistency checks. To that end, the main emphasis is placed on compliance checking of (i) technical realization data (e.g., the interconnection of flanges with different internal diameters or process parameters exceeding the stability limits of some plant equipment) as well as of (ii) product specification data (e.g. product X must not exceed temperature T_{max} to avoid isomerisation of X to X*).
In the following, we give a simplified example on how the proposed information model can be employed to guarantee the required data consistency by considering example (ii) referring to product specification constraints. Such constraints, usually formulated at the very beginning of the design project, may initially be described, according to the principles stated in OntoCAPE: The *temperature property* characterizing *material* (chemical) X may take at most the *value* T_{max}, regardless of the *(sub)system* comprising the *material*. This requirement has to be formally defined by the classes and relations provided by OntoCAPE and stored in the knowledge base. It can then be used for consistency checks with the data generated in the following phases in the design project. The data captured in a PFD, as provided at the end of conceptual design, may be adequately described by the *aspect system* of *function*. Typically a PFD is an aggregation of blocks and streams which can be represented by special (sub)systems in OntoCAPE. Blocks may be represented by so-called *process steps*, i.e. a certain material processing procedure, and streams are depicted by so-called *process states* which represent the collection of *properties* of a certain *material* produced in the associated *process step* [19]. Accordingly, all *process steps*/*process states* referring to *material* X, which are characterized by *temperature properties,* can now be reconciled to the former formally stated *requirements*. Proceeding to the next stage of the design project, i.e. front-end engineering/basic engineering, the data are documented by process data sheets and (skeleton) P&IDs. At this point of the design project the data correspond to *plant*

items actually to be constructed. Accordingly the *realization* aspect system is employed subsequently. The realization viewpoint generally consists of the major plant items such as *equipment*, *instrumentation* and *piping* which are again special subsystems and thus conform to the principles stated before [19]. Since the equipment process data are typically determined in this design phase, each *equipment* processing *material* X can be checked against the *process steps* specification for consistency to guarantee that the range of the *temperature property* is within limits. Finally, if it comes to the stage of basic engineering/detail engineering, the equipment is further refined to a level of detail to allow construction. Also the remaining plant items, essentially instrumentation and piping, are specified. To meet the consistency constraints relating to temperature, the technical specification data of *equipment* may be reconciled against the process data and the configuration of *instrumentation* and *piping* respectively may be checked against the *requirements* alternatively. Despite the fact that the technical specification of equipment, instrumentation and piping is achieved in the same design phase, it is typically done by different disciplines and thus different engineers. Hence, the alignment of technical realization data is an additional issue which, however, can be checked with our approach as well. Due to the lack of space we refer to [7,8] for a more detailed description on that issue.

Thus, coherent information integration, in the sense of obtaining consistent data sets, can be achieved by the formulation of constraints. Engineers have to handle a variety of constraints and their exceptions in the course of a typical design project which have to be classified in categories. The formulation of constraints and their exceptions is best achieved expressing them declaratively as rules. Hence, all aforementioned consistency checks are realized via rules. These rules represent declarative knowledge in the form "if A then B", where A and B are statements about the information that is expressed by means of ontological terms. By choosing a suitable ontology language [20] and inference engine (reasoner) [21], which is capable of deductive reasoning, the rules can be executed. Considering example (ii) above such a rule may state (informal representation): **IF** X is a *chemical material* **AND** it is contained in a *(sub)system* **AND** the *material* is characterized by a *temperature property* **AND** the corresponding *value* exceeds T_{max} **THEN** the *(sub)system* is declared inconsistent.

5. Conclusions

This contribution presents an outline of the semantic information model of a novel, ontology-based approach for information integration for process engineering design data. The model is based on the formal ontology OntoCAPE, which provides the necessary principles and vocabulary to formally express information such that automatic data reconciliation by means of suitable software components (reasoners) is achievable across the phases of a design process. Therefore, we introduced the representation of certain viewpoint on the design data within a design project such that design data stemming from different phases and disciplines may be defined according to its information content. Hence, this approach is capable of successfully handle: (i) hidden dependencies between the contents of the documents applied in the design project; and (ii) compliance checking for the design data based on the categorization in so-called aspect systems. So far, the compliance checking between the viewpoints of realization and requirements took priority for the development of the software prototype. It has been tested against small to medium size amounts of real plant data and has been able to fulfill all needs with respect to technical realization as well as product property constraints. Even highly complex real world data could be inferred correctly. Thus, the

approach constitutes the semantic basis for successful information integration as envisioned by industrial practitioners.

The research described herein has been funded by the German National Science Foundation (DFG) as part of the Transfer Unit 61 of CRC (SFB) 476 “IMPROVE” [22].

References

[1] B. Bayer, (2003), Conceptual Information Modeling for Computer Aided Support of Chemical Process Design. Fortschritt-Berichte VDI: Reihe 3, Nr. 787. VDI-Verlag, Düsseldorf.

[2] B. Bayer and W. Marquardt, (2004), Towards integrated information models for data and documents, *Comp. Chem. Eng.*, 28 1249-1266.

[3] ISO10303, (1998), Standard for the Exchange of Product data

[4] ISO15926, (2003), Industrial Automation Systems and Integration – Integration of Life-Cycle Data for Process Plants Including Oil and Gas Production Facilities

[5] W. Marquardt, J. Morbach, A. Wiesner, A. Yang, (2009), OntoCAPE – A (Re)Usable Ontology for Chemical Process Engineering. Springer, Berlin. In preparation.

[6] J. Morbach, A. Wiesner, W. Marquardt, (2009), OntoCAPE 2.0 – a (Re)Usable Ontology for Computer-Aided Process Engineering . *Comp. Chem. Eng.*, in press.

[7] J. Morbach and W. Marquardt (2008), Ontology-based integration and management of distributed design data. In: Nagl, M., Marquardt, W. (eds.): *Collaborative and Distributed Chemical Engineering*. Springer, Berlin, pp. 647–655.

[8] A. Wiesner, J. Morbach, W. Marquardt, (2008), Semantic data integration for process engineering desgin data. In: *Proceedings of the 10th International Conference on Enterprise Information Systems – ICEIS 2008*, pp. 190-195

[9] M.S. Peters, K.D. Timmerhaus, (1991), Plant Design and Economics for Chemical Engineers. McGraw-Hill

[10] J.M.Douglas, K.D. Timmerhaus, (1988), Conceptual Design of Chemical Processes. McGraw-Hill

[11] M. Uschold and M. Grüninger, (1996), Ontologies: principles, methods and applications. *Knowl. Eng. Rev.* 11 (2), 93–136.

[12] J. Morbach, B. Bayer, A. Wiesner, A. Yang, W. Marquardt, (2009), OntoCAPE 2.0 – The Upper Level. Technical Report (LPT-2008-25), Lehrstuhl für Prozesstechnik, RWTH Aachen University.

[13] B. Chandrasekaran, J.R. Josephson, V.R. Benjamins, (1999), What are ontologies, and why do we need them? *IEEE Intell. Syst.* 14 (1), 20–26.

[14] S. Szykman, R.D. Sriram, W.C. Regli, (2001), The role of knowledge in next-generation product development systems. ASME J. *Comp. Inf. Sci. Eng.* 1 (1), 3–11.

[15] Y. Kitamura, and R. Mizoguchi, (2003), Ontology-based description of functional design knowledge and its use in a functional way server. *Expert Syst. Appl.* 24 (2), 153–166.

[16] L.v. Bertalanffy, (1968), General System Theory: Foundations, Development, Applications. Braziller, New York.

[17] B. Thomé, ed. (1993), Systems Engineering: Principles and Practice of Computer-based Systems Engineering. Wiley, New York.

[18] IEEE, (2000), IEEE Recommended Practice for Architectural Description for Software-Intensive Systems. IEEE Standard 1471-2000, Institute for Electrical and Electronics Engineering, New York.

[19] A. Wiesner, J. Morbach, B. Bayer, A. Yang, W. Marquardt, (2009), OntoCAPE 2.0 – Chemical Process Systems. Technical Report (LPT-2008-29), Lehrstuhl für Prozesstechnik, RWTH Aachen University.

[20] M. Kifer, G. Lausen, J. Wu, (1995), Logical foundations of object-oriented and frame-based languages, *Journal of the ACM* (42:4).

[21] Ontobroker, (2009), http://www.ontoprise.de/en/home/products/ontobroker/

[22] M. Nagl and W. Marquardt, eds. (2008), Collaborative and Distributed Chemical Engineering: From Understanding to Substantial Design Process Support. Springer, Berlin.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A Parallel Computing Scheme for Large-scale Logistics Network Optimization Enhanced by Discrete Hybrid PSO

Yoshiaki Shimizu, Takatobu Miura and Masashi Ikeda

Department of Production Systems Engineering, Toyohashi University of Technology, 1-1 Hibarigaoka, tenpaku-cho,Toyohashi 441-8580, Japan

Abstract

Recently, we have concerned the strategic optimization on logistic network design problems and developed an efficient two-level solution method using a meta-heuristic method like tabu search. To cope with extremely large-scale problems, in this paper, we propose an extended algorithm for the parallel computing that utilizes the analogy between the algorithm and the master-worker configuration of PC cluster. To enhance the efficiency, we developed a modified discrete algorithm of particle swarm optimization (PSO) that can deal with binary decision variables. Then, we applied it to a parallel procedure that requires only very small overhead for the parallel computing. Finally, we confirmed that the proposed method can bring about high performance for the parallel computing that is suitable for the present goal and circumstance through numerical experiments.
Keywords: Logistic network optimization, PSO, Parallel computing algorithm, Master-worker PC cluster.

1. Introduction

Noticing the growing importance of supply chain management in manufacturing, we have engaged in a logistics network optimization using a method termed hybrid tabu search, and deployed it in various situations (Wada, Shimizu & Yoo, 2005; Shimizu, et al., 2006; Shimizu, et al., 2007). Compared with previous studies made extensively (for example, Drezner, 2002), these studies can cope with large-scale problems in a practical and flexible manner. To cope with extremely large-scale problems in a numerically effective manner, in this paper, we propose an algorithm for the parallel computing aimed at reducing the computation time and improving the optimality at the same time. For this purpose, we have given a discrete algorithm of particle swarm optimization (PSO) so as to deal with binary variables that stand for open or close of sites and introduced a mutation operation like genetic algorithm into the conventional PSO. Eventually, we will show such a population-based algorithm is particularly suitable for the parallel computing targeted at global logistics optimization. After showing the algorithm of such modified discrete PSO, we will outline the procedure of the parallel computing and discuss its properties. Finally, we provide a few numerical experiments to validate the effectiveness of the proposed method.

2. Parallel Computing for Logistic Optimization

2.1. Strategic Logistics Network Model

Let us take a logistic network composed of plant, distribution center (DC), and customer as shown in the left side of Fig.1. Then consider an optimization problem formulated

as the following mixed-integer programming problem.

$$Min \quad \sum_{i\in I}\sum_{j\in J}(T1_{ij}+C_i)\cdot e_{ij}+\sum_{j\in J}\sum_{k\in K}(T2_{jk}+H_j)\cdot f_{jk}+\sum_{j\in J}F_j\cdot x_j$$

$$\text{subject to} \quad \begin{cases} \sum_{j\in J} f_{jk} = D_k \quad \forall k\in K & (1)\\ \sum_{i\in I} e_{ij} = \sum_{k\in K} f_{jk} \quad \forall i\in I & (2)\\ \sum_{j\in J} e_{ij} \le S_i \quad \forall i\in I & (3)\\ \sum_{k\in K} f_{jk} \le U_j\cdot x_j \quad \forall j\in J & (4)\\ x_j\in\{0,1\} \quad \forall j\in J & \\ e_{ij}, f_{jk}\ge 0,\ \forall i\in I,\ \forall j\in J,\ \forall k\in K & \end{cases}$$

where notations denote as follows:

C_i: production cost per unit amount at plant i
D_k: demand of customer k
F_j: fixed-charge cost for opening DC j
H_j: holding cost per unit amount at DC j
S_i: upper bound for production at plant i
$T1_{ij}$: transport cost from plant i to DC j per unit amount
$T2_{jk}$: transport cost from DC j to customer k per unit amount.
U_j: upper bound of holding capacity at DC j
e_{ij}: shipped amount from plant i to DC j
f_{jk}: shipped amount from DC j to customer k
x_j: take 1 if DC j is open, otherwise 0
I, J, K: index sets of plants, DCs and customers, respectively

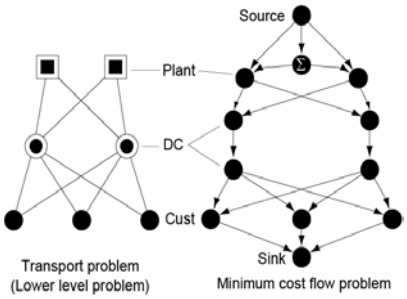


Fig.1 Transformation procedure into MCF graph

The objective function is the total cost composed of transportation costs, production costs at plants, holding costs at DCs, and fixed-charge costs for opening the DCs. On the other hand, we impose the constraint on demand satisfaction for every customer, Eq.(1); input-output balance at each DC, Eq.(2); available amount of product from each plant, Eq.(3); and upper bound of holding capacity at each DC, Eq.(4). In addition, binary decision variables are introduced for selecting opening DCs and non-negative real variables for deciding transport amounts.

To solve this kind of problem, we successfully applied the hybrid tabu search. Its development relies on the fact that meta-heuristic method is amenable for solving the upper-level sub-problem that decides the available DCs and refers to a 0-1 programming problem. On the other hand, the route selection problem from plants to customers via DCs in the lower-level reduces to a linear program after binary variables are fixed. Additionally, as illustrated in the right side of Fig.1, it can be further transformed into the minimum cost flow problem by introducing a few virtual nodes and edges. Consequently, we can solve the lower-level sub-problem by the graph algorithm much faster than by the linear programming method.

Then, noticing the analogy between such a two-level algorithm and the master-worker configuration of the PC cluster (See Fig.2), we proposed a suitable framework for the parallel computing for this logistics optimization. There, the master PC engages in deciding the DC location. That corresponds to the upper-level sub-problem while each worker tries to solve the route selection sub-problem in the lower-level with DC sites appointed by the master. Generally speaking, the global optimum is achievable more efficiently by the population-base algorithm than by the single-start local search algorithm like tabu search. For this purpose, we employ the PSO after giving its modified algorithm mentioned below. It can readily realize a multi-walk search in the parallelism since each worker can entirely engage in the route selection problem as shown in Fig.3.

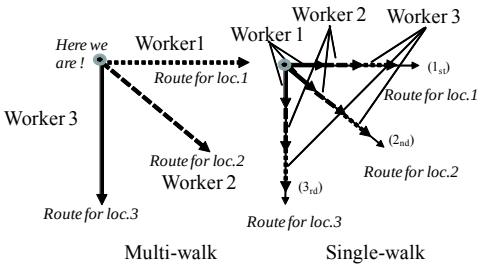


Fig.3 S*chematic idea of single and multi*-walk searches

2.2. Modified PSO for 0-1 Programming Problem

The PSO is a meta-heuristic optimization technique developed recently (Kennedy and Eberhart, 1995, 1997) after the behavior of bird flocking or fish schooling (swarm), and known as a powerful global optimization method of real variables.

Members of swarm communicate with each other and adjust their own positions and velocities based on the information regarding good positions both of their own and the swarm. In practice, the position and the velocity are updated through the following formulas, respectively.

$$x_i(t+1) = x_i(t) + v_i(t+1), \tag{5}$$

$$v_i(t+1) = w \cdot v_i(t) + r_1 b(p_i - x_i(t)) + r_2 c(y_n - x_i(t)), (i = 1,2,\ldots,N_p), \tag{6}$$

where t denotes a generation, N_p a swarm size, and w an inertial constant. Moreover, b and c are constants that will properly guide each member to a good position. On the other hand, r_1 and r_2 are random values in the range [0,1], p_i is the best position seen by the member i (personal best), and y_n is the global best position seen by the swarm (net best).

The algorithm is simple and outlined below.

Step 1: Set t=1. Initialize $x_i(t)$ and $v_i(t)$ randomly within the admissible range of these

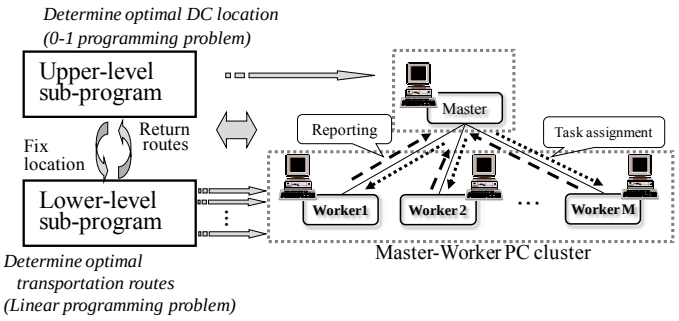


Fig.2 Analogy between two-level algorithm and Master-worker PC cluster

values, each p_i to the current position, and y_n to the position having the best fitness among the swarm.

Step 2: For each member, do the following: obtain $x_i(t+1)$ and $v_i(t+1)$ according to Eqs.(5) and (6), respectively, and evaluate the new position. If it outperforms p_i, update p_i, and if it outperforms y_n, update y_n.

Step 3: If the stopping condition is satisfied, stop. Otherwise, let t:=t+1, and go back to Step 2.

Now, to cope with 0-1 programming problems, we modified the original method as a discrete PSO. First, let us give a binary code of decision vector so that it can describe whether the site is open (1) or close (0) as exemplified in Fig.4. By noting that the bit-wise calculation of $p_i - x_i(t)$ or $y_n - x_i(t)$ in Eq.(6) takes anyone of (-1, 0, -1), we update the velocity at the next generation as shown in Table 1.

Table 1 Decision method of the next velocity instead of Eq.(6)

Pattern	$v_i(t+1)$	Probability
$p_i(t) - x_i(t) = y(t) - x_i(t) = v_i(t)$	$v_i(t)$	r_a (= 0.997)
$y(t) - x_i(t) \neq p_i(t) - x_i(t) = v_i(t)$	$v_i(t)$	r_b (= 0.7)
$p_i(t) - x_i(t) \neq y(t) - x_i(t) = v_i(t)$	$v_i(t)$	r_c (= 0.6)
$p_i(t) - x_i(t) = y(t) - x_i(t) \neq v_i(t)$	$y(t) - x_i(t)$	r_d (= 0.5)
$p_i(t) - x_i(t) \neq y(t) - x_i(t) \neq v_i(t)$	Select randomly	-

Apparently, these probabilities must satisfy the following relations such that: $1>r_a>r_b>r_c>r_d>0$. Likewise, the position is decided by the bit-wise calculation between $x_i(t)$ and $v_i(t+1)$. That is, $x_i(t+1)$ becomes 0 if the resulting value is less than 0, while 1 if greater than 1.

To enhance the performance of the algorithm, we introduced an idea to guide forcibly some inferior members to appropriate positions (we call this "warp" hereinafter). Since this warp operation is not ruled by Eqs.(5) and (6), it can be viewed as a mutation operation in the genetic algorithm. We prepared the following three methods to carry out the warp.

Warp 1: Let move the worst member to the vicinity of the net best. The vicinity is given by the flip-flop of only one locus of the net best. This presents the intensification to the search at the expense of the loss of diversification.

Warp 2: Let move the member chosen randomly at an arbitrary position. This makes increase the diversification but disturb the intensification.

Warp 3: First let choose an arbitrary member, and then randomly copy codes from net best. This is viewed as another local search around the net best.

2.3. Algorithm for Parallel Computing by Modified Discrete PSO

In what follows, we outline the tasks assigned to master and worker PCs, respectively.

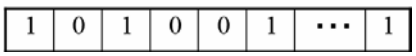


Fig.4 Code of location (open:1, close:0)

Master PC

Step 1: Notify the number of DC site to every worker PC.

Step 2: Enter "waiting" mode

Step 3: Compare the reported personal best with the current net best. Then update if the reported personal best outperforms the current best.

Step 4: If the reported personal best is inferior to the net worst, rewrite the current net worst, and indicate the worker to carry out the warp as far as a certain condition is satisfied. Otherwise, return the current net best to the worker.

Step 5: If a certain convergence condition is satisfied, let every worker stop, and let the net best be the final solution. Otherwise, go back to Step 2.

Worker PC

Step 1: According to the notified DC number, decide the initial location.

Step 2: Solve the route selection problem. If the current search could find the personal best, update it. Otherwise, go to Step 4.

Step 3: Contract with the master and follow its indication, i.e., either warp or to receive the latest net best.

Step 4: Re-locate the DC sites based on the PSO algorithm mentioned already and go back to Step 2.

2.4. Evaluation of the Parallel Computing

There are several factors that are popularly known to affect on the performance of the parallel computing, i.e., the load imbalance between master and worker, the granularity and the frequency of communication, and the overhead for the parallelism. Regarding these factors, the proposed framework has nice properties. First, it can provide a wonderful timing that enables us to almost completely avoid the idle time due to the load imbalance. Regarding the information exchange between the master and worker, not only its amount is small but also its frequency is low due to the multi-walk implementation (The search starting from the different point is totally referred to each worker PC). Moreover, synchronization between every PC is quite unnecessary regarding information exchange. Consequently, each member is to be controlled by the different net best by the virtue of asynchronization. Without paying any particular attentions, this brings about the increase in manifoldness that is essential for the meta-heuristic algorithm. Due to these effects, we can make the overhead for the parallelism very small and improve the performance of the algorithm at the same time.

To evaluate the total performance of the parallel computing, the following two indices are commonly used in the homogeneous computing environment. When it takes $T(P)$ CPU time using P number of PCs, speedup rate is defined by $S(P) = T(1)/T(P)$, and the efficiency by $\eta = S(P)/P$. Hence, these values ideally become $S(P)=P$ and $\eta = 1$, respectively.

3. Numerical Experiments

Using up to 9 PCs with various specifications, one of which works as the master while the others as workers, we performed numerical experiments. They are run under the Debian Linux and using the library of MPICH (Pacheco, 1997).

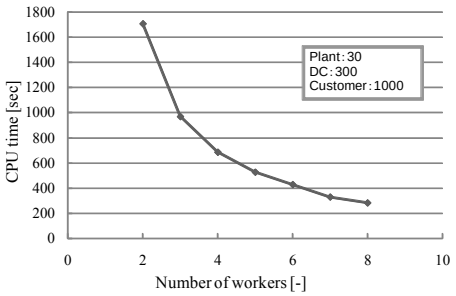


Fig.5 Effect of number of workers on speed up

We prepared benchmark problems with the problem size such as $|I|$=30, $|J|$=300, and $|K|$=1000. When number of evaluation has reached at 8000, we stop the search. To properly instruct the poor worker regarding the warp, we gave a condition that relates to the improvement rate, i.e., $\Delta Cost / \Delta Time \leq D$ where $\Delta Cost$ and $\Delta Time$ denote the difference of objective function between the solutions and of evaluation time, respectively. Moreover, D is an appropriate constant controlling the timing of the warp. Adjusting this condition, we can avoid the undesirable warps,

i.e., at the beginning of the search and during the active updating stage. Table 2 summarizes the results (averaged over ten trials) among various warp methods. From Table 2, we know every warp method achieved the better performance than the original method without any warp, i.e., less computation time and objective value. Besides the better convergence property, the speedup rate shown based on the wall-clock time in Fig.5 also validates the effectiveness of the proposed method. Thereupon, the parallelism effect is splendid and moderately degraded with the increase in the number of workers. From these facts, we can expect to efficiently solve much larger-scale problem using the larger cluster by the proposed method.

Table 2 Comparison of performance among the warp methods

	Objective value	CPU time
Original (No warp)	33601	275.9
Warp 1	33464	223.2
Warp 2	33581	219.6
Warp 3	33569	212.7

4. Conclusions

To solve strategic optimization problems for large scale logistic network design, we have developed a practical and efficient method for the parallel computing and implemented it into a master-worker configuration. Thereupon, we are interested in the PSO and applied it after giving its modified algorithm to handle binary variables.

This is a novel approach of PSO for realizing information sharing and exchange and multi-walk in the parallel computing. Additionally, due to the analogy between the two-level solution algorithms and the master-worker configuration of PC cluster, we have shown the developed procedure can provide very nice properties regarding the reduction of the overhead and the idle time. As a side effect, the idea can increase manifoldness that is essential for the population-based meta-heuristic approach.

Through numerical experiments, the effectiveness of the proposed procedure is confirmed. Relying on these results, we can conclude that the proposed approach is promising for large-scale and complicated real world applications, targeting at global logistics of chemical industries.

References

J. Kennedy and R. C. Eberhart, 1995, Particle Swarm Optimization, Proceedings of IEEE International Conference on Neural Networks, Piscataway, NJ., 1942-1948

J. Kennedy and R. C. Eberhart, 1997, A Discrete Binary Version of the Particle Swarm Algorithm, Proceedings of World Multiconference on Systemics,Cybernetics and Information, Piscataway, NJ., 4104-4109

Z. Drezner and H. W. Hamacher (eds.), 2002, Facility Location, Springer, Berlin

P. S. Pacheco, 1997, Parallel Programming with MPI, Morgan Kaufmann Publisher

Y. Shimizu, S. Matsuda and T. Wada, 2006, A Flexible Design of Logistic Network against Uncertain Demands through Hybrid Meta-Heuristic Method, Proc. 16th ESCAPE, 2051-2056

Y. Shimizu, T. Wada and Y. Yamazaki, 2007, Logistics Optimization Using Hybrid Meta-heuristic Approach under Very Realistic Conditions, Proc. 17th ESCAPE, Bucharest, Romania

T. Wada, Y. Shimizu and J. K. Yoo, 2005, Entire Supply Chain Optimization in Terms of Hybrid in Approach, Proc. 15th ESCAPE, 1591-1596

Y. Shimizu, H.Kawamoto, 2008, An Implementation of Parallel Computing for Hierarchical Logistic Network Design Optimization Using PSO, Proc. 18th ESCAPE, 605-610

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Supply chain optimization for bioethanol production system in Northern Italy: environmentally conscious strategic design

Andrea Zamboni[a], Fabrizio Bezzo[a], Nilay Shah[b]

[a]*DIPIC – Dipartimento di Principi e Impianti di Ingegneria Chimica,Università diPadova, via Marzolo 9, Padova, I-35131, Italy (andrea.zamboni@unipd.it).*
[b]*CPSE – Centre for Process Systems Engineering, Imperial College London, South Kensington Campus, London, SW7 2AZ,UK*

Abstract

This work proposes a spatially explicit modeling framework developed for the strategic design of biomass-based fuel supply networks. A Mixed Integer Linear Programming (MILP) model is formulated and solved for the multi-objective optimization of the whole supply network in terms of both operating costs and greenhouse gas (GHG) emissions. The economics are assessed by means of Supply Chain Analysis (SCA) techniques. The environmental performance is evaluated in terms of greenhouse gas emissions, by adopting a Well-to-Tank (WTT) approach. The framework capabilities in steering the strategic design path are assessed through a real-world case study.

Keywords: Supply Chain Design, Multi-objective Optimization, Life Cycle Analysis.

1. Introduction

Concerns related to oil depletion as well as environmental issues like global warming have been driving a worldwide debate centered on the global energy supply question. In the transport sector, liquid biofuels have been identified as a viable and realistic alternative to fossil energy in achieving energy supply security and climate benefits in the near future. Italy has complied with the European guidelines (Directive 2003/30/EN) by setting the minimum blending of biofuels at 3% in energetic content for 2009 and 5.75% for 2010. Bioethanol is globally considered the current best viable alternatives for a short-term partial gasoline replacement. However, some doubts still persist on whether ethanol production from starchy biomass brings any effective economic and environmental benefits [1]. In particular, production costs and greenhouse gas emissions strongly depend on the geographical conditions and technological experience of each country which the production system is operating in. As a consequence, the transition path from an oil-based fuel system to a biomass-based one should be driven by specific tools capable of assessing the economic, technological and social interactions along the entire supply chain (SC).

This paper proposes a spatially explicit modeling framework developed for driving the strategic design of biofuels supply networks under economic and environmental criteria. The design task is formulated as a Mixed Integer Linear Program (MILP) that accounts for the simultaneous minimization of the SC operating costs as well as the environmental impact in terms of greenhouse gas (GHG) emissions. The economics have been assessed by means of Supply Chain Analysis (SCA) techniques, focusing on

biomass cultivation site locations, ethanol production capacity assignment and facilities location as well as transport system optimization. The environmental performance of the system has been evaluated in terms of GHG emissions, by adopting a Well-to-Tank (WTT) approach. The proposed tool has been applied to design the forthcoming Italian corn-based ethanol system. The reported results demonstrate the framework capabilities in providing valuable insights in steering the design of strategic infrastructures.

2. SC and LC analysis

The economic as well as the environmental assessment of a production system by means of SCA and LCA techniques requires a rigorous preliminary work focused on the characterization of the supply chain components as well as of the logistic nodes in terms of infrastructure availability, operating costs and environmental impact. Fig. 1 depicts the general structure of the system. A biomass-based fuel supply chain can be divided into two main substructures: the first one concerns with the fuel upstream production network and involves biomass cultivations, biomass delivery and fuel production sites; the latter is related to the downstream product distribution to the demand centers.

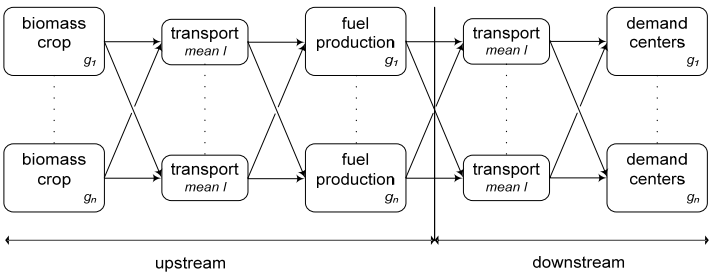


Figure 1. Bioethanol supply chain superstructure

2.1. Biomass cultivation

Considering the first generation production technology as the best solution over a short-term horizon, corn has been identified as the most convenient biomass for ethanol production in Italy. Province specific data regarding yield and land availability for corn crops have been collected from Governmental institution databases (www.istat.it; www.apat.it). Production costs have been derived from actual data (www.crpv.it) and then divided into fixed costs and yield dependent ones, so as to create a grid-dependent set of parameters. This approach has been validated by comparing the obtained results with actual production costs collected from regional databases. The same procedure has been applied to adapt specific data regarding Italian corn cultivation practice [2], in order to obtain a set of grid specific global emission factors resulting from an interactive spreadsheet based tool [3] conceived according to the IPCC guidelines.

2.2. Transport system

The Northern Italy infrastructure includes a full-scale range of transport options available for industrial purposes. The whole set has been considered defining trucks, rail, barges and ships as possible delivery means. Assuming the delivery of goods as an additional service provided by existing actors already operating within the infrastructure, transport costs have been evaluated by multiplying the freight loads by a unit transport cost [€/(t·km)] gathered from the literature [4] and then validated comparing the resulting outcomes with actual data coming from confidential information. Global emission factors specific for the different transport means have

been taken from [5]. The availability of each transport option has been characterized through the definition of feasibility constraints on transport means suitability as well as through a tortuousity factor accounting for the different routes that each specific transport means has to go through.

2.3. Ethanol production

The dry grind process has been considered in characterizing the production facilities. Ethanol production costs are sensitive to plant capacity due to the economy of scale effect on capital and operating costs. This important issue has been taken into account considering four different discrete capacity ranges. For each range specific production and capital costs have been calculated by means of a purpose-designed financial model [6]. Also in this case, the global emission factors have been calculated as in § 2.1.

2.4. Cost allocation and emission credits

In calculating the operating costs as well as the environmental impact related to the entire supply chain operation, cost allocation and emission credits have been taken into account. Therefore, in evaluating the operating costs a quote for DDGS has been detracted considering a 20% allocation factor deduced from literature data [7]. According to [8], no credits have been assigned for land use. Hence, GHG emission credits have been only associated to products displacement by using DDGS as a valuable product for animal feed substitution.

2.5. Demand centers

Internal gasoline depots have been assumed as demand centers for bioethanol. Data about provincial gasoline demand perspectives for 2010 as well as internal depots location and maximum distribution capacity have been collected from Governmental databases (dgerm.sviluppoeconomico.gov.it). Given the demand driven nature of the optimization problem considered, demand assignment to depots must be solved as a secondary distribution problem before the supply chain optimization. The distribution problem has been based on the modeling approach commonly applied in the optimization of fuel distribution systems [9].

3. Optimization framework

Strategic decisions in designing a biofuel production network deal with geographical location of biomass cultivation sites, logistic definition of transport system and location as well as capacity assignment of production facilities. Therefore the key variables are:

1. biomass production in each cell;
2. distribution processes for biomass to arrive at the production facility;
3. location and capacity of production facilities;
4. distribution processes for fuel to be sent to blending terminals.

The core of the modeling framework is based on the MILP approaches commonly adopted in optimizing the strategic design of multi-echelons supply networks for renewable fuels: the work of [10] has been taken as standard reference in defining the superstructure of the problem; whilst concerning with the environmental issue the approach of [11] has been adapted into a more flexible mathematical formulation devised to avoid excessive efforts in both calculation and data collection.

Eq. 1 shows the economic objective function representing the SC operating costs (*TDC)*:

$$\min TDC = \frac{FCC}{a} CCF + PC + TC \qquad (1)$$

where the facilities capital costs (*FCC*) are annualized through a capital charge factor (*CCF*) and divided by the SC operating period (*a*); the additional terms are the production costs (*PC*), accounting for both biomass and fuel production, and the transport costs (*TC*).
The environmental objective function accounting for the GHG emissions in the supply chain (*TI*) results from the sum of the life cycle stage contributions as shown by Eq. 2:

$$\min TI = \sum_p f_p F_p \qquad (2)$$

where F_p is the reference flow for each life cycle stage p and f_p is the global emission factor, representing the overall GHG emitted at stage p per unit of reference flow. The global emission factor (f_p) has been obtained by grouping the emission coefficient of each substance, together with the related damage factor based on the concept of 100-year global warming potentials as specified by the IPCC. All the cost variables as well as the stage-specific reference flows depend on design variables related to the ethanol and biomass production (P^T_{ig}), the product demands (D^T_{ig}) and the mass flows between grids ($Q_{ilgg'}$). The SC behavior is then captured through logical constraints and mass balances. Eq. 3 and Eq. 4 are demonstrative examples of this feature:

$$P^T_{ig} = D^T_{ig} + \sum_{l,g'} (Q_{ilgg'} - Q_{ilg'g}) \qquad \forall i,g \qquad (3)$$

$$Q^{\min}_{il} X_{ilgg'} \le Q_{ilgg'} \le Q^{\max}_{il} X_{ilgg'} \qquad \forall i,l,g,g' \qquad (4)$$

where i identifies the product (ethanol or corn), l the transport option and $X_{ilgg'}$ is the binary decision variable that is assigned the value 1 if the transportation of product i by means of l is allowed from g to g', or 0 otherwise.

4. Case study

The modeling framework has been implemented to assess the 2010 demand scenario derived from the ethanol market penetration imposed by the current Italian regulations. Solving the two objective functions problem by adopting a multi-parametric approach [12] results in the trade-off set of Pareto optimal solutions reported in Fig. 2.

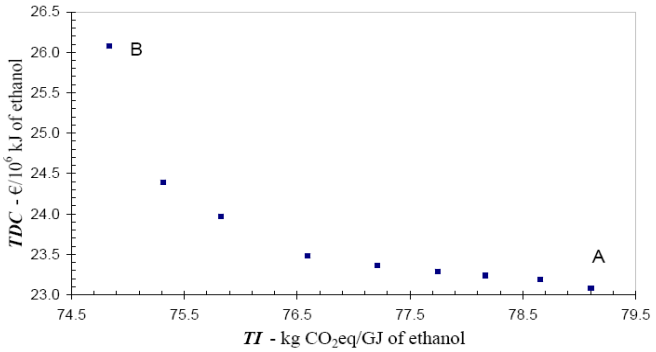


Figure 2. Multi-objective optimization: Pareto curve.

The shape of the curve reveals the expected conflict existing between environmental and economic performance. Indeed, in reaching the best optimum in terms of operating cots (point A as reported in Fig. 2, corresponding to a marginal operating costs of 23.1 €/10^6 $kJ_{ethanol}$) the environmental impact turn out to be 79.10 kg CO_2-eq/$GJ_{ethanol}$,

corresponding to a GHG reduction of about 8% compared to gasoline. This is not enough to meet the latest EU standards requiring biofuels to have a minimum of 35% of GHG emissions saving for them to be counted towards the target. However, even minimizing the impact (point B of Fig. 2) the resulting GHG emissions are anyhow too high to meet the minimum requirements, albeit the substantial economic effort required to reach the target: reducing the marginal impact down to 74.83 kg CO_2-eq/$GJ_{ethanol}$ (equal to 13% of GHG reduction) results in an increase of the overall operating costs up to 26.1 €/10^6 $kJ_{ethanol}$.
Further remarks also concerning with the market penetration of bioethanol can be based on the assessment of the performance indicators coming from the model implementation. For instance, promoting the maximum GHG mitigation in establishing the bioethanol industry would entail the need for more subsidies to fill the gap between the environmentally optimum costs performance and the economic ones (roughly amounting to 3 €/10^6 $kJ_{ethanol}$, about 0.1 €/$L_{ethanol}$). However, the corresponding environmental benefit is still not enough to satisfy with the EU standards. Therefore, in the particular case of the Italian corn-based ethanol production a well-advised strategy would address the design process under economic criteria. In this way a double benefit would be achieved: firstly the market penetration of bioethanol would be eased, so as to open the way to the oncoming second generation system; secondly, the resulting financial saving entails a bigger founds availability that might be assigned to support other energy sectors characterized by more effective environmental performance.

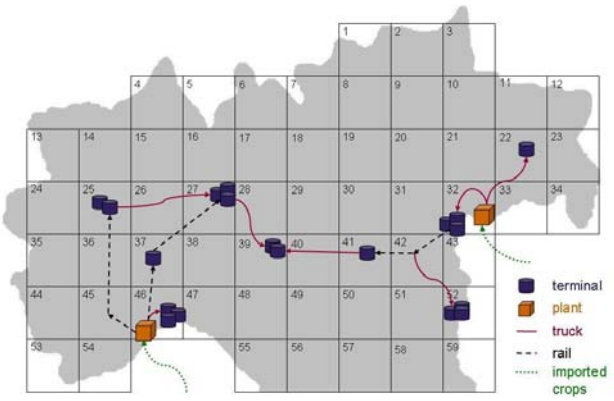


Figure 3. Costs optimization: supply network configuration.

Finally, an illustrative example of the network configuration corresponding to the costs-optimal solution is shown in Fig. 3. The graphical representation reveals important issues that both the environmental and the economic performance are dealing with. The biomass needs are met by importing corn from Eastern European Countries. Corn is directly shipped to the two production plants of the maximum capacity (about 250 kt/y) and located within the industrial areas close to the main ports of Venice and Genoa. This configuration allows the best economic performance in terms of both biomass supply costs, due to the lower price of the imported corn, and of ethanol production costs, positively affected by the scale factor. However, the longer delivery distance along with sub-optimal cultivation practice, that is often characterizing corn production in the Eastern area, are those factors that negatively affect the environmental performance. Indeed, moving along the curve shown in Fig. 2 from point A to point B, we assist to a gradual transition toward a network configuration characterized by a more decentralized fuel production system (four plant of the smaller capacity) and by

domestic biomass supply: this ensures better environmental performance in terms of goods distribution and corn production impact in spite of a drastic worsening of the system economics due to the negative scale factor as well as to the unprofitable biomass supply conditions.
A possible alternative in meeting the fuel demand would be the establishment of production plants abroad. However, this option has not been considered here mainly because it would not comply with the national policy (common in most EU countries) aiming at increasing national security in energy production.

5. Final remarks

A spatially explicit modeling framework for the strategic design of biofuel systems has been developed. The aim of the study was to build a general modeling tool that might be helpful to steer an economic and environmentally conscious design for biofuels supply chains. The optimization model applied in assessing the forthcoming Italian corn-based ethanol supply chain can be used to steer the strategic policy. The results coming from the case study application outline how the corn-based bioethanol supply network design should be based on economic criteria rather than environmental ones. This would grant an easier market penetration, as a direct consequence of the production costs minimization, and would require less governmental subsidies that thus could be deployed in more environmentally effective sectors.

6. Acknowledgement

Partial support from the University of Padova under Progetto di Ateneo 2007 (cod. CPDA071843): “Bioethanol from lignocellulosic biomass: process and equipment development” is gratefully acknowledged.

References

[1] C. B. Granda, L. Zhu, M. T. Holtzapple, 2007, Sustainable liquid biofuels and their environmental impact. Environmental Progress, 26, 233–250.
[2] C. Grignani, L. Zavattaro, 2000, A survey on actual agricultural practices and their effects on the mineral nitrogen concentration of the soil solution, Europ. J. Agronomy, 12, 251–268.
[3] G. Brown, J. Wood, A. Estrin, 2005, Bioethanol greenhouse gas calculator: Users’ guide, HGCA.
[4] L. Buxton, 2008, All Aboard, Biofuel International, 1.
[5] DEFRA, 2008, Guidelines to Defra’s GHG conversion factors: Methodology paper for transpor emission factors.
[6] G. Franceschin, A. Zamboni, F. Bezzo, A. Bertucco, 2008, Ethanol from corn: a technical and economical assessment based on different scenarios, Chem. Eng. Research and Design, 86, 488–498.
[7] Hammerschlag, R. (2006). Ethanol’s energy return on investment: a Survey of the literature 1990–present, Environ. Sci. Technol., 40, 1744–1750.
[8] M. A. Delucchi, 2006, Lifecycle analyses of biofuel, Draft Report, University of California, Davis.
[9] M. T. Kong, 2002, Downstream oil products supply chain optimisation. PhD Thesis, University of London.
[10] A. Almansoori, N. Shah, 2005, Design and operation of a future hydrogen supply chain: Snapshot model, Chem. Eng. Research and Design, 84, 423–438.
[11] A. Hugo, E. N. Pistikopoulos, 2005, Environmentally conscious long-range planning and design of supply chain networks, J. Cleaner Production, 13, 1471– 1491.
[12] V. Dua, E. N. Pistikopoulos,2000, An algorithm for the solution of multi-parametric mixed integer linear programming problems. Ann. Oper. Res., 99, 123–139.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

An Effective Decomposition Approach for Solving Large Supply Chain Oriented Pick-up and Delivery Problems

Rodolfo Dondo, Carlos A. Méndez, Jaime Cerdá
INTEC (UNL - CONICET) - Güemes 3450 - 3000 Santa Fe - ARGENTINA

Abstract

Pick-up and delivery problems (PDP) are receiving a growing attention in process systems engineering due to its close relationship with major supply chain issues. Its aim is to discover the best routes/schedules for a vehicles fleet fulfilling a number of transportation requests at "minimum cost". In the conventional PDP, each request defines the shipping of a given load from a specified pickup site to a given customer. However, in order to account for a wider range of logistics problems, the so-called supply-chain oriented PDP (SC-PDP) problem has been defined as a three-tier network of interconnected factories, warehouses and customers. Multiple products are to be efficiently delivered through this network in order to meet a set of given demands. The selected vehicle routes/schedules must satisfy capacity and timing constraints while minimizing transportation costs. The pickup points for each demand are decision variables rather than problem specifications and several commodities can be transported between sites. Moreover, every customer can be visited several times. The general SC-PDP has been represented as an MILP formulation that is able to address moderate size instances. In order to efficiently address large-scale SC-PDP problems, a decomposition method based on a column generation (CG) procedure is introduced in this work. In contrast to traditional CG approaches lying on dynamic-programming-procedures as route generators, an MILP formulation is here proposed to implicitly create the set of feasible routes/ schedules at the slave level of the method.

Keywords: supply-chain, pick-up and delivery, logistics, columns generation.

1. Introduction

A typical supply chain covers the procurement of raw materials from suppliers and their shipping to one or more factories, the conversion of such inputs into intermediate and final products, their shipping to warehouses or depots for intermediate storage and the delivery of products to retailers and customers (Simchi-Levi et al., 2004). To this end, the supply chain management field and the emerging area of enterprise wide optimization have focused the attention on developing tools to efficiently coordinate factories, warehouses and customers so that product requirements are all satisfied with the specified service level at the minimum system cost. The pickup and delivery problem (PDP), widely studied in the literature, involves orders matching the pick-up of a given load at one or more locations and the subsequent delivery to one/several locations. Despite the importance of this problem, it is clear that some new features must be added to the standard PDP to obtain a better representation of realistic supply chain networks. Some of the new features to be considered are: (1) each request may involve the shipment of specific quantities of commodities to one or more destinations; (2) several commodities may be transported on the same vehicle; (3) several alternatives

sources for each commodity may be available; (4) a shipment of a given commodity from a single source may have several destinations; (5) each vehicle may be operated on more than a single route if the total time spent on these routes is less than the specified maximum service time; (6) the problem may involve multiple depots with known inventories of commodities; (7) depots may arise as intermediate stops on the vehicles routes. To face these new features, Méndez et al. [2008] proposed a variation of the PDP which involves a set of facilities from which multiple products are to be efficiently delivered to many consumers in order to meet some demands while satisfying capacity and timing constraints. This problem, named as supply-chain-PDP (SC-PDP), was addressed through a new MILP formulation that defines the choice of pickup nodes for each product demand as a set of additional decisions to be made based on the geographical locations of product demands and inventories. The higher flexibility in the routes generation process allows to find more efficient operational strategies to manage complex multi-site distribution systems. However, in addition to the development of rigorous optimization methods for logistics problems, a lot of effort of the research community has been directed to the development of heuristic procedures capable of dealing with the inherent complexity of large-scale problems. Desrochers et al. [1992] presented a new technique widely known as the *Dantzig-Wolfe decomposition + column generation* approach that remains as the most efficient optimization technique for large routing problems. Although the original SC-PDP model can be directly used to find the optimal solution for instances of moderate size, this work embeds a columns generation procedure into the model in such a way that larger examples can be solved. It is worth noting that most Dantzig-Wolfe decomposition + column generation algorithms are based on the use of dynamic-programming-label-setting algorithms for generating routes. On the contrary, the proposed procedure generates fewer columns by using a MILP-based routes-generator that arises as the adaptation to the CG structure of the model proposed by Méndez et al. [2008].

2. Problem statement

Consider a three-tier network represented by a graph G[$\{I^+, \Gamma, B\}$; A], where $I^+ = \{i_1, i_2, ..., i_n\}$ denotes the set of pick-up nodes (factories and warehouses), $\Gamma = \{j_1, j_2, ..., j_{n'}\}$ is the set of delivery nodes (customers), $B = \{b_1, b_2, ..., b_l\}$ represents the set of vehicle bases, and $A = \{a_{ij} / i, j \in I^+ \cup \Gamma \cup P\}$ defines the net of minimum cost arcs among nodes. Several nodes $i \in (I^+ \cup \Gamma)$ may be related to the same geographical location if multiple visits may be considered. A distance-based traveling-cost matrix $C=\{c_{ij}\}$ and a travel-time matrix $\Gamma=\{t_{ij}\}$ are associated to the net A. The service times on pickup/delivery nodes $i \in (I^+ \cup \Gamma)$ are denoted by st_i. To fulfill the pick-up and delivery tasks, a set of vehicles $V = \{v_1, v_2, ..., v_m\}$ is available. The solution to the problem will provide a number of sequences of arcs, commonly called *routes*, such that: (1) each vehicle starts and ends the trip at a vehicle base *b;* (2) for each commodity type, all the customer demands are exactly satisfied through one or multiple visits; (3) the total quantity of products provided by pick-up nodes must not exceed the initial inventory; (4) each node $i \in (I^+ \cup \Gamma)$ is at most assigned to a single route; (5) the load transported by a vehicle must never exceed both the weight q_v and the volumetric qv_v restrictions; (6) a pick-up or delivery node $i \in (I^+ \cup \Gamma)$ must be serviced within its time-window $[a_i, b_i]$; (7) the duration of the trip for any vehicle v must be shorter than its maximum allowed routing time t^{max}. All these features must be simultaneously considered while minimizing the total cost of providing pickup or delivery services to every node $i \in (I^+ \cup \Gamma)$.

3. The column generation-based approach for the SC-PDP

The SC-PDP can be formulated as a set partitioning problem (SPP) as follows: let $P = \{\rho_1, \rho_2, \ldots, \rho_R\}$ be the set of *all* feasible routes. Let c_ρ be the cost of route ρ and let y_ρ be a binary variable denoting that route ρ is included in the optimal solution if y_ρ is equal to 1. Then, the objective is to select the set of minimum-cost feasible routes such that each request is fulfilled. So, the optimization problem can be formulated as follows:

$$Min \sum_{\rho \in P} c_\rho y_\rho \tag{1}$$

$$\sum_{\rho \in P} a_{i\rho} y_\rho \geq 1 \qquad \forall i \in \left(I^+ \cup I^-\right) \tag{2}$$

$$y_\rho = \{0,1\}$$

The parameter $a_{i\rho}$ indicates that node i belongs to route ρ if $a_{i\rho} = 1$. Otherwise, $a_{i\rho} = 0$. Since feasible routes may easily run into billions, it is not possible to realistically generate all routes (here also called *columns*), the CG approach handles this complexity by *implicitly* considering all routes through the solution of the linear relaxation of the SPP. Then, a portion of all feasible routes is enumerated and the problem linear relaxation is solved just considering this partial set. The solution to this problem is used to determine if there are routes not included in the partial set that can reduce the objective function value. Using the value of the optimal dual variables with respect to the partial routes set, new routes are generated and the linear relaxation is solved again. This procedure continues until the optimal solution to the linear problem cannot be improved with the addition of a new route. For the master problem of the SC-PDP, this basic SPP, widely used in many routing problems, must also consider inventory constraints for each product as well as the number of available vehicles. These new features leads to a new ***master problem,*** where the dual-variable-values are obtained by enumerating a feasible solution comprising the partial set of routes $P' = \{\rho_1, \rho_2, \ldots \rho_{P'}\}$. This formulation, named as the *reduced master problem* (RMP), is given below.

$$Min \sum_{\rho \in P'} c_\rho y_\rho \tag{3}$$

$$\sum_{\rho \in P'} a_{i\rho} y_\rho \geq 1 \qquad \forall i \in \left(I^+ \cup I^-\right) \tag{4}$$

$$\sum_{\rho \in P'} a_{i\rho} \alpha_{i\rho} y_\rho \leq l_{ip} \qquad \forall i \in I^+, p \in P \tag{5}$$

$$\sum_{\rho \in P'} y_\rho \leq n \tag{6}$$

$$0 \leq y_\rho \leq 1$$

where l_{ip} is the inventory of product p on supply site $i \in I^+$ and n is the number of available vehicles. Now, let us assume that $\bar{y}$ is the optimal solution to the RMP and let $\mathbf{J\Pi} = [\pi_c, \pi_l, \pi_n]$ be the corresponding vectors of optimal dual variables values for constraints (4), (5) and (6), respectively. Then, if we can find a new route minimizing the quantity $(c_\rho - \Sigma_{i \in I} a_{i\rho} \pi_c^i - \Sigma_{i \in I^+} \Sigma_{p \in P} a_{i\rho} \alpha_{i\rho} \pi_l^{ip} - \pi_n)$ and, this quantity is negative, a cheaper route has been found. Consequently, $\bar{y}$ and $\mathbf{J\Pi}$ are not optimal for the RMP. The column just found is added to the partial set and the problem is solved again. The procedure iterates until no columns with negative reduced costs are found. Finally, the

integer RMP can be solved for finding the best set of routes. Although in some cases the solution generated may not be the global optimal solution to the SC-PDP, it is generally very close. To find the optimal one, the procedure should be embedded into a branch & bound algorithm because some routes that are not generated when solving the relaxed RMP may be needed to solve the integer one.

The feasible routes are generated by solving the ***slave problem*** introduced below. The objective is to find a route ρ minimizing the quantity stated by eq. (7) subject to constraints (8) to (18). So, the *route-generator* problem can be formulated as follows:

$$Min \quad \left[CV - \left(\sum_{i \in I} \pi_c^i Y_i + \sum_{i \in I^+} \pi_l^i Y_i + \pi_n \right) \right] \tag{7}$$

Cost-based constraints: Eqs. (8) and (9) compute the total cost for the route generated. While eqs. (8) are focused on the start and end of the trip, eqs. (9) accumulate the cost of the intermediate trips along the route. So, if nodes i and j are allocated to the route ($Y_i = Y_j = 1$), the ordering of both nodes is determined by the value of the sequencing variable S_{ij}. If location i is visited before j ($S_{ij} = 1$), according constraints (9.a), the travel cost up to location j (C_j) must be larger than C_i by at least c_{ij}. In case node j is visited earlier, ($S_{ij} = 0$), the reverse statement holds and constraint (9.b) becomes active. If one or both nodes are not allocated to the tour, eqs. (9.a)-(9.b) become redundant. M_C is an upper bound for variables C_i and CV.

$$\left\{ \begin{array}{c} C_i \geq c_{bi} \\ CV \geq C_i + c_{ib} - M_C(1 - Y_i) \end{array} \right\} \qquad \forall i \in I^+ \cup I^- \qquad \begin{array}{l} (8.a) \\ (8.b) \end{array}$$

$$\left\{ \begin{array}{l} C_j \geq C_i + c_{ij} - M_C(1 - S_{ij}) - M_C(2 - Y_i - Y_j) \\ C_i \geq C_j + c_{ij} - M_C S_{ij} - M_C(2 - Y_i - Y_j) \end{array} \right\} \qquad \forall i, j \in I^+ \cup I^- : i < j \qquad \begin{array}{l} (9.a) \\ (9.b) \end{array}$$

Time-based constraints: Eqs. (10) and (11) define visiting-time constraints that are similar to eqs. (8) and (9) but apply to the time dimension. M_T is an upper bound for variables T_i and TV. Eq. (12) forces the service time on any node $i \in (I^- \cup I^+)$ to start at a time T_i bounded by the interval $[a_i, b_i]$. Also, eq. (13) defines that the routing time TV must be lower than t^{max}.

$$\left\{ \begin{array}{c} T_i \geq t_{bi} \\ TV \geq T_i + st_i + t_{ib} - M_T(1 - Y_i) \end{array} \right\} \qquad \forall i \in I^+ \cup I^- \qquad \begin{array}{l} (10.a) \\ (10.b) \end{array}$$

$$\left\{ \begin{array}{l} T_j \geq T_i + st_i + t_{ij} - M_T(1 - S_{ij}) - M_T(2 - Y_i - Y_j) \\ T_i \geq T_j + st_j + t_{ji} - M_T S_{ij} - M_T(2 - Y_i - Y_j) \end{array} \right\} \qquad \forall i, j \in I^+ \cup I^- : i < j \qquad \begin{array}{l} (11.a) \\ (11.b) \end{array}$$

$$a_i \leq T_i \leq b_i \quad \forall i \in I \tag{12}$$

$$TV \leq t^{max} \tag{13}$$

Cargo constraints: Eq. (14.a) states that the cargo of product p (α_{ip}) to be picked-up from any customer $i \in I^+$ must be smaller than the inventory l_{ip}, while the cargo β_{ip} to be delivered to this client must be zero. Conversely, the cargo β_{ip} to be delivered to any customer $i \in I^-$ must be l_{ip} while the cargo α_{ip} to be picked-up from this customer must be zero (eq. 14.b).

$$\left\{ \begin{array}{c} \alpha_{ip} \leq Y_i l_{ip} \\ \beta_{ip} = 0 \end{array} \right\} \quad \forall i \in I^+, p \in P \tag{14.a}$$

$$\left\{ \begin{array}{c} \alpha_{ip} = 0 \\ \beta_{ip} = Y_i l_{ip} \end{array} \right\} \quad \forall i \in I^-, p \in P \tag{14.b}$$

The cargo of product p carried by the visiting vehicle up to the node $i \in (I^+ \cup I^-)$ is computed as the difference (L_i - U_i) between variables L_{ip} (total cargo of p loaded by the visiting vehicle up to the node i) and U_{ip} (total cargo of p unloaded by the visiting vehicle up to the node i). So, eq. (15) states that the current load of p up to the node $i \in (I^- \cup I^+)$, must be larger than zero. In turn, eqs. (16) state that the sum of these

differences must be smaller than the volumetric and weight vehicle capacity. Constraints (17) set the accumulated loaded and unloaded cargo of product p in a way similar to eqs. (9). Eq. (18.a) states that the load of product p available after visiting node i (L_i) must be larger than the quantity α_{ip} to be picked-up from the node and smaller than the quantity of goods collected on the generated tour. Constraint (18.b) is similar to (18.a) but for the cargo unloaded after servicing node i (U_i). M_L is an upper bound for L_i and U_i.

$$L_{ip} - U_{ip} \geq 0 \qquad \forall i \in I^+ \cup I^-, p \in P \qquad (15)$$

$$\left\{ \begin{array}{l} \sum_{p \in P} w_p \left(L_{ip} - U_{ip} \right) \leq q_v \\ \sum_{p \in P} v_p \left(L_{ip} - U_{ip} \right) \leq q v_v \end{array} \right\} \qquad \forall i \in I^+ \cup I^- \qquad \begin{array}{l} (16.a) \\ (16.b) \end{array}$$

$$\left\{ \begin{array}{c} L_{jp} \geq L_{ip} + \alpha_{jp} - M_L \left(1 - S_{ij}\right) - M_L \left(2 - Y_i - Y_j\right) \\ U_{jp} \geq U_{ip} + \beta_{jp} - M_L \left(1 - S_{ij}\right) + M_L \left(2 - Y_i - Y_j\right) \\ L_{ip} \geq L_{jp} + \alpha_{ip} - M_L S_{ij} - M_L \left(2 - Y_i - Y_j\right) \\ U_{ip} \geq U_{jp} + \beta_{ip} - M_L S_{ij} + M_L \left(2 - Y_i - Y_j\right) \end{array} \right\} \qquad \forall i, j \in I^+ \cup I^-, p \in P : i < j \qquad \begin{array}{l} (17.a) \\ (17.b) \\ (17.c) \\ (17.d) \end{array}$$

$$\left\{ \begin{array}{l} \alpha_{ip} \leq L_{ip} \leq \sum_{j \in I^+} \alpha_{jp} \\ \beta_{ip} \leq U_{ip} \leq \sum_{j \in I^-} \beta_{jp} \end{array} \right\} \qquad \forall i \in I^+ \cup I^- \qquad \begin{array}{l} (18.a) \\ (18.b) \end{array}$$

The relaxation of the RMP is not necessarily integer and it may exist a column that would price favorably but it is not present in the columns pool. Consequently, in order to find the optimal solution we must generate columns after branching. Ryan and Foster [1981] proposed a rule that is equivalent to branch in node-to-tour assignment relationships. Following this rule, we just branch on variables Y_i. Rather than adding explicitly the branching constraints to the master problem, the infeasible columns are eliminated from the pool and branching constraints are enforced at the slave level. Summarizing, given an initial feasible solution, the series of RMP and slave problems is solved until no new columns can be found. Then, if the solution of the integer RMP (or global upper bound GUB) is higher than the solution of the relaxed RMP (or global lower bound GLB), we must branch on variables Y_i to generate the missing routes. At each tree-node, the mechanism is repeated and the bounds are compared. So, if the local lower bound is higher than the GUB, the node is fathomed; otherwise it is divided into two child-nodes that are included in the database of unsolved subspaces. Afterwards, the next subspace is fetched from the database until the base is empty. In such a case, the algorithm shows the solution and terminates. The column generation procedure is run at all nodes of the branch & bound tree and accelerating tricks as heuristic generation are used whenever it is possible.

4. An illustrative example

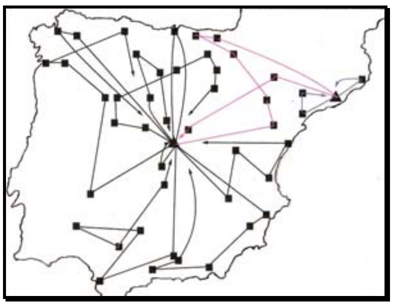
The proposed decomposition approach was developed with ILOG OPL Studio 3.7. A large supply chain distribution problem was solved in a 2 Ghz Pentium IV PC. The example comprises two factories (▲) located in Madrid and Barcelona from where two types of final goods (P1 and P2) are to be delivered to the 45 main Spanish cities (■). In addition, a certain quantity of auto-parts (P3) is delivered to a warehouse located in Bilbao, from where they must be transported to both factories for its posterior use. The factories host a number of identical vehicles with q_v = 15tn and qv_v = 45 m^3 that can perform pick-up/delivery tasks on customer-nodes during specified time-windows. The problem-data is summarized on Table 1. Furthermore, the average trucks travel-speed is

70 km/h, t^{max} = 120 h, service times on pick-up nodes are constants (8 h.) and unload times at delivery nodes are modeled as the sum of a constant (1 h.) and a time proportional to the quantity to deliver ($\Sigma_{p \in P}$ $0.005\ P_i\ h;\ i\ \in \Gamma$). The solution, detailed in Figure 1 provides the detailed routes for all vehicles. The solution, found in just 1316.4 s of CPU time, involves the use of 11 vehicles for fulfilling 12 trips with a total travelled distance of 13007 km.

Table 1. Problem data

Site	Albacete;Alicante;Almeria;Avila;Badajoz;Barcelona;Bilbao;Burgos;Caceres;Cadiz;Castellon;CReal; Cordoba;Coruna;Cuenca;Gerona;Granada;Guadalajara;Huelva;Huesca;Jaen;Leon;Lerida;Logrono;L ugo;Madrid;Malaga;Murcia;Orense;Oviedo;Palencia;Pamplona;Pontevedra;Salamanca;SSebastian;S antander;Segovia;Sevilla;Soria;Tarragona;Teruel;Toledo;Valencia;Valladolid;Vitoria;Zamora;Zarag oza
a_i	32;32;80;8;56;0;60;80;64;56;104;104;80;88;56;16;88;104;80;88;32;88;32;104;80;0;88;64;80;88;104; 56;80;32;64;88;104;88;104;8;32;8;56;64;88;80;32
b_i	40;40;88;24;64;120;68;88;72;64;112;112;88;96;64;24;96;112;88;96;40;96;40;112;88;120;96;72;88;9 6;112;64;88;40;72;96;112;96;112;16;40;16;64;72;96;88;40
$P1_i$	0;-600;-200;-300;-50;5000;-400;-200;-50;0;-300;-100;-200;-1000;-200;-400;-100;-400;-150;-200;0;- 200;-300;-500;-100;5000;-400;0;-300;0;-50;-400;-200;-800;-400;-700;-100;-600;-300;-200;-300;- 1200;-400;-200;-100;0;-700
$P2_i$	-500;0;-250;-100;0;2500;0;0;-50;-100;-200;-200;-200;-800;0;0;-250;0;-50;-200;-150;-700;0;0;- 250;2500;0;-400;-150;-600;-50;-400;-500;0;-200;-700;-50;-400;-100;-700;0;-600;-400;-500;-150;- 300;0
$P3_i$	;0;0;0;0;0;600;1000;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;400;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0

Figure 1. The detailed best solution



Barcelona-Gerona: 201 km/20.4 h
Barcelona-Tarragona-Lérida: 346 km/36.7 h
Madrid-Santander Madrid-Toledo: 930 km/101.6 h
Madrid-Albacete-Cuenca-Valencia-Castellón: 1116 km/113.45 h
Madrid-Lugo-Coruña: 1219 km/106.7 h
Madrid-Valladolid-León-Palencia-Segovia: 709 km/110.7 h
Madrid-Jaén-Cádiz-Córdoba-Huelva-Sevilla-CReal:1830 km/109.2 h
Madrid-Alicante-Murcia-Almería-Málaga- Granada: 1499 km/101.8 h
Madrid-Badajoz-Cáceres-Orense-Pontevedra-Oviedo: 1990 km/105.2 h
Madrid-Teruel-Zaragoza-Pamplona-Bilbao-SSebastian-Barcelona-Huesca-Guadalajara: 2138 km/118.8 h
Madrid-Ávila-Salamanca—Zamora-Burgos-Vitoria-Logroño-Soria: 1029 km/115.3 h

5. Conclusions

This work has introduced an effective optimizing decomposition approach lying on a column generation algorithm that utilizes an MILP formulation as routes generator. The proposed method is able to deal with realistic transportation problems which aim at minimizing distribution costs while keeping a high level of efficiency in complex three-tier networks comprising factories, warehouses and customers. The applicability of the approach has been proven by solving a large-scale example comprising two factories, one warehouse and forty five customers with very low computational effort.

References

Méndez C., R Dondo and J.Cerdá. Proceedings FOCAPO-2008 (2008).
Desrochers M., J. Desrosiers and M. Solomon. Op. Res. 40, 642-354 (1992)
Simchi-Levi, D., Kaminsky, P., Simchi-Levi, E., Man. the S.Chain, McGraw-Hill (2004).
Ryan D. and B. Foster. Wren, A., (Ed.), Comp. Sch. of Public Transp. Urban Passenger. Vehic. and Crew Sch., North-Holland, Amsterdam, pp. 269–280 (1981).

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

A novel approach to policy design using process design principles

Araz Taeihagh, René Bañares-Alcántara and Zun Wang

Department of Engineering Science, Oxford University, Parks Road, Oxford OX1 3PJ, UK

Abstract

Policy design, as it is currently practiced, relies on the manual formulation of policies. This fact, combined with an increase in the complexity of the systems under study, results in a large portion of the possible choices for action being left unexplored. The introduction of a systematic approach for exploring alternative policies using a computational methodology has the potential to help in the decomposition of complex problems into subproblems with more manageable size, and can accelerate and improve the effectiveness of the policy-making process. A new approach to facilitate policy design is proposed; it uses principles widely used in chemical engineering design combined with mathematical and artificial intelligence techniques. In particular, our ideas are based on the Hierarchical Design Method for the conceptual design of chemical processes (Douglas, 1988) and use of techniques such as network analysis and agent-based modelling. A six-step framework has been proposed and is being implemented as a prototype decision support system (DSS). An agent-based modelling approach, due to the opportunistic nature of agents is used as the programming paradigm in the design of policies. The DSS creates a network of policy measures and extracts information from the network during the development of policy alternatives. A set of agents then combine policy measures and goals within a tree structure at different hierarchical levels during the creation of the overall policy. The results from the research are a fresh contribution to the methodological development of policies using a multidisciplinary approach and have the potential to accelerate the design of policies and improve their success rate.

Keywords: Decision support systems, process design, policy design, transport emission reduction.

1. Introduction

As our socio-economical systems are increasingly becoming more complex, solving major national and international problems are becoming more challenging. Such complexities arise from increased number of stakeholders and the consequences of phenomena such as globalization, result in energy and environmental constrains (both intentional, e.g. health and safety standards, and unintentional, e.g. global warming). Recent events such as global warming, and the credit crunch highlight the increased complexities in our systems and the interconnections between different geographical locations and entities. For instance, in the case of global warming, as the effect of man-made emissions on climate has been acknowledged, tackling the climate change issue has become a serious priority for governments and international organizations. Although the problem has been recognized, an understanding of how to develop effective, acceptable and detailed policies has yet to be attained. There is a wide debate on the level at which the emission reduction targets should be set, but once values are

agreed, there are a number of possible alternative strategies (policies) to achieve them. Addressing such a complex problem requires the formulation of integrated policies that are well coordinated and reciprocally reinforced. The precise nature and scope of policies designed to achieve their respective targets are necessarily geographically and culturally dependent given the variability of resources, of access to technology and of political constraints at different locations and times. For this reason, a one-for-all and static policy is unlikely to achieve the desired targets and there is a need for bespoke policies that are able to accommodate periodic revisions. The introduction of a systematic approach for exploring alternative policies using a computational methodology will help in decomposing such complex problems into subproblems with more manageable size, accelerate and improve the effectiveness of the process of policy-making. The transport sector has been chosen for the case study because it is the second largest growing source of greenhouse emissions (IPCC, 2007). A new framework for policy formulation is being developed and implemented as a prototype DSS using the backcasting approach around a case study: the formulation and analysis of the policies required to achieve CO_2 emission targets for the UK transport sector.
The background to policy design and backcasting methodology is discussed in Section 2. Section 3 describes the proposed framework for policy design and the details of the DSS. The results achieved in the development of the system are presented in Section 4 followed by a description of future work and the conclusions in Section 5.

2. Background to policy design and backcasting

2.1. Policy Design

A policy is a principle or guideline for action in a specific everyday-world context (Pohl, 2008) and policy design is the step in which the components of a policy are selected and the overall policy is formulated. Policies may be related to technological, economic, political and social aspects. Some technological and economic factors can be modelled mathematically, resulting in complex simulation/prediction models. Mathematical models provide valuable insights, but are only part of the required inputs to a general policy-making process because, after all, decisions about desirable futures, and the policies to attain them, are essentially a question of social values and political choice (Robinson et al., 2006).
The focus within the transport domain has been on the development of mathematical models and tools for assessment of large-scale infrastructure projects and analysis of transport policies. Monetary-based techniques and Multi-Criteria Analysis techniques (especially Multi-Criteria Decision Analysis techniques (MCDA)) form the basis of the models and tools. Often, risk analysis techniques and probabilistic models are used to further refine and fine-tune the models. Both in the case of policy design in general and in the transport sector specifically, the decisions on what to include in the policies (their *synthesis*) is done manually. This fact, together with the size of the space of possible policies, results in a large portion of the design space left unexplored. A systematic approach to explore the space of alternative policies using a computational methodology will accelerate the task of policy-making and improve policy effectiveness.

2.1.1. Process design and policy design

Similar to the transport domain, there are powerful simulators used for chemical process design (Finlayson, 2006). However, in contrast to the transport sector, process engineers have developed a set of tried-and-tested synthesis methods, which help in the selection of the items of equipment and how they are to be interconnected (Westerberg, 2004). In

process design, synthesis and simulation steps are applied in tandem and iteratively: a synthesis step generates alternative designs and the output from a simulation step is used to compare those alternatives and inform the application of the next synthesis step (see Figure 1(a)). The introduction of the aforementioned techniques changed the way in which chemical processes were designed. The plan is to adapt/adopt existing process synthesis and design methods to the design of policies, i.e. to provide practical frameworks and tools for better policy design and support decision-making regarding which measures to use and how to combine them to achieve the targets (the set of tasks included inside the dashed box in figure 1(b)).

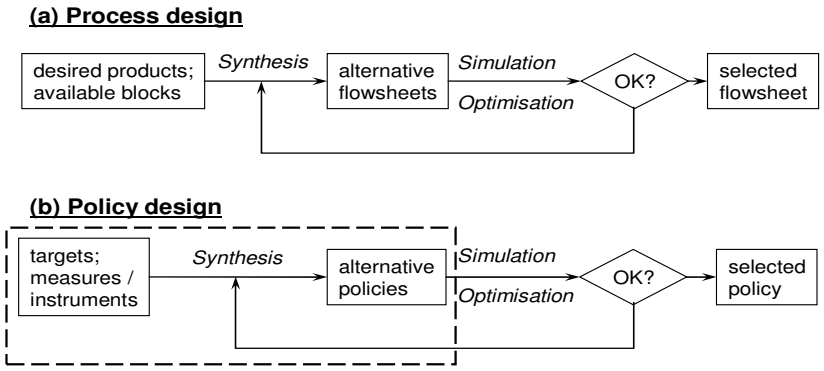


Figure 1 Analogy between (a) process design and (b) policy design

The exploration-based model of design (Smithers & Troxell, 1990) and the use of a hierarchical organization are reflected in the widely accepted Hierarchical Design Method (Douglas, 1988), and have been embodied in different design support systems. Our hypothesis is that the process by which the collection of operators is synthesized is similar. However, process and policy design are not identical and, as a result, different approaches will have to be used to take into account their differences, in particular, the pervasiveness of non-quantifiable factors in policy-making.

2.2. Backcasting

Backcasting was first proposed in (Robinson, 1990) and has become a well-established methodology. It involves the development of normative scenarios aimed at achieving desired end-points, i.e. working backwards from a desirable state to determine what policies would be required to reach it. The Visioning and backcasting for UK transport policy project (VIBAT) studied the potential for 60% reduction of emissions due to transport by the year 2030 using a backcasting approach (Banister & Hickman, 2006).

3. A new framework for design of policies

The purpose of the development of the proposed framework and accompanying DSS is to facilitate policy design to achieve environmental targets. The designed system should be applicable to different targets in different sectors and geographical scopes.

3.1. The proposed policy formulation framework is broken down into six steps:

1- Identification of the relevant concepts such as targets and goals.
2- Development of a library of policy measures.
3- Specification of relations among policy measures.

4- Generation of policy packages.

5- Application of planning techniques in the scheduling of policy packages.

6- Evaluation and comparison of alternative policy clusters

3.2. Software architecture and implementation details of the DSS

Figure 2 illustrates the software architecture of the DSS. The connection between the database and the core is achieved through the Java Database Connectivity application programming interface (JDBC, 2008). The Java code currently uses Mathematica's kernel as a computational and visualization engine (Mathematica, 2008). In addition, Mathematica is used for accessing the discrete mathematics package Combinatorica (Pemmaraju & Skiena, 2003). The integration between Java and Mathematica is established using J/Link (J/Link, 2008). The information acquired through user input and analysis via Mathematica is channeled towards the agent-based modeling toolkit. An agent-based model is a system for the simulation of interactions between autonomous entities. Due to the opportunistic nature of agents, it is possible to combine the top-down and bottom-up approaches of design to better address complex problems such as the one under study, allowing for an enhanced utilization of the available information for more effective policy design. Moreover, an agent-based modeling approach is well suited for developing flexible systems that can handle incremental discovery, design and development (North & Macal, 2007) and provides a natural method for decomposing the overall problem.

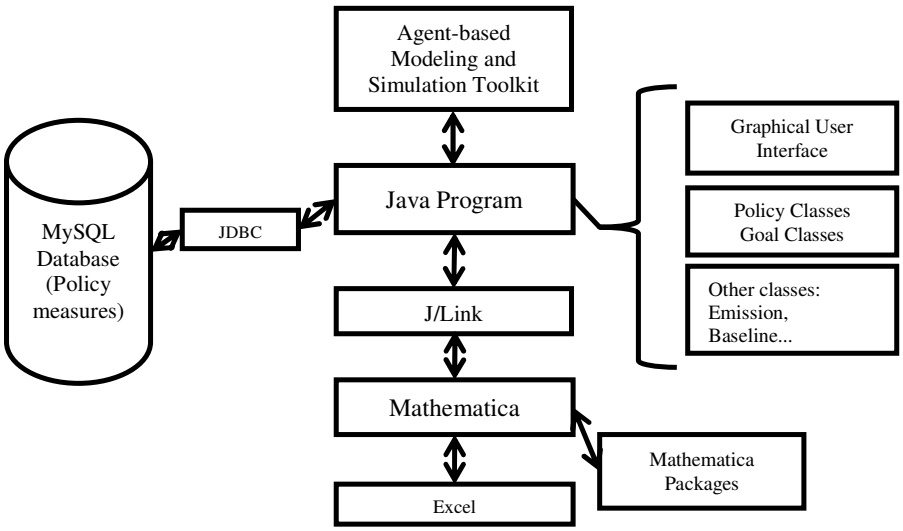


Figure 2 Software architecture of the decision support system

4. Results

4.1. Generation and analysis of the policy measures network

The 122 policy measures from the VIBAT project constitute the core of the library of policy measures with properties such as the policy type, effectiveness and timeframe of implementation. Based on specified preconditions among policy measures, a graph structure was formed with policy measures as nodes (vertices) and the relations as directed edges between the nodes, and several analyses were performed on this network.

Network characteristics, such as degree, betweenness, closeness, and eigenvector centralities were examined. Based on these characteristics, a set of preference indices were created such as *total cost vs. effectiveness* (includes the cost of preconditions), *node cost vs. effectiveness*, *implementation time vs. effectiveness*, and *combined weighted costs and time vs. effectiveness*, which are used later by the agents during the selection and comparison of different policy measures. In addition, graph colouring was used to identify the minimum number of policy packages required for the implementation of all the selected policy measures, also a number of contradictions were defined among policy measures.

4.2. Generation and analysis of the tree of policy alternatives

The next step in the development of the DSS is the generation of desirable combinations of policy measures. A tree structure is utilized for exploring the space of alternatives by generating different policies using the policy measures as building blocks. A multi-layered agent-based architecture is applied, where agents select policy measures as building blocks for creating policy packages, policy clusters and the overall policies. Results from the agent working on creating policy packages are fed to another agent working on policy clusters. This last agent, in turn, creates a policy cluster from the pool of available packages based on its internal rules and interests. In the same manner, results from the policy clusters are fed to the agent which was in charge of creating the overall policies.

The concept of hybrid tree is used to combine goals and policy measures in the same structure. Previously the goals defined by the user were connected to the policy measures through hardwired relations. By formalising the goals, it has been made possible for the agents to connect policy measures to the relevant goals. The goals are formally defined using a method similar to the one proposed by Hashmi et al. (Hashmi, 2003), where each goal has a number of properties such as emission reduction target, a desired time frame for implementation, a level of priority, and a context for the actions. An example goal is to reduce the average level of emissions of the vehicle fleet from 189 g/km in the year 2000 to 90 g/km by the year 2030 with a high priority, which would result in a reduction of 11.8 million tonne of Carbon (MtC) emissions. A set of agents then combine policy measures and goals within a tree structure at different hierarchical levels during the creation of the overall policy.

5. Future Work and Conclusions

The generation and evaluation procedures for the tree structure will be enhanced in the future, and the results from the system will be validated. As it is not possible to compare the results with a real system, since it would require the implementation of the policies beforehand, we intend to compare the results with the VIBAT study and seek transport policy experts' opinion.

It is intended to increase the number of agents that participate in the creation of the tree structure and to use negotiation between agents. Furthermore, general directives such as the DTLR multi-criteria analysis manual (Dodgson, 2001), and tentative guidelines on choosing appropriate multi-criteria decision analysis techniques (Guitouni & Martel, 1998), will be utilized to improve the evaluation of policy packages and clusters.

In the area of network analysis, the effect of the type of relations between nodes with respect to their immediate and non-immediate neighbors will be explored.

The purpose of the research is to facilitate the design of policies by using knowledge gained mainly from process design and synthesis. The focus has been directed towards the similarities between process and policy design with the specific aim of introducing a

new framework and systematic thinking to the problem of policy formulation. A working prototype DSS has been developed; it facilitates the design of transport policies that aim to achieve environmental targets. In particular, the software will help decision makers in selecting appropriate policy measures to achieve a reduction in transport related CO_2 emissions. This research constitutes the first step towards the development of a general family of computer-based systems that support the design of policies to achieve environmental targets for areas such as transport, energy, biofuels/food security, water, etc. The results from the research will be a fresh contribution to the methodological development of policies and have the potential to: (a) accelerate the design of policies. (b) improve the chance of policy success through exploration of more alternatives. (c) facilitate the specialisation of policies.

6. Acknowledgments

The authors would like to acknowledge the support of Prof. David Banister at the Transport Studies Unit of the University of Oxford and Dr. Robin Hickman from Halcrow Group Ltd for providing additional details regarding the VIBAT project.

References

D. Banister, & R. Hickman, 2006, Visioning and backcasting for UK transport policy (VIBAT), the Bartlett school of planning & Halcrow group ltd, http://www.ucl.ac.uk/~ucft696/vibat2.html

J. Dodgson, M. Spackman, A. Pearman, L. Phillips, 2001, DTLR multi-criteria analysis manual, National Economic Research Associates, UK, Retrieved 1/18/2008 www.communities.gov.uk/documents/corporate/pdf/146868.pdf

J. M. Douglas, 1988, Conceptual design of chemical processes. McGraw-Hill New York.

B. A. Finlayson, 2006, Process simulation, Chapter 7 in: Introduction to chemical engineering computing. Wiley-Interscience.

A. Guitouni, & J. M. Martel, 1998, Tentative guidelines to help choosing an appropriate MCDA method. European Journal of Operational Research, 109(2), pp. 501-521.

IPCC. 2007, Climate Change 2007: Mitigation. Contribution of the Working group III to the 4th Assessment Report of the Intergovernmental Panel on Climate Change, B. Metz, O.R. Davidson, P.R. Bosch, R. Dave and L.A. Meyer (eds.). Cambridge University Press.

J/Link, 2008, Java toolkit: J/Link: Integrating Mathematica and java. Retrieved 7/8/2008, http://www.wolfram.com/solutions/mathlink/jlink/

JDBC, 2008, The java database connectivity (JDBC) overview, Sun Microsystems inc. Retrieved 7/8/2008 http://java.sun.com/products/jdbc/overview.html

N. Hashmi, A. Boxwala, D. Zaccagnini, J. Fox, 2003, Formal Representation of Medical Goals for Medical Guidelines, Medinfo, pp. 1663.

Mathematica. 2008, Wolfram Mathematica: Home page. Retrieved 7/8/2008 http://www.wolfram.com/products/mathematica/index.html

M.J. North, & C.M. Macal, 2007, Managing business complexity: Discovering strategic solutions with agent-based modeling and simulation. Oxford University Press, New York.

S. V. Pemmaraju, & S. S. Skiena, 2003, Computational discrete mathematics: Combinatorics and graph theory with Mathematica. Cambridge University Press.

C. Pohl, 2008, From science to policy through transdisciplinary research. Environmental Science and Policy, 11(1), 46-53.

J. Robinson, 1990, Futures under glass: A recipe for people who hate to predict. Futures, 22(8), 820.

J. Robinson, M. Bradley, P. Busby, D. Connor, A. Murray, B. Sampson, et al., 2006, Climate change and sustainable development: Realizing the opportunity. Ambio, 35(1), 2-8.

T. Smithers, & W. Troxell, 1990, Design is intelligent behaviour, but what's the formalism? AI EDAM, 4(2), 89-98.

A. W. Westerberg, 2004, A retrospective on design and process synthesis. Computers and Chemical Engineering, 28(4), 447-458.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Continuous-time Representation Approach to Hybrid Process Scheduling of Single-product Production

Lijie Su, Lixin Tang

Liaoning Key Laboratory of Manufacturing System and Logistics, The Logistics Institute, Northeastern University, Shenyang, 110004, China

Abstract

This paper investigates the scheduling problem of single-product production on hybrid batch and continuous production processes where the batch process is composed of several parallel reactors and followed by a continuous production line. There is an intermediate storage tank with limited capacity available between two processes. The parallel reactors may simultaneously supply material into the tank. Based on global-event continuous-time modeling method, the scheduling is formulated as a mixed integer nonlinear programming (MINLP) which can be solved by a conventional solver. Two real hybrid processes with certain demand ranges are examined to illustrate the efficiency and applicability of the presented mathematical model.

Keywords: scheduling, hybrid process, global event, continuous-time representation

1. Introduction

The research area of process scheduling has received great attention in the recent 30 years. The complicated process topology structure and variable characters of products make process scheduling different from classical scheduling on discrete manufacture, and mathematical programming approach is often adopted. There are some prominent theoretical modeling fruits in this research field (Méndez, Cerdá, et al 2006). It would be valuable to extend these obtained modeling fruits to more practical production. However, there exists a certain gap between theory methods and practical application for the limitation of methods and the diversity of practical process. For instance, most research work focus on batch process or continuous process, there exist the hybrid process in practice. Hybrid process with batch and continuous modes is one kind of typical processes, which is popular in polymerizing process. Shah, Liberis, et al 1996 presented one integrated approach to design the batch plants. Kang, Kang, et al 2006 formulated the integrated scheduling for the Polyvinyl Chloride (PVC) process. These two references all introduced discrete-time modeling method, and the scheduling was formulated as a mixed integer linear programming (MILP) model where batch processing time is assumed fixed.

Taking PVC process as a research background, we investigate the scheduling of the hybrid process. The new feature of the problem is that scheduling with batching decision is considered and batch processing time is varied. The failure of the discrete-time modeling to tackle the variable batch processing time motivates us to develop the continuous-time method to model the problem under our consideration. Here, global time point is adopted for all units to represent event.

This work is partially supported by NSF Grant # 60674084.

2. Problem Description

The hybrid process here is composed of three parts sequentially: parallel batch reactors, one storage tank and one continuous production line. Multiple tanks and continuous production lines could be regarded as one for the single-product production. And continuous production line could be reduced into one continuous processing unit. Batch size could vary in certain range, and batch processing time is linear function of batch size. The simultaneously unloading operations from multiple reactors to tank are allowed. Once completing batch operation, semi-finished product must transfer into tank at once in order to assure the quality of product. The capacity of tank is limited. The production rate of continuous production line could be adjusted in some range. And the continuous production line would be expected to be running smoothly.

The problem description for the single-product scheduling on one hybrid process is as follow: given the upper and lower bound of the batch size, the variable batch processing time dependent on batch size, the capacity of intermediate storage tank, the production-rate range of continuous production line, total demand quantity of one order. Determine the detail scheduling of the whole process in order to minimize the *makespan*. One assumption is that the delivery durations from reactors to storage tank and storage tank to continuous production line are ignored.

3. Mathematical Formulation for Scheduling Hybrid Process

Concerning with simultaneous multiple-to-one operations from batch units to finite storage tank, global-event continuous-time representation approach is adopted here, which predefines certain number of time points for all units including tank, and any task should happen at some time point (Zhang&Sargent, 1996). Here, the tasks are defined as the finishing operation of batch units and adjusting one of production rate. Compared with unit-specific event based approach (Ierapetritou&Floudas, 1998), global-event one would be more simpler on expression of material balance constraints, and easy to settle storage.

Nomenclature

Indices:

$j = 1,\ldots,J$: batch reactors.

$n = 1,\ldots,N$: event/time points.

Parameters:

D : demand quantity.

$B_j^{\min}$, $B_j^{\max}$: minimal and maximal capacity of batch reactor *j*.

α_j, β_j : constant and variable processing time of batch reactor *j*.

$S^{\max}$: maximal capacity of intermediate storage tank.

$S^{\min}$: minimal required left intermediate product in tank during production process.

$R^{\min}$, $R^{\max}$: adjusted production-rate range of continuous process.

Variables:

x_{jn} : binary variable, 1 if reactor *j* completes one batch on time point *n*, otherwise 0.

y_n : binary variable, 1 if the continuous production process starts working on time point *n*, otherwise 0.

B_{jn}: continuous variable, batch size which is completed in reactor *j* on time point *n*.

P_{jn}: continuous variable, batch processing time of reactor *j* completed just on time point *n*.

T_n: continuous variable, the time of the *nth* time point on all units.

SB_n: continuous variable, left intermediate storage quantity at the *nth* time point before receiving new material from reactors.

S_n: continuous variable, intermediate storage quantity at the *nth* time point after receiving intermediate product from some reactors.

R_n: continuous variable, production rate of continuous production process between the *nth* and *n+1th* time point.

MK: continuous variable, the final finishing time of continuous production line, namely, the *makespan*.

The mathematical formulation for the single-product scheduling on the hybrid process is presented as follow.

For the parallel batch reactors, there are allocation, capacity, processing time and time sequence constraints.

There are at most |*J*| batches completed at any defined time point.

$$\sum_{j=1}^{J} x_{jn} \le |J| \qquad \forall n = 1...N \tag{1.1}$$

With the balance of work load, the batch number for every reactor should satisfy some limitation.

$$\sum_{n=1}^{N} x_{jn} \le \left\lceil D \Big/ \left(B_j^{\min} \cdot |J| \right) \right\rceil \qquad \forall j = 1...J \tag{1.2}$$

The processed batch amount is bounded by the maximum and minimum capacities of that batch reactor according to reactor capacity and operation safety.

$$x_{jn} \cdot B_j^{\min} \le B_{jn} \le x_{jn} \cdot B_j^{\max} \qquad \forall j = 1...J, n = 1...N \tag{1.3}$$

The processing time of one batch is composed by fixed part and variable one, and the later part is dependent on the batch size.

$$P_{jn} = \alpha_j \cdot x_{jn} + \beta_j \cdot B_{jn} \qquad \forall j = 1...J, n = 1...N \tag{1.4}$$

The time of the *nth* time point is less than that of the *n+1th* time point.

$$T_n \le T_{n+1} \qquad \forall n = 1...N-1 \tag{1.5}$$

For any reactor *j*, the time of time point *n* must be larger than the summation of all batches processing time just before time point *n*.

$$\sum_{m=1}^{m=n} P_{jm} \le T_n \qquad \forall j = 1...J, n = 1...N \tag{1.6}$$

For the storage tank, there are material balance and capacity constraints.

The left intermediate-product quantity on time point *n* equals to the intermediate storage quantity on *n-1* minus the quantity processing by continuous production line.

$$SB_n = S_{n-1} - R_{n-1} \cdot \left(T_n - T_{n-1} \right) \qquad \forall n = 2...N \tag{1.7}$$

The intermediate storage quantity on time point n is the left intermediate-product quantity on time point n plus completed batch quantity from reactor at the same time point.

$$S_n = SB_n + \sum_{j=1}^{J} B_{jn} \qquad \forall n = 2...N \tag{1.8}$$

The reason for using two variables defining the storage quantity on time point n is that consumption of continuous production has been lasted between the *n-1th* and *nth* time points, however, production of batch reactor happens at the *nth* time point. Two material balance equations would constraint the consumed quantity was not the just produced one at time point n. For the first time point, we use the following constraint

$$S_1 = SB_1 = \sum_{j=1}^{J} B_{j1} \tag{1.9}$$

It is the limit on maximal intermediate storage quantity on any time point.

$$S_n \le S^{\max} \qquad \forall n = 1...N \tag{1.10}$$

In order to ensure the continuity of next procedure, some minimal quantity of intermediate product must be remained in storage tank.

$$SB_n \ge S^{\min} \qquad \forall n = 1...N \tag{1.11}$$

This constraint would make the gotten schedule more flexible, for the minimal quantity of intermediate product could be regarded as a buffer between batch reactors and continuous production line.

For the continuous process part, there are processing rate, smoothing production and *makespan* constraint.

This is the adjusted range of production rate when continuous process runs smoothly.

$$y_n \cdot R^{\min} \le R_n \le y_n \cdot R^{\max} \qquad \forall n = 1...N \tag{1.12}$$

The followed constraint would assure the uninterrupted production for continuous production line, which means once the continuous production process starts up, it would continue until final completion.

$$y_n \ge y_{n-1} \qquad \forall n = 2...N \tag{1.13}$$

The completion time of one order with certain demand, the *makespan,* is greater than the summation of the time of last time point N plus the processing time of quantity S_N.

$$MK \ge T_N + S_N / R^{\max} \tag{1.14}$$

The total production quantity must satisfy demand. Here, D means the quantity of intermediate product converted from one of final product.

$$\sum_{j=1}^{J} \sum_{n=1}^{N} B_{jn} \ge D \tag{1.15}$$

The objective function is minimization of makespan, that is

$$\min \quad MK \tag{1.16}$$

Proposition 1. For one certain demand quantity, the upper bound of batch number is $\lceil D/B^{\min} \rceil$, and the lower bound of it is $\lceil D/B^{\max} \rceil$.

Here, $B^{\min} = \min_j \left\{ B_j^{\min} \right\}$, $B^{\max} = \max_j \left\{ B_j^{\max} \right\}$.

According to proposition 1, the valid cut inequalities would be presented.

$$\sum_{j=1}^{J}\sum_{n=1}^{N} x_{jn} \le \left\lceil \frac{D}{B^{\min}} \right\rceil \tag{1.17}$$

$$\sum_{j=1}^{J}\sum_{n=1}^{N} x_{jn} \ge \left\lceil \frac{D}{B^{\max}} \right\rceil \tag{1.18}$$

Proposition 2. The lower bound for the problem *makespan* is

$$MK^{LB} = \min_{j}\left\{\alpha_j + \beta_j \cdot B_j^{\min}\right\} + D / R^{\max}. \tag{1.19}$$

The shortest duration for completing demand is $D/R^{\max}$, if continuous production line starts at the beginning of horizon. Concerning batch processing time, the earliest time getting intermediate product is $\min_{j}\left\{\alpha_j + \beta_j \cdot B_j^{\min}\right\}$, as the condition for the beginning of continuous production line. Therefore, the idea objective for the whole production process is the summation of the two parts.

The lower bound valid inequality is

$$MK \ge MK^{LB} \tag{1.20}$$

4. Case Studies

All the following data from process structure to processing time is gotten from real PVC production processes. The number of parallel units is 4 and 8, which operational capacity varies from 35m^3 to 39m^3. The batch processing time is calculated by $\alpha_j = 22, \beta_j = 0.05 \quad \forall j = 1,\cdots,J$. Other data would be within each case.

LINGO 8.0 is used to solve the MINLP formulation, and the selected solver is Branch and Bound. In order to accelerate convergence, higher level of heuristics and probing strategies are set. The used computer is Intel Core 2 Due CPU, and the CPU is 2.33 GHz. And the operating system is Windows XP professional 2002.

Proposition 3. The least number of predefined time points is $\left\lceil D/\left(S^{\max} - S^{\min}\right)\right\rceil$.

From the least number of predefined time points, iterative procedures would be applied until the value of the optimal solution not decreases, or decreases little.

The hybrid process I includes 4 reactors in batch part. The capacity of storage tank is 80 ton, the minimal storage quantity of intermediate product is 5 ton, and the production rate varies from 3.6 t/h to7.2 t/h. The computational results are listed in Table 1.

Table 1 Statistic results for Hybrid process I

Demand	Events	Int vars	Cont vars	Lower bound	Solution	CPU s
300	5	25	143	65.4	68.4	3
350	6	30	170	72.4	81.7	38
400	7	35	197	79.3	86.5	19
500	9	45	251	93.2	105.4	113
600	10	50	278	107.1	115.9	103
800	11	55	305	134.9	152.9	112
900	14	70	386	148.8	163.3	104
1000	16	80	440	162.6	181.5	117
1500	23	115	629	232.1	258.9	110
2000	28	140	764	301.5	343.9	129

The hybrid process II includes 8 reactors in batch part. The capacity of storage tank is 160 ton, the minimal storage quantity of intermediate product is 4 ton, and the production rate varies from 7.2 t/h to 14.4 t/h. The demand changes from 500 ton to 3000 ton. The computational results are listed in Table 2.

Table 2 Statistic results for Hybrid process II

Demand	Events	Int vars	Cont vars	Lower bound	Solution	CPU s
500	5	45	227	58.5	61.0	55
800	7	63	313	79.3	84.2	180
1000	7	63	313	93.2	103.1	114
1200	8	72	356	107.1	114.2	131
1500	11	99	485	127.9	137.8	114
1800	16	144	700	148.8	164.4	115
2000	16	144	700	162.6	177.5	109
2200	16	144	700	176.5	196.2	105
2500	20	180	872	197.4	221.8	177
3000	24	216	1044	232.1	258.8	170

Here, the solutions in table 1-2 are local optimums gotten by LINGO.

Observing the statistic results of numerical experiments, we find that the proper number of time points usually is near to the least number of predefined time points. The objective function reduces little with the increase of time points, the changes are negligible and multiple local optimal solutions are supplied, which would be flexible for the production management. Comparing with lower bounds, the local optimal solutions are good enough. And the CPU running time is sustainable.

5. Conclusion

This work introduced the global-event based formulation to the hybrid process scheduling. The formulation is applied to the hybrid process of flexible batch size and batch processing time, simultaneous multiple-to-one operations, finite storage, adjusted production rate. The numerical experiments validate the efficiency and practicability of the formulation, which could be extended into the similar practical production process with little modifications. The presented model could also be used to optimize the capacity of storage and maximal constant production rate for certain demand within given time horizon.

References

M. G. Ierapetritou and C. A. Floudas, 1998, Effective Continuous-Time Formulation for Short-Term Scheduling. 1. Multipurpose Batch Processes, Industrial and Engineering Chemistry Research, Vol. 37, 4341-4359.

M. G. Kang, S. Kang, and S. W. Park, 2006, Integrated Scheduling for Polyvinyl Chloride Processes, Industrial and Engineering Chemistry Research, Vol. 45, 5729-5737.

C. A. Méndez, J. Cerdá, I. E. Grossmann, I. Harjunkoski, M. Fahl, 2006, State-of-the-art Review of Optimization Methods for Short-term Scheduling of Batch Processes, Computers and Chemical Engineering, Vol. 30, 913–946.

N. Shah, L. Liberis, E. Izumoto, R. Henson, 1996, Integrated Batch Plant Design: A Polymer Plant Case Study, Computers and Chemical Engineering. Vol. 20, S1233-S1238.

X. Zhang, R. W. H. Sargent, 1996, The Optimal Operation of Mixed Production Facilities-A General Formulation and Some Solution Approaches for the Solution, Computers and Chemical Engineering, Vol. 20, 897-904.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Evaluation of Synergy Effect in the Merger of Companies in a Petrochemical Complex

Sung-Geun Yoon[a], Sunwon Park[a], Jeongseok Lee[b], Peter M. Verderame[c], and Christodoulos A. Floudas[c]

[a]*Department of Chemical and Biomolecular Engineering, KAIST, 373-1, Guseong-dong, Yuseong-gu, Daejeon, 305701, Korea*
[b]*Corporate R&D, LG Chem Ltd., 104-1, Moonji-dong, Yuseong-gu, Daejeon, 305380, Korea*
[c]*Department of Chemical Engineering, Princeton University, Princeton, New Jersey, 08544, U.S.A.*

Abstract

Mergers and Acquisitions (M&A) have been actively carried out in the petrochemical industry. However, the synergy created by the merger of petrochemical companies has seen relatively little study, despite being the primary goal of a merger. This study deals with the horizontal merger of petrochemical companies located within a single complex. Synergies considered in this paper stem from integration of the process network and the utility plant, fixed cost reduction, and contracts in purchasing and selling. A novel mathematical model that represents the operation of a process network and a utility plant and the decision for purchasing and selling contracts is formulated. Four contracts for purchasing and selling are considered. The proposed model is applied to three Korean Naphtha Cracking Center (NCC) companies located in the same industrial complex. The results show that synergy effects from integration of the process network and the utility system, fixed cost reduction, and increased market share together increase profit by fifty percent.

Keywords: Mergers and acquisitions, synergy, petrochemical company

1. Introduction

Recently, the Korean petrochemical industry has experienced prosperity due to rapid growth of the neighbouring Chinese market. That China and Middle Eastern countries are rapidly developing their own petrochemical industries, however, poses a major threat to the Korean petrochemical industry. In order to sustain competitiveness, the Korean petrochemical industry has recently come to consider M&A as an important strategy. At present, too many companies are participating in the petrochemical industry. Companies and government should therefore consider M&A as a necessary strategy to improve competitiveness.

However, there is no blueprint for M&A strategies from a holistic view. Some companies have announced mergers of their subsidiary companies. The government also has not provided direction regarding M&A. Many stakeholders are now debating on the M&A issue. The debate has focused on how much M&As will impact the petrochemical industry and the national economy. In order to assist stake holders, in an earlier study the present authors proposed an optimization model to quantify synergy in

the merger of petrochemical companies in an industrial complex, considering purchasing and selling advantages (Yoon et al., 2008). This paper deals with the same target companies that were considered in our previous work. We propose a new and improved optimization model to reflect more realities. The model can quantify synergy from the case in which a target company has one utility plant respectively. Investment cost to enable transport of chemicals and steam between individual companies is added to the optimization model. The single period model is extended to a multi-period model. And four representative contracts for purchasing and selling are applied to the model.

2. Problem Definition

This paper deals with three companies, A, B, and C, having a naphtha cracking center (NCC) within a complex in Korea. Each company has NCC and BTX processes and different downstream processes. Each company has one utility plant having one boiler, one SS steam turbine, and one HP steam turbine. The capacities of the boilers and turbines are different. There are steam and electricity transportations between the process network and the utility plant. Some processes generate steam, which is transported to the utility plant. The utility plant produces steam and electricity and then supports them to the process network according to the demands of each steam grade and electricity. NCC companies can sell steam to other companies. Sometimes NCC companies sell electricity generated by their steam turbine to a regional power company.

If the three companies are merged, the new merged company can operate all process networks and utility plants of the three companies. We assume that company D is the aggregate of companies A, B, and C. The process network and utility plant diagram of company D are shown in Figure 1.

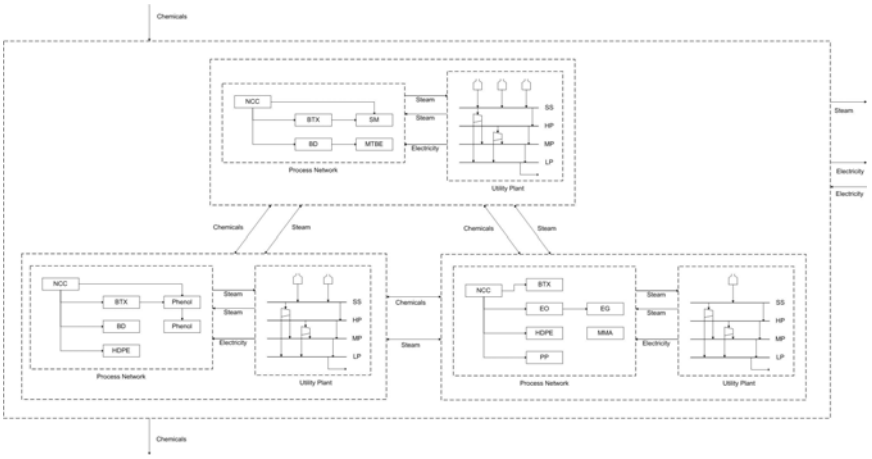


Figure 1. Process network and utility plant diagram of Company D

3. Mathematical Modeling

A novel MILP model is developed for optimization of a process network and a utility plant. The model maximizes the profits of target companies under given price and parameter data. The nomenclature is presented in Appendix.

3.1. Objective Function

$$\begin{aligned}\text{Maximize } z = & \sum_t \sum_h \sum_j REV_{hjt} - \sum_t \sum_h \sum_j PUR_{hjt} - \sum_t \sum_n \sum_k FP_{nkt} \cdot fuel_{nkt} \\ & - \sum_t \sum_n \sum_i (U_{nit}(0.4 \cdot b_{nit} + 0.6 \cdot r_{nit}) + F_{nit})\overline{X}_{nit} - \sum_t PE_t \cdot ep_t + \sum_t SE_t \cdot es_t \\ & + \sum_t \sum_n \sum_o stms_{not} \cdot stmp_{ot} - \sum_t \sum_n \sum_{n'} \sum_j invct_{n\,n'\,jt} - \sum_t \sum_n \sum_{n'} \sum_o invst_{n\,n'\,ot}\end{aligned} \quad (1)$$

The objective is to maximize the net present value of the profit. A company intends to operate their plant optimally based on a given business condition. The profit is calculated by subtracting chemical purchasing, operating cost and investment cost from the chemical sales.

3.2. Purchasing and Selling

$$Qp_{jt} = \sum_n nqp_{njt},\ \forall j,t \ (2),\ Qs_{jt} = \sum_n nqs_{njt},\ \forall j,t \ (3)$$

Equations (2) and (3) calculate amount of chemical purchasing and selling, Qp_{jt} and Qs_{jt}, of the merged company. The types of contracts for purchasing include (1) fixed price, (2) discount after a certain amount, (3) bulk discount, and (4) fixed duration. Equations to determine purchasing cost for each type of contract are given by:

$$PUR_{jt} = \sum_h PUR_{hjt},\ \forall j,t \ (4),\ Qp_{jt} = \sum_h qpc_{hjt},\ \forall j,t \ (5)$$

$$0 \le qpc_{hjt} \le M \cdot bqpc_{hjt},\ \forall h, j, t,\ M = \text{large positive} \ (6),\ \sum_h bqpc_{hjt} \le 1,\ \forall j,t \ (7)$$

Equation (4) means that PUR_{jt}, purchasing cost of chemical j at time t, is defined as the sum of PUR_{hjt}. Equation (5) represents that Qp_{jt}, purchasing amount of chemical j at time t, is the sum of qpc_{hjt}, purchasing amount of chemical j in contract h at time t. Equation (6) activates binary variable for qpc_{hjt}. Large M used in this study is 10^6. The number of contracts that can be made at time t is constrained by equation (7). In the same way to the purchasing contracts, a company can sell chemicals through different types of contracts. The types of contracts for selling include (1) fixed price, (2) discount after a certain amount, (3) bulk discount, and (4) fixed duration. Detailed models for contracts, operation of process network and utility system, and investment are refereed to works of Park et al. (2006) and Yoon et al. (2008, 2009).

4. Optimization Study

In this section, we apply the proposed mathematical model to individual companies, A, B and C and a merged company D. Synergy of merger is defined as difference between profit of company D and the sum of profits of companies A, B and C. The case study is based on five years time periods. An interest rate used in the optimization is five percent. Table 1 shows model statistics for the optimization study.

Optimization results show the NPV of profits of three individual companies and a merged company. Synergy is calculated using the results. Table 2 shows the results and synergy. The results exhibit that the merger of three NCC companies creates synergy of 1,851 million dollars. The synergy increases the profit by almost fifty percent.

Investment for chemical and steam transportation is 2.93 million dollars that is much smaller than the synergy created.

Table 1. Model statistics for the optimization study

Company	Equations	Continuous variables	Binary variables	Solution time (s)	Iterations
A	24158	18936	6905	0.406	618
B	24158	18936	6905	0.343	368
C	24158	18936	6905	0.234	397
D	49773	36826	11860	3.000	5220

Table 2. Optimization results

Profit (mil. $)	Individual Companies	A	1,333.93
		B	1,093.84
		C	1,305.13
		Sum	3,732.89
	Merged Company	D	5,584.36
Synergy (mil. $)			1,851.46
Synergy Effect (%)			49.60
Investment (mil. $)			2.93

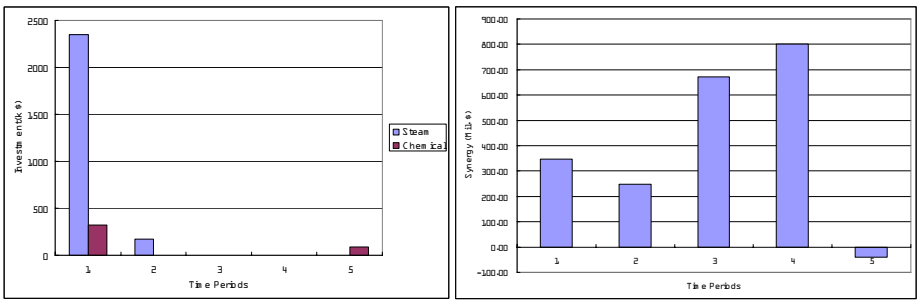


Figure 2. Investment cost (left) and Synergies in each time period (right)

A merger needs rationalization of production facility. For petrochemical companies, the rationalization means integration of process networks and utility systems. The integrations require capital investment for chemical and steam transportations. Figure 2 (left) shows investment cost for chemical and steam transportations. Investment for steam transportation is much larger than investment for chemical transportation. Most investment is made in time period 1.

Synergies in each time period are shown in Figure 2 (right). Except in time period 5, positive synergy is created in every time period. This study considers four kinds of synergy factors: process network integration, utility system integration, fixed cost reduction, and contracts for purchasing and selling. Contributions of four factors to the synergy need to be analyzed. In order to identify contributions of the factors, optimization studies are carried out without fixed cost reduction and chemical and steam transportations respectively. The chemical transportation represents process network integration to create synergy. The steam transportation represents utility system integration to create synergy. The contribution of contracts is calculated by subtracting

the contributions of fixed cost reduction, process network integration and utility system integration from the total synergy. Table 3 shows results of the optimization studies and contribution of each factor.

Table 3. Contributions of four factors (Mil. $)

Profit of the merged company, D	5584.36
without fixed cost reduction	5427.23
without process network integration	5423.44
without utility system integration	5538.68
Contributions of	
Fixed cost reduction	157.13 (8.49 %)
Process network integration	160.92 (8.69 %)
Utility system integration	45.68 (2.47 %)
Contracts	1487.74 (80.35 %)

The results represent that contracts in purchasing and selling contribute the most to synergy. Contract affects the revenue and the purchasing cost. Figure 3 shows that the amount of naphtha purchasing in Company D is larger than the sum of the amounts of individual companies. Unit naphtha purchasing cost of D is lower than those of individual companies on average. Selling amount of ethylene of D is larger than the sum of those of A, B and C in Figures 4. Unit selling price of ethylene, however, is almost the same in A, B, C and D. In results, the larger purchasing amount of naphtha in the merged company leads to a better price of the naphtha and then, produces more products.

5. Conclusion

This paper proposes the novel mathematical model to quantify various synergies in the merger of petrochemical companies within a complex. We focus on synergy from contracts in purchasing and selling, process network integration, utility system integration and fixed cost reduction. The investment term is included to calculate the cost to make chemical and steam transportation real. The proposed model integrates planning of the contracts for purchasing and selling and operation of the process network and utility system with M&A of petrochemical companies. Specific planning for post merger situation leads to precise estimation of the synergy. The optimization study is carried out on Korean NCC companies. The results show that the merger increases profit by fifty percents. The purchasing contract contributes the most to creating synergy. The larger purchasing amount leads to lower unit cost of naphtha. Korean petrochemical industry seriously needs to consider the merger as a growth strategy.

References

S. Yoon, SB. Park, S. Park, J. Lee, PM. Vederame, CA. Floudas, 2008, Synergy in Mergers of Petrochemical Companies within a Complex Considering Purchasing and Selling Advantage, Ind. Eng. Chem. Res., 47, 5556

M. Park, FD. Mele, S. Park, 2006, Modeling of Purchase and Sales Contracts in Supply Chain Optimization, Ind. Eng. Chem. Res., 45, 5013

S. Yoon, SB. Park, S. Park, J. Lee, PM. Vederame, CA. Floudas, 2009, Selecting the Optimal Target Company Based on Synergy Calculation for the Vertical Merger in a Petrochemical Complex., Ind. Eng. Chem. Res., 48, 1511

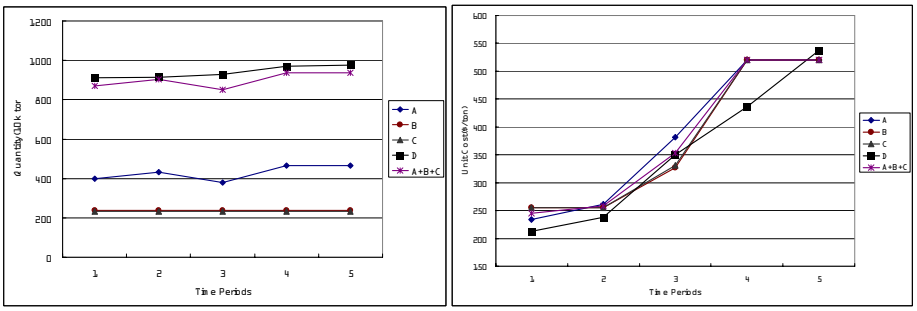


Figure 3. Naphtha purchasing amount (left) and unit purchasing cost (right)

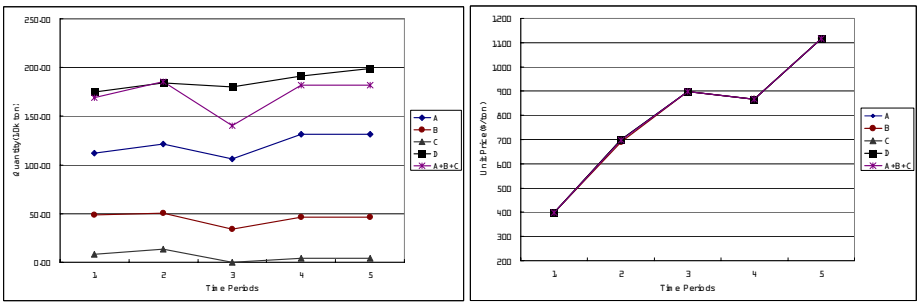


Figure 4. Ethylene selling amount (left) and unit selling price (right)

Appendix: Nomenclature

Sets

n: Company, i: Process, j: Chemical, k: Boiler, h: Contract, o: Steam grade, t: Time

Parameters

ep_t, es_t: electricity purchasing and selling price ($/MWh)
F_{nit}, U_{nit}: fixed cost and and utility cost for process ($/ton)
FP_{nkot}: fuel price of boiler ($/toe)
$stmp_{ot}$, $stms_{not}$: steam selling price and amount ($/ton)
$\overline{X}_{nit}$: process capacity (ton/yr)

Variables

b_{nit}: binary variable for process operation
$fuel_{nkot}$: fuel consumption in boiler (toe/yr)
$invct_{nn'jt}$, $invst_{nn'ot}$: investment cost for chemical and steam transportations ($)
PE_t, SE_t: electricity purchasing and selling (MWh/yr)
PUR_{hjt}, REV_{hjt}: purchasing cost and sales revenue of chemical in contract h ($/yr)
r_{nit}: operating ratio of process (ton/yr)

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal Assignment of Plant Operators on Basis of Shift's Ability Evaluation

Masaru Noda and Hirokazu Nishitani

Graduate Scool of Information Science, Nara Institute of Science and Technology, 8916-5 Takayama, Ikoma 630-0192, Japan

Abstract

A mathematical programming approach is proposed for solving a shift operator assignment problem on the basis of an evaluation of each shift's abilities. To evaluate the abilities of the workers of a shift, the minimum required skill levels for all processes are defined for three plant conditions: normal situations, abnormal situations, and emergencies. The ratio of successful combinations to all the assumed combinations of abnormal processes is defined as an index of the shift's ability to cope with the situation. The optimal assignment of operators to each shift is formulated as a nonlinear integer programming problem, where the minimum ability among all shifts is maximized. The optimal solution of the operator assignment problem shows us the limitation of operations by the assigned shift members under the specified situation. The proposed approach is helpful for managing abnormal situations and for improving operator training.

Keywords: Operator assignment, Multi-skilled operator, Shift ability evaluation, Abnormal situation management

1. Introduction

Most plants in the chemical industry are operated continuously by shifts that are formed with a certain number of operators. Each shift is composed of multi-skilled operators who have different skill levels for different processes. As a result, when the condition of a process becomes abnormal, a higher-skilled operator on the shift who can cope with the abnormality must temporarily take the place of a lower-skilled operator. Thus, for abnormal situation management, it is necessary to quantitatively evaluate a shift's ability to cope with plant abnormalities.

Previously, we proposed a quantitative method to evaluate the abilities of a shift under specified abnormal plant conditions (Noda *et al.*, 2008). We defined an index for a shift's ability to cope with abnormal situations by the ratio of successful combinations to all possible combinations of abnormal processes. We can find bottlenecks of the shift by investigating all of the patterns that the shift cannot cope with. In order to resolve these bottlenecks, we need to optimally assign operators to shifts on the basis of the shift's ability evaluation or to improve the skill levels of operators.

A human resource management problem in an organization operated by shifts had been discussed by many researchers. For example, Ikegami *et al.* (2003) proposed a new mathematical model and an approach to the nurse scheduling problem. In this study, we formulate a new mathematical programming problem for solving a shift operator assignment problem in a plant on the basis of the shift's ability evaluation. In the formulation, the minimum ability among all shifts is maximized by optimally assigning

operators to shifts. The proposed method is useful for abnormal situation management that ensures safe plant operations under specified abnormal situations.

2. Evaluation method of shift's ability

2.1. Definition of operator's skill level

A shift is composed of multi-skilled operators, who have different skill levels for different processes. In order to assess the ability of a shift to cope with abnormal plant conditions, we first need to quantitatively evaluate the skill level of each operator in each process. The definition of skill levels of course depends on the company and the plant (Takeda, 2002), but in this study, we use the following five skill levels.

Skill level 1: Beginner
Skill level 2: Able to do routine operations
Skill level 3: Able to cope with minor problems
Skill level 4: Able to cope with abnormal conditions
Skill level 5: Able to cope with emergencies

2.2. Formulation of minimum requirements

To quantitatively assess the ability of a shift to cope with a specified abnormal situation, we formulate a linear integer programming problem, where a typical continuous chemical plant with several processes operated by different shifts is used as an example to explain the formulation. In this plant, there are N_p processes, and each process is operated by one operator. Thus, the total number of operators in a shift is N_p. We define the minimum required skill level for all processes under three types of plant conditions: normal, abnormal, and emergency. The minimum skill levels required for operating normal, abnormal, and emergency processes are 2, 4, and 5, respectively.

To formulate an optimal shift assignment problem, we introduce an integer variable $z(i, p)$. When operator i is assigned to process p, $z(i, p) = 1$. When operator i is not assigned to process p, $z(i, p) = 0$. Because each operator can be assigned to only one process, the following constraint must be satisfied for all operators:

$$\sum_{p=1}^{N_p} z(i, p) = 1 \quad \text{for all } i, \tag{1}$$

and because each process is operated by one operator, the following constraint must be satisfied for all processes:

$$\sum_{i=1}^{N_p} z(i, p) = 1 \quad \text{for all } p. \tag{2}$$

To quantitatively assess the ability of a shift to cope with a specified situation, we introduce the objective function given by Eq.(3), where $c(i, p)$ is a cost coefficient calculated by Eq.(4). Here, $s(i, p)$ is the skill level of operator i for process p, and $e(p)$ is the minimum skill level required for operating process p. Eq. (3) denotes the total sum of deficiencies of operator's skill levels from satisfying the minimum requirements for all processes.

$$\sum_{p=1}^{N_p} \sum_{i=1}^{N_p} c(i, p) \cdot z(i, p) \tag{3}$$

$$\left.\begin{aligned} c(i, p) &= 0, && \text{if } s(i, p) \geq e(p) \\ c(i, p) &= e(p) - s(i, p), && \text{if } s(i, p) < e(p) \end{aligned}\right\} \tag{4}$$

When the skill levels of all operators in a shift and all combinations of abnormal processes are given, an optimal assignment problem of shift operators can be formulated as the following linear integer programming problem (LIP):

(LIP)
Minimize Eq. (3)
subject to Eqs. (1) and (2).

If the value of the objective function for the optimal solution is zero, the shift can satisfy the minimum requirements. When it is not zero, the shift is unable to cope with the specified situation.

2.3. Shift's ability evaluation

The LIP can be used to find combinations of abnormal processes that the shift cannot cope with. The ratio of combinations of abnormal processes with which a shift can cope to all the assumed combinations of abnormal processes is used as an index of the shift's ability to cope with abnormal situations. When this index is 100%, the shift can cope with all the assumed combinations of abnormal processes.

In this study, the following three evaluation indices of a shift's ability are defined by Eqs. (5) – (7).

$$A_{\mathrm{n}} = \frac{N'_{\mathrm{n}}}{N_{\mathrm{n}}} \times 100\ [\%] \tag{5}$$

$$A_{\mathrm{a}} = \frac{N'_{\mathrm{a}}}{N_{\mathrm{a}}} \times 100\ [\%] \tag{6}$$

$$A_{\mathrm{e}} = \frac{N'_{\mathrm{e}}}{N_{\mathrm{e}}} \times 100\ [\%] \tag{7}$$

Where, A_{n}, A_{a}, and A_{e} are evaluation indices of the shift's abilities under normal, abnormal, and emergency plant situations, respectively, and N'_{n}, N'_{a}, and N'_{e} indicate the total numbers of combinations of normal, abnormal, and emergency processes with which a shift can cope. The values N_{n}, N_{a}, and N_{e} represent the total numbers of all the assumed combinations of normal, abnormal, and emergency processes. The possible combinations of abnormal and emergency processes should be determined on the basis of the history of problems and on the characteristics of each process. If the possible combinations of processes are unknown, it may be worth investigating the maximum number of processes in the plant that the shift can handle safely.

3. Optimal assignment of operators based on shift's ability evaluation

We consider an integrated control room, where the plant is operated by N_{g} shifts. There are N_{p} processes in a plant, and each process is operated by one operator. Thus, the total number of operators in a shift is N_{p}. To formulate a mathematical programming problem for solving a shift operator assignment problem on the basis of the shift's ability evaluation, we introduce the integer variable $x(i, g)$ as an optimization variable. When operator i is assigned to shift g, $x(i, g) = 1$. When operator i is not assigned to shift g, $x(i, g) = 0$. Because each operator can be assigned to only one shift, the following constraint must be satisfied for all operators:

$$\sum_{g=1}^{N_g} x(i,g) = 1 \quad \text{for all } i \,. \tag{8}$$

Furthermore, because the total number of operators in each shift is N_p, the following constraint must be satisfied for all shifts:

$$\sum_{i=1}^{N_i} x(i,g) = N_p \quad \text{for all } g \,, \tag{9}$$

where N_i is the total number of operators in the plant. Because all the shifts have to cope with a normal operating condition, the following constraint must be satisfied for all shifts.

$$A_n(g) = 100 \quad \text{for all } g \tag{10}$$

When the minimum shift's ability in abnormal situations is given from the viewpoint of safe plant operations, Eq. (11) is introduced in the optimization problem as a constraint.

$$A_a(g) \geq A_{a,\min} \quad \text{for all } g \tag{11}$$

Here, $A_{a,\min}$ denotes the minimum limit for A_a.

In this study, the minimum ability in emergencies among all shifts is maximized by optimally assigning operators to all shifts. The operator assignment problem is newly formulated as a nonlinear integer programming problem (NIP).

(NIP)
Maximize (min $A_e(g)$)
subject to Eqs. (8), (9), (10), and (11).

4. Case study

4.1. Shift's ability evaluation

The proposed method is demonstrated through a case study, where a chemical plant with seven processes is operated by shifts, 1 and 2. Each shift consists of seven operators, and each process is operated by one operator. Thus, the total number of operators in this plant is fourteen. The skill level of operator i for process p is listed in Table 1. Fourteen operators were assigned to two shifts so that average skill levels of operators in the shifts became nearly equal (3.4).

Table 1 Skill levels of all operators in shifts 1 and 2

	Shift 1								Shift 2						
i \ p	1	2	3	4	5	6	7	i \ p	1	2	3	4	5	6	7
1	5	5	5	5	5	4	4	8	4	4	4	4	5	5	5
2	5	5	4	4	3	4	4	9	5	5	4	4	3	4	4
3	4	4	3	3	1	4	4	10	4	4	3	3	3	5	5
4	4	4	3	4	1	3	3	11	4	4	3	4	3	4	4
5	4	4	3	4	1	3	3	12	4	4	3	4	1	1	1
6	4	4	3	4	1	1	1	13	4	4	4	4	1	1	1
7	4	4	3	4	1	1	1	14	3	3	3	3	1	1	1

First, the abilities of these two shifts were evaluated for the three plant conditions: normal, abnormal, and emergency situations. In the normal situation, all processes in

the plant are normal. In an abnormal situation, it is assumed that three processes become abnormal simultaneously. In an emergency, it is assumed that emergency conditions occur in two processes simultaneously. The total numbers of patterns for normal, abnormal, and emergency situations are 1, 35 ($_7C_3$), and 21 ($_7C_2$) respectively, and are listed in Tables 2, 3, and 4. In the tables, 0 means that the process is normal, 1 means that the process is abnormal, and 2 means the process is in an emergency. The minimum skill levels required for operating normal, abnormal, and emergency processes are 2, 4, and 5 respectively.

Table 2 Pattern of normal plant condition

No	P1	P2	P3	P4	P5	P6	P7
1	0	0	0	0	0	0	0

Table 3 Patterns of abnormal conditions

No	P1	P2	P3	P4	P5	P6	P7
1	0	0	0	0	1	1	1
2	0	0	0	1	0	1	1
3	0	0	0	1	1	0	1
4	0	0	0	1	1	1	0
5	0	0	1	0	0	1	1
6	0	0	1	0	1	0	1
7	0	0	1	0	1	1	0
8	0	0	1	1	0	0	1
9	0	0	1	1	0	1	0
10	0	0	1	1	1	0	0
11	0	1	0	0	0	1	1
12	0	1	0	0	1	0	1
13	0	1	0	0	1	1	0
14	0	1	0	1	0	0	1
15	0	1	0	1	0	1	0
16	0	1	0	1	1	0	0
17	0	1	1	0	0	0	1
18	0	1	1	0	0	1	0
19	0	1	1	0	1	0	0
20	0	1	1	1	0	0	0
21	1	0	0	0	0	1	1
22	1	0	0	0	1	0	1
23	1	0	0	0	1	1	0
24	1	0	0	1	0	0	1
25	1	0	0	1	0	1	0
26	1	0	0	1	1	0	0
27	1	0	1	0	0	0	1
28	1	0	1	0	0	1	0
29	1	0	1	0	1	0	0
30	1	0	1	1	0	0	0
31	1	1	0	0	0	0	1
32	1	1	0	0	0	1	0
33	1	1	0	0	1	0	0
34	1	1	0	1	0	0	0
35	1	1	1	0	0	0	0

Table 4 Patterns in emergencies

No	P1	P2	P3	P4	P5	P6	P7
1	0	0	0	0	0	2	2
2	0	0	0	0	2	0	2
3	0	0	0	0	2	2	0
4	0	0	0	2	0	0	2
5	0	0	0	2	0	2	0
6	0	0	0	2	2	0	0
7	0	0	2	0	0	0	2
8	0	0	2	0	0	2	0
9	0	0	2	0	2	0	0
10	0	0	2	2	0	0	0
11	0	2	0	0	0	0	2
12	0	2	0	0	0	2	0
13	0	2	0	0	2	0	0
14	0	2	0	2	0	0	0
15	0	2	2	0	0	0	0
16	2	0	0	0	0	0	2
17	2	0	0	0	0	2	0
18	2	0	0	0	2	0	0
19	2	0	0	2	0	0	0
20	2	0	2	0	0	0	0
21	2	2	0	0	0	0	0

The evaluation results of the shift's ability under the three plant conditions for two shifts in Table 1 are summarized in Table 5. The evaluation results indicate that the abilities of the two shifts are high in normal and abnormal conditions, but very low in

emergencies. The ability of shift 1 in an emergency is lower than that of shift 2, although the average skill levels of the two shifts are equal. The reason for the low ability of shift 1 in an emergency is due to an unbalanced assignment of operators to the shifts. In shift 1, only two operators have skill level 5 and no operator can cope with an emergency in process 6 or 7.

Table 5 Shift abilities in three plant conditions

Shift no.	Operator's no.	A_n	A_a	A_e
1	1,2,3,4,5,6,7	100 %	97.2 %	9.5 %
2	8,9,10,11,12,13,14	100 %	100 %	42.9 %

4.2. Optimal assignment of plant operators

In this section, all operators were optimally assigned by using the proposed method. The optimal operator assignment problem was solved using a standard NIP solver, where $A_{a,min}$ was set as 100% for two shifts. Twenty-four patterns of operator assignment were found as optimal solution, and the optimal value of the objective function A_e was 33.3%. Table 6 presents one of the optimal operator assignments and abilities of shifts 1' and 2'. By optimally assigning operators into shifts, the ability of shift 1 in an emergency was drastically improved from 9.5% to 33.3% without sacrificing anything from shift 2. The simulation results proved that we can improve shift's abilities by appropriately assigning operators to each shift.

Table 6 Example of optimal assignment of plant operators

Shift no.	Operator's no.	A_n	A_a	A_e
1'	2,3,5,8,9,12,14	100 %	100 %	33.3 %
2'	1,4,6,7,10,11,13	100 %	100 %	47.6 %

The optimal solution of the operator assignment problem shows us the limitations of operating in shifts when a plant experiences abnormal conditions or in emergency. The evaluation results are helpful for abnormal situation management and for operator training. For example, when a plant condition moves outside of operable conditions of the shift in charge, the shift must ask for assistance to continue operating or shut down the plant safely. It is also worth planning training programs to improve bottlenecks that are found by the optimization.

5. Conclusion

We formulated an optimal operator assignment problem as the mathematical programming problem on the basis of a shift's ability evaluation for industrial chemical plants. The proposed approach is very useful for improving a shift's abilities under abnormal and emergency situations.

The objective function and constraints in the optimization problem can be easily modified according to the characteristics of individual chemical plants.

References

Ikegami, A., Niwa, A., 2003, A Sub-problem-centric Model and Approach to the Nurse Scheduling Problem, *Mathematical Programming*, 97, 517-541.

Noda, M., Kuratsune, K., and Nishitani, H., 2008, Optimal Assignment of Shift Operators for Safety Operation and Skill Advancement, *Proceedings of FOCAPO 2008*, 311-314.

Takeda, M., 2002, Activity for Environment and Safety in the Iwakuni – Ohtake Works, *Human Factors in Japan*, 7(1), 54-64. (in Japanese)

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

An eco-efficiency study for a WEEE recovery network: the Portuguese case

Maria Isabel Gomes-Salema[a], Ana Barbosa-Povoa[b] and Augusto Q. Novais[c]

[a]*Centro de Matematica e Aplicacoes, FCT – UNL , Monte da Caparica, 2829-516 Caparica, Portugal. e-mail: mirg@fct.unl.pt*
[b]*CEG-IST, Av. Rovisco Pais, 1049-001 Lisboa, Portugal.*
[c]*Dep. Modelacao e Simulacao, INETI, Est. do Paço do Lumiar, 1649-038 Lisboa, Portugal.*

Abstract

The rapid growth of electric and electronic equipment waste (WEEE) transformed this waste stream into a worldwide problem. The Directive 2002/96/EC on electrical and electronic waste aims at the reduction of the environmental impact of WEEE, encouraging end-of-life management, eco design, life cycle analyses and extended producer responsibility. However, this legislation may not produce the results the legislator aimed for. In this work we analyse the environmental impact of a WEEE recovery network in the Portuguese context. With this aim, a model, previously developed by the authors for the optimal design of this network using economic indicators (Salema, 2007), was now adapted to design the network subjected to the minimization of environmental performance indices.
The original mathematical formulation was found to be flexible and easily adapted to the two types of indices and the major differences between the optimal network configurations obtained, were identified and discussed.

Keywords: Recovery network, environmental impact, optimization.

1. Introduction

The rapid growth of electric and electronic equipment waste (WEEE) transformed this waste into a worldwide problem. The European Union estimates a growth of 3 to 5% per year for the WEEE, a figure three times greater than for general waste. Around 90% of WEEE is going to landfills. In order to deal with the problem the Directive 2002/96/EC on electrical and electronic waste was defined that aims at the reduction of the environmental impact of WEEE, encouraging end-of-life management, eco design, life cycle analyses and extended producer responsibility. Under this directive, producers are responsible not only for new products placed on the market, but also for those equipments sold before this directive. This represents a new driving force that will compel producers to support collecting and recycling costs of their products.
In Portugal, a non-profit organization (Amb3e) was created by a group of 57 equipment producers, whose mission is to design and manage the integrated system for the recovery of WEEE.
Within this organization producers were able to deal with this challenge in a more efficient way since it allows for economies of scale, reducing the usual uncertainty associated to the quantity and quality of end-of-life equipments.

Although these recovery systems aimed at the reduction of the environmental impact, the collection of products to reduce the amount of hazardous materials in landfills and the recovery of valuable disposed materials, may still lead to a potential increase of global warming, since products have to be transported to the recovery/treatment facilities.

Several tools have been developed to evaluate the environmental impacts of products and processes. Life-Cycle Assessment is one of them, which evaluates the emissions generation and material consumption during the entire product life. The eco-indicator'99 (Goedkoop et al. 2000) aggregates into one single index the information on the multiple aspects involved in an environmental impact assessment, taken from the life cycle analysis methodology Also carbon footprint measures in terms of greenhouse gas emissions are very useful, amongt others. More recently, the eco-efficiency concept has been gaining importance in the business community since it combines the economical and environmental objectives. All these tools provide environmental indicators that can be very useful in the design of environmental friendly recovery networks. Such tools have not yet been explored in what concerns the definition and operability of supply chains. Traditionally, the scientific community has been studying supply chain designs through the optimization of economical indicators. This situation, however, is now changing, since current environmental problems ask for new designs that also account for environmental impacts of transportation, production, water consumption and raw-material depletion, recycling, among others.

Few works have been published that evaluate the environmental impact of these structures. Some have nonetheless been published that consider a single type of electric and electronic equipment (Krikke et al. 2003). Neto et al. (2008) proposed a methodology to assess the eco-efficiency in logistics networks and developed a multi-objective model, which was applied in Germany to a recovery network for entertainment equipments, such as TV sets and Wi-Fi equipments.

In this work we analyse the environmental impact of a WEEE recovery network in the Portuguese context. A previous work from the authors on the design of this network using economic indicators (Salema, 2007) is now extended to treat this new problem. The two network structures are compared and the most important aspects affecting each performance indicator highlighted.

This paper is organized as follows. In the next section the context for the Portuguese network is presented. Next, in section 3, the methodology for the impact assessment and a detailed description of the model are given. This is followed by an analysis on the results. Lastly, some final remarks will be drawn.

2. The Portuguese recovery network details

The Portuguese recovery network for WEEE has a two echelon network structure formed by the WEEE sources (that given the strategic nature of this problem are the 278 municipalities of the mainland Portuguese territory), the collection centres that will collect and sort products according to five pre-defined categories, and the recycling facilities already operating in Portugal.

The recycling facilities are located at Tondela and Setúbal. The latter only recycles lighting equipment, which leaves the former with the whole of the remainder. Amb3e estimates that lighting equipment represents just about 1.5% of the total volume to be collected. Tondela facility will therefore process the remaining 98.5%. This will have a great impact on the network, as it will be shown in section 4.

Amb3e needed to perform contractual agreements with companies and organizations that will act as collection centres. Therefore, the major goal is to determine which municipalities are the best locations for collection centres. Together with the locations, some planning information is also needed to better operate this network.
For the year of 2006, Amb3e set the target of recovering 3 kg per inhabitant per year. This objective would allow them to achieve 75% of the legal target imposed by the European directive. This implied that about 30000 tonnes had to be collected and at least 80% of it should be recovered. A maximum distance was imposed so that each Portuguese municipality had a collection centre within a 65 km radius.
In this work, the primary objective is to choose the best locations for the collection centres. In addition, the optimal planning of the network, for the year 2006, will also be determined. This will be possible by modelling a time horizon that is divided in planning units. For the case below the planning units is the trimester.

3. An Eco-efficient Network Model

3.1. Assessing the environmental impact

Several methodologies have been developed over the last two decades to assess environmental impacts of products and services. Life cycle assessment (LCA) is the most applied methodology. It analyses the entire product life, starting at the extraction of raw-materials down to its use by the final customer, and measures each and every impact it has on the environment (Daniel et al., 2004). This involves a very complex process and in order to measure it, some indicators have been developed: eco-indicator99 that measures impacts using a specific unit called points, the carbon footprint which measures impacts in terms of CO_2 emissions, the eco-cost2007 that converts impacts into Euros, etc.
The EU has done an extensive analysis of the impacts that the EU directive on WEEE had on the Member States (Huisman et al., 2008). This report evaluates environmental impacts with several indicators and methodologies. One of the greatest challenges of any work related to the recovery of WEEE is the diversity of products that fit this designation. Products can go from a small hairdryer or a kid's toy, to a refrigerator, LCD television or even an automatic dispenser machine. Therefore it is very difficult to provide an accurate measurement of the environment impact of each of the 10 categories established by the legislation. In Huisman et al. (2008) subcategories were created and impacts were assessed considering average values within each subcategory. Each of these groups of products is then characterized in terms of average weight and by several environmental impact categories.
For the present work, a higher level of product aggregation is needed since Amb3e only considers five groups of products: large household appliances (C1), cooling and freezing equipment (C2), small household appliances (C3), monitors and televisions (C4), lighting equipment (C5).
In terms of environmental measures, we considered four major impact sources: **the impact of establishing the collection centres**: this impact reflects the construction of a facility similar to a warehouse; no assessment was made regarding the impact of the land use; **the transportation impact between WEEE sources and collection centres** (by km and tonne): this value was based on the impact values provided by TU Delft University (Eco-costs); **transportation impact between collection centres and recyclers** (by km, product and tonne): this value was based on the values of transportation given in the EU report; in categories with more than one subcategory, the average value is weighted by each product own weight; **the recovery impact**: this

reflects the gain in terms of environment for the recycling materials, therefore this is a negative value; and **the disposal impact**: that reflects the impact of sending to proper disposal products that could not be recovered; this operation does not have to be bad for the environment since some products can be incinerated, producing energy.
No impact value was given to products kept in stock. It was assumed that products were well conditioned and, since at the end of the planning period, all collect products have to be recycled, these stocks do not represent any harm to the environment. Nonetheless and since legal targets have to be meet, it was assumed that in the last period of the planning horizon.

3.2. The model

A model, previously developed by the authors (Salema, 2007), was adapted to design the network according to the environmental performance measures. The same model is used to minimize impacts and (but not simultaneously) to minimize costs. Therefore, the parameters defined for the objective function have in each case a different meaning, either impacts or costs. For reasons of confidentiality the real units are omitted and values presented are expressed in terms of monetary units (m.u.) or impact units (i.u.). Due to lack of space, only a general description of the model is next given. The model can be stated as: Given: WEEE volume to be collected at each source; recovery target set by legislation; sorting criteria to be performed at centres; initial stock levels at facilities; maximum storage capacity; maximum and minimum processing capacities; upper and lower bounds for flows; unit impact/ compensation fee given to recyclers and sorting centres; unit transportation impact/cost; unit processing and storage impact/cost. Determine, locations of sorting centres; flows amounts between sources of WEEE and sorting centres and between these and the recycling facilities; storage volumes at sorting centres and recycling facilities; processed and disposed volumes by recycling facilities, so as to minimize the total environmental impact or total cost of the network. In terms of constraints, the model is characterised by nine types of constraints related to: material balances; collection satisfaction; target levels assuring that the legal target is met; disposal, since it is estimated that 5% of the collected products are not proper to recycle; maximum stock levels; maximum and minimum values for flows; and finally sorting centres capacities limitations. In terms of variables, the model has three types of continuous variables (flow amounts, stock levels and non-satisfied return), and one type of binary (choice of entities to integrate the network). For further details regarding the model refer to (Salema 2007).

4. Results and discussion

The model described above was run in its original and modified versions: the former considering the economic parameters (cost model) and the latter the environmental impacts (impact model). All results presented were computed on a Pentium 1.7 GHz, using GAMS/CPLEX (built 22.9). First, the results obtained for the impact model will be presented, and then the results of both models will be discussed.

4.1. The environmental impact network

The optimal location of collection centres for the environmental impact model is depicted in Figure 1. The network is formed by 110 centres spread all over the country. The total impact of this network is $-3312*10^3$ i.u. Note this is a negative value, which means that the recovery of products is environmentally friendly. In a “doing nothing” context, the disposal of products in the municipal waste system in a uncontrolled landfill without energy recovery, would have an positive impact of about $371*10^6$ i.u. (this value was calculated using the data provided at Huissman et al (2008)).

Looking closer to the positive impacts shown in Figure 2, one sees that the transportation between collection centres and recycling facilities (flow F2) is responsible for 97% of the total positive impact ($157*10^3$ i.u.).
Comparing flows F1 and F2 (F1 refers to the connection between sources and collection centres), the first conclusion taken it that the location of recycling facilities is of extreme importance and, for this case in particular, at least a second unit should be considered. Although the construction and operation of a recycling unit has a positive contribution in terms of impacts, the decrease in the transportation impact most certainly will overshadow this negative effect.
This model produces a large amount of information regarding network planning: recycling, transportation, disposal, and storage. Due to the lack of space, only a few examples will be shown.
The first is related to the recycling plan, Figure 3. This figure illustrates very clearly the difference between both recycling facilities in terms of processed volume. Another important aspect are the disposal amounts. As mentioned above it is estimated that at least 5% of collected products will not be suitable for recycling, and should be properly disposed in landfills and, when possible, energy should be recovered. The optimal disposal volume is exactly 5%. The reason is that recycling activities reduce the environmental impact of the network. It will be shown below that the opposite scenario characterizes the cost model solution.

Figure 1: Environmental impact recovery network

4.2. Comparison between cost and environmental networks

Using the same model structure, the formulation mentioned above was used to determine the optimal solution in terms of economic objectives. The environmental impact of the optimal solution for the cost model was computed, as well as the cost related to the optimal solution of the impact model. All values are shown in Table 1.
This table shows some interesting aspects. The transportation costs between WEEE sources and collection centres are much higher in the impact solution. This is explained by the number of centres which is much smaller: in the impact solution there are 110 centres, while in the cost solution these are 169. Therefore, for the latter more travelling is needed.

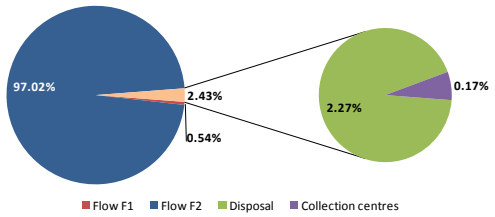


Figure 2: Weight of the environmental parameters in the total impact value.

With regard to transportation between collection centres and recycling facilities, environmental impacts show a larger difference than costs. More centres reduce the total travelling. Disposal costs present also significant differences. Note that these values (both for environmental impacts and costs) are four times higher in the cost model. The legal target imposes that 80% of collected products to be recycled. In the solution of the cost model, only 80% of products are recycled while in the impact model this percentage is 95%. The reason is that in the cost model, the recycling represents a cost and in the impact model this operation has a negative impact. Stocks only appear in the impact model. As mentioned above, no impact was assessed for products kept in stock. Therefore, the impact solution has high levels of stocks. These are, however, accounted for when costs are calculated.

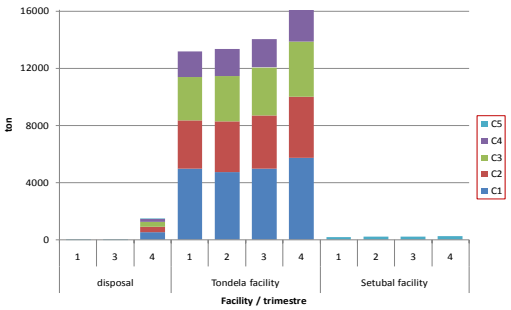


Figure 3: Recycling and disposal plans for the entire time horizon.

Table 1: Environmental impacts and costs in the optimal values of both models.

	Impacts (u.i.)		Costs (m.u.)	
	Impact model	**Cost model**	**Impact model**	**Cost model**
Recycling	-3 463 106	-2 916 300	9 169 466	7 721 656
Collection	253	253	782704	782 704
Transp. F1	820	58	1 241 304	88 483
Transp. F2	146 268	174 992	5 176 062	6 176 314
Disposal	3 415	13 659	73 840	295 360
Stocks	0	0	1 062 436	0
Total	-3 312 350	-2 727 338	17 505 812	15 064 517

5. Final remarks

In this work, a recovery network for WEEE recovery in the Portuguese context is designed considering environmental performance metrics. The optimal location for collection centres is presented and some planning details discussed. With the original model, the optimal network was also designed and planned considering this time an economic objective function. Both solutions were presented and compared.

As mentioned above, it is very hard to estimate accurately the environmental burden of each product category. Therefore, one way to overcome this limitation is to perform similar studies using different methodologies for impact assessment. Further work is being done to develop a methodology for assessing the environmental impact of global recovery networks.

References

Daniel, S. E., et al. (2004). *Ecological Indicators*, 4(2), 125-138.
Eco-costs, http://www.ecocostsvalue.com/httpdocs/content/html/5-data.html.
Neto, J. Q.F., et al. (2008). *IJPE*, 111(2), 195-208.
Goedkoop, M., et al. (2000). PRé - product ecology consultants.
Huisman, J., et al. (2008). United Nations University, Bonn, Germany.
Krikke, H., et al. (2003). *IJPR*, 41(16), 3689-3719.
Salema, M. I. (2007). PhD thesys, Instituto Superior Técnico, UTL, Lisbon.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Valuation of Clean Technology Projects: An application of Real Options Theory

Marcelo C. M. de Souza[a], Cristiano H. O. Fontes,[a] Silvio A.B. Vieira de Melo,[a] Antonio Francisco A. Silva Junior[a]

[a] *Programa de Engenharia Industrial - PEI, Escola Politécnica, Universidade Federal da Bahia – UFBA, R Aristides Novis, n. 2, Federação, Salvador - BA, 40210-630, Brasil*

Abstract

The valuation of investment in technologies that minimize the environmental impacts is not an easy task. Investment and environmental damages are partially or totally irreversible, the forecast period for the analysis of such investment is usually long and there are economic and environmental uncertainties which are still enhanced by the fact that environmental damages is not a linear function of the impacts. A better understanding of the behavior of these uncertainties and of the strategies to incorporate these uncertainties in the design valuation would be useful in the decision making associated to the investment with relevant environmental impacts. This work presents an extensive review about the application of the Real Options Analysis in the investment decision associated to the environmental problems and discusses the potential of this technique in the investment analysis of these cases. The aspects discussed denote that the approach of Real Options Analysis is better than the traditional methods and possibly is capable of predicting the benefit obtained with the reduction of the environmental risks provided by the use of clean technology.

Keywords: Real Options Analysis, Clean Technology, Environmental Evaluation.

1. Introduction

The environmental sustainability debate, focusing on the ability to combine economic growth with the environment preservation, has been broad and deep. According to Jeffrey Sachs (2008), until now, just one thing is true: the current trajectory of human activity is not sustainable. If the humanity maintains the same current technologies considering the scale up due to the fast economic growth of big population centers such as China and India, the environment and humanity's welfare will collapse.

The traditional production process has caused social and environmental problems with increasing impact potential (SEIFFERT, 2007). Only more recently, the practice of pollution prevention and clean production strategies has begun to be adopted in order to reconcile economic performance together with improvements in the environmental performance (KIPERSTOK, 2008). In most cases the adoption of sustainable practices have been employed due to the pressure imposed by the society and many investors still consider the sustainable development a necessary (but not desirable) strategy which involves the adoption of a series of procedures, regulations and onerous responsibilities. They do not consider sustainable practices as a strategic business opportunity (HART and MILSTEIN, 2004).

Generally investor's motivation is just the increase of wealth. The adoption of strategies based on clean technologies can imply more investment when compared with traditional processes, and can also affect the future cash flows and the risk of business through changes in costs and revenues and the reduction of environmental risk. It is usually difficult to predict the real benefits of these practices for the company. Therefore, the environmental risks, usually difficult to quantify, are not considered in decision-making process (BEKEFI and EPSTEIN, 2007).

Traditional investment analysis methods use an approach based on discounted cash flow. This is questioned by professionals and academics due to its inability to capture the administration's flexibility to adapt and revise subsequent decisions caused by market changes, being incapable to evaluate all value sources of an investment (TRIGEORGIS, 1993, SIEGEL et al. 1987, TRIGEORGIS, 2007). In these methods the uncertainty related to the project and the reactions to the market conditions changes are treated superficially, changing the rates to update the risk (MORCK et al. 1989). Thus, the higher the uncertainty and flexibility of management to react to market changes, the more undervalued will be the project value using traditional methods of evaluation (COPELAND and ANTIKAROV, 2001).

More recently, a new framework for investments evaluation, called as Real Options Analysis (ROA), has treated uncertainties in investments design. The purpose of this article is to review the state of the art and evaluate the application of the Real Options Analysis in the investment decision associated to environmental problems and clean technology projects. Section 2 presents the concept of ROA and presents some arguments that justify its use in environmental analysis. Section 3 presents a review about ROA applications in works about environmental valuation and clean technology and also shows why the ROA application in environmental analysis is still incipient. Finally, the last section discusses the basic aspects that suggest the adoption of ROA in the valuation of clean technology projects.

2. Why use Real Options in the Environmental Investment Analysis?

Traditional methods of investment analysis based on discounted projected cash flows have some limitations with regard to the capture of the management flexibility. Initially these methods consider that the investment is reversible. Even when this is not true, if the firm does not undertake the investment now, it will not be able to in the future (DIXIT and PINDYCK, 1994). In addition the Net Present Value rule assumes that the decision initially taken can not change in the future.

The Real Options Analysis does not replace the Net Present Value (NPV), but rather represent a complement of this technique. The ROA can capture the value of the project flexibility resulting in a new NPV which some authors call the NPV expanded. These management flexibilities are basically related to the strategic positioning of the company. So, the analysis of real options can provide a connection between the strategic planning and the finance of the company. The possibility of a new entrant in the market, the possibility of a delay in the implantation of the project due to the wait for more information that would reduce the uncertainty, expand, contract or even abandon the business are all management flexibilities of the project. In fact, if a project has

flexibility options, the investment opportunity value is equal to the Net Present Value plus the value of the options.

The investment in projects associated with clean technologies is difficult to evaluate because it represents a direct cost for the investor and the return is not direct. The benefits provided by clean technologies comprise the reduction of pollution levels, the reduction of the costs for its treatment, the compliance to the environmental laws, the reduction of probability of environmental damages, the reduction of shutdown events and reputation risks that could imply in loss of market share and even the solvency of the company among others. But it is not so clear that the benefits will arise in the next production cycle. The environmental damage risk or the pollution level represents an externality. The externality is considered in some way a possible contingency in the project. An environmental contingency can be seen as an obligation to be exercised or not. If it is possible to calculate the price of this option it is possible to evaluate the project considering its externality.

Real Options Analysis (ROA) can provide a consistent method for valuation of projects with high uncertainties through the analogy between the financial options and the options for business management. It uses financial options pricing methods to evaluate this latter and has been presented as an alternative, which leads to better results in the evaluation of investment, especially in environments of high uncertainty (DIXIT and PINDYCK, 1995) (COPELAND and ANTIKAROV, 2001).

Considering that the effects of environmental impacts may be irreversible and that the value of environment preservation for future generations is uncertain, the benefit in the environment protection today must include a value of option. When this option is exercise, the managerial flexibility is lost. The social costs deriving from the anticipated exercise of this option can be large and should be measured.

3. ROA applications in Environmental Valuation.

The first works of environmental preservation presented in the literature that established the concept of irreversibility are due to Arrow and Fisher (1974) and Henry (1974). The option value concept in the economic environmental approach was developed and explained by Fisher and Hanemann (1987) and Hanemann (1989).

Recently some applications of real options theory in issues about environmental policy have been suggested, often to solve optimal timing problems due to the uncertainties and the irreversibility associated to the environmental degradation, its prevention and its economic consequences (Pindyck, 2000, 2002). Pindyck (2002) shows that the traditional approach using the net present value of the costs and benefits flows as a rule of decision, ignore three important features of the most environmental problems: i) the uncertainty of the future costs and benefits, ii) the irreversibility associated to the environmental policy, and iii) the possibility to delay some action and to wait for more information. He used a two-period model to illustrate the optimal timing problem and its implications on the environmental policy.

Dotsis et al (2005) extend the analysis of Pindyck (2000) allowing the state variables to follow jump diffusion processes. They analyzed the features of optimal timing decisions

related to the adoption of an environmental policy when there is a possibility of abnormal changes either in the evolution of the stock of pollutant, or in the future costs and benefits for environmental damage.

Lin, Ko, Yeh (2007) propose a continuous time model using the real option approach, (extension of Pindyck model), capable to determine when and how much greenhouse gas emissions should be reduced. The model uses a geometric Brownian motion to describe the economics and ecological uncertainties and also considers the existence of correlation between these variables.

Martins and Melo (2007) use Real Options Theory to analyze if a larger urban mangroves area must be used for the construction of a highway. They also use a continuous time model that considers a geometric Brownian motion to describe the benefits flow provided by the mangroves area.

There are few works using Real Options Analysis approach to evaluate environmental questions associated to private investment analysis. Sarkis and Tamrkin (2005) apply real options analysis in a real case study (*British Petroleum – Amoco*) for valuation one specific project with carbon credit generation. Espinoza and Luccioni (2007) present a model derived from an option to estimate the value of brownfields considering the environmental and market risks.

In another recently study, Pindyck (2007) shows the importance of uncertainties in the evaluation of environmental policies and attests that traditional models of cost benefits analysis are not suitable for evaluation of these policies because these have some limitations, namely, the quantification of the reduction in the environmental damage, the benefit from this reduction are unknown, the physical and ecological impacts are uncertain and there is relevant uncertainty related to the costs and benefits associated to the policy that will be employed; The discount rate to be used is not know exactly. Uncertainty about costs, benefits and discount rates are common in investment analysis, but, according to Pindyck (2007), three additional difficulties arise:

1. The environmental cost and benefit functions tend to be highly nonlinear and expected values are difficult to be used in this case.
2. Environmental policies usually include significant irreversibility, and they sometimes interact in complex way with uncertainty. There are sunk costs for the society (such as investment or as a discrete flow of expenditure). Furthermore, environmental damage is often partially or completely irreversible..
3. Environmental policies often comprise long time horizons. This increases the uncertainty associated to the cost, benefit and discount rate.

An irreversible investment opportunity is similar to a financial Call Option. A Call Option gives to the holder the right to acquire an active at a fixed price. The exercise of this option is irreversible. If the active value increases, the net return of the investment increases, however if the active value falls the investor does not exercise the option (to invest) and loses it was paid to have the right to invest.

A literature search was carried out in the database *Engineering Village* and the results are shown in Table 1 in number of publications identified. Combinations for "clean technology" or "clean production" and key words which mean processes of economic or financial assessment, such as: "Real Option" Investment Decision"," Investment Analysis" were not found.

Keywords Subject, Title and Abstract
Periodo 1999 - 2008
All Languages

And		"Real Option*"	"Investment decision"	"Investment Analysis"	"Clean* Technology"	"Clean* Production"	"Enviromental Valuation"	"Sunstainable Development"	"Enviromental Economic*"	"cost benefit analysis"
"Real Option*"	211		29	5	0	0	0	5	0	11
"Investment decision"	400			10	0	0	0	16	0	14
"Investment Analysis"	107				0	0	0	3	0	8
"Clean* Technology"	147					5	0	18	0	2
"Clean* Production"	94						0	14	0	2
"Enviromental Valuation"	47							1	1	7
"Sunstainable Development"	17.815								70	217
"Enviromental Economic*"	216									12
"cost benefit analysis"	6.820									

Table 1 - Results of literature search

"Sustainable development" and "Cost benefit analysis" produced the largest number of 217 results with great concentration in the years 2003 to 2005, still high in 2006 and 2007, confirming that most of the analysis of investment models that address issues of sustainability are still based on the method of cost-benefit analysis. Only 5 works are directly related to the ROA application in sustainable development.

First of all, the environmental issues are relatively recently. After then, there are many difficulties in modeling. There are not many data to assess the stochastic parameters, the process of learning and emergence of new information and new technologies is very dynamic. The analysis tools based on ROA is much more complex than the traditional strategies and sometimes is difficult for the investor understanding. The stochastic behavior is difficult to modeling and often involves jump diffusions and optimal stopping problems solution.

4. Conclusion.

Based on the literature review and on the intrinsic features associated to the evaluation of environmental projects, this paper suggests that the traditional cost-benefits analysis often employed in these tasks have some misleading. The main difficulties presented in these problems, namely, total or partially irreversibility over investment, cost and benefit, uncertainty over the future costs and benefit and managerial flexibility to change the choices in function of market and technological changes are the same as those in which the ROA is based on. Therefore, Real Option can represent a potential tool to support the decision making in environmental projects, specifically related to clean technologies. The few works that present ROA application in this kind of problem attest the complexity of the modeling considering all the tools provided by this method.

References

K. J. Arrow and A.C. Fisher (1974) Environmental Preservation, Uncertainty, and Irreversibility. The Quartely Journal of Economics, Vol. 88, No. 2 p. 312-319.

T. Bekefi and M. Epstein (2007) The whole picture CMA Management; 80, 8 pg.23-26.

T.E. Copeland and V. Antikarov (2001) Opções Reais: um novo paradigma para reinventar a avaliação de investimentos. Tradução de Maria José Cyhlar. – Rio de Janeiro: Campus.

A.K. Dixit and R.S. pindyck (1994) Investment Under Uncertainty. Princeton, New Jersey: Princeton University Press.

A. K. Dixit and R.S. Pindyck (1995) The Options Approach to Capital Investment. Harvard Business Review, May.

G. Dotsis, V. Makropoulou and D. Psychogios (2005) The Timing of Environmental Policies in the Presence of Extreme Events. Real Options – Theory meets practice - 9th Annual International Conference. Paris – France, June 22-25. Available: http://www.realoptions.org/abstracts/abstracts05.html.Accessed in 29/10/2007.

R.D. Espinoza and L.X. Luccioni (2007) An approximate solution for perpetual American option with time to build: the value of environmental remediation investment projects. International Journal of Business. Volume 12 number 3 summer. Available: http://www.craig.csufresno.edu/IJB/Volumes.htm. Accessed in 24/08/2007.

A.C. Fisher and W.M. Hanemann (1987) Quasi-option Value: Some misconceptions Dispelled, Journal of Environmental Economics and Management, 22, pp. 197-204.

W.M. Hanemann. (1989) Information and the Concept of Option Value, Journal of Enviromental Economics and Management 16, pp. 23-37.

S. L. Hart e M. B. Milstein (2004) Criando Valor sustentável. Revista de Administração de Empresas – FGV, Vol 3 n° 2 maio/jul.

C. Henry (1974) Investment Decisions Under Uncertainty: The "Irreversibility Effect". The American Economic Review, Vol 64. No. 6 Dec. Pp. 1006.1012.

A. Kiperstok (2008) O papel da Universidade e da Rede Teclim na Introdução de Práticas de Produção Limpa na Bahia. In KIPRSTOK, A. (org.) Prata da Casa: construindo produção limpa na Bahia. – Salvador, Rede de Tecnologias Limpas da Bahia – Teclim/Universidade Federal da Bahia.

T. Lin, C. Ko, H. Yeh. (2007) Applying real options in investment decisions relating to environmental pollution. Energy Policy 35 p. 2126 – 2432.

G. M. Martins & A.S.S.A.Melo (2007) O Valor Da Preservação Do Parque Dos Manguezais Em Recife-Pe: Uma Utilização Do Método De Opções Reais, Anais do XXXV Encontro Nacional de Economia [Proceedings of the 35th Brazilian Economics Meeting] 146, ANPEC - Associação Nacional dos Centros de Pósgraduação em Economia [Brazilian Association of Graduate Programs in Economics]. Avalilable: http://ideas.repec.org/p/anp/en2007/146.html. Accessed in 05/01/2009.

R. Morck, E. Schwartz and D.Stangeland (1989) The valuation of Forestry Resources under Stochastic Prices and Inventories. Journal of Finance and Quantitative analysis. Vol. 24, N° 4, PP.473-487.

R. S. Pindyck (2000) Irreversibilities and the timing of environmental policy. Resource and Energy Economics. Vol. 22 PP. 233-259.

_____ (2002) Optimal timing problems in environmental economics. Journal of Economic Dynamics & Control. Vol. 26 PP. 1677-1697.

_____(2007) Uncertainty in Environmental Economics. Review of Environmental Economics and Policy, volume 1, issue 1, PP. 45-63, winter.

J. Sachs (2008) Common Wealth: Economics for a Crowded Planet. Great Britain, Allen Lane.

J. Sarkis and M. Tamarkin (2005) Real Options Analysis for "Green Trading": The Case of Greenhouse Gases. The Engineering Economist, 50 p. 273 – 294.

M. E.B. Seiffert (2007) Gestão Ambiental: instrumentos, esferas de ação e educação ambiental – São Paulo: Atlas.

D.R. Siegel, J.L. Smith, J.L. Paddock. (1987) Valuing Offshore Properties With Option Pricing Models. Midland Corporate Finance Journal, Vol. 5, Spring. PP. 22-30.

L. Trigeorgis (1993) The nature of option interactions and the valuation of investments with multiple real options. Journal of Financial and Quantitative Analysis, p. 1-20, mar.

_____ (2007) Opções Reais e interações com a flexibilidade financeira. Revista de Administração de Empresas – FGV, Vol 47 n° 3 jul/set.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Crude Oil transshipment using Floating, Storage, and Offloading Platforms (FSOPs)

Sangeeta Balram and I.A. Karimi

Department of Chemical & Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117576

Abstract

Floating Storage and Offloading Platforms (FSOPs) have been around for forty years now and are attracting attention due to increasing offshore fields and facilities for oil and gas. Although in a conceptual stage, this technology can enable more efficient crude delivery to various demand centers onshore. The concept of FSOP has been prevalent for some time now in terms of its design and sustenance offshore. Although, the use of FSOPs for crude transshipment has been existential in the logistics industry, yet no work has so far addressed the management of such units for crude oil handling and transportation. To address this research gap, we combine a mix of techniques used by Cheng and Karimi (2006) as well as Reddy and Karimi (2004) specifically aiming at building continuous time models.

Keywords: FSOP, MILP, VLCCs/ ULCCs

1. Introduction

World crude oil demand is expected to increase to 90 million barrels per day by 2010, a growth of around 43% in a span of 20 years (1990-2010). In order to meet this heavy demand, world offshore oil production has risen three times to 53 million barrels per day in a span of 10 years whereas the world offshore gas production has more than doubled during the same span (PennWell Corporation, 2009). Offshore oil production now accounts for one third of the total world production and is growing at a much faster rate than the onshore sector. Over the past many years, Oil and Gas industry has thrived heavily on offshore drilling and storage units for ease of transportation of crude oil onshore. A technology, which has already been proposed in industrial literature, is the use of Floating Production & Storage Platforms (FPSOP) offshore. Although in a conceptual stage, this technology can enable more efficient crude delivery to various demand centers onshore. Crude oil transportation and delivery is a major trade activity in the oil and gas industry, largely done using Very Large Crude Carriers (VLCCs) or Ultra Large Crude Carriers (ULCCs). So far, crude oil transshipment by tanker lightering has been in practice in many fields. This practice is however, limited by the inability of VLCCs or ULCCs in entering ports with shallow draughts. The FPSOP can not only act as an intermediary between the oil market and refineries, in coordinating oil trade between the two, but can also act as an intermediate storage unit for crude parcels during transshipment, reducing onshore inventory levels. FPSOPs especially facilitate crude storage and transportation in remote or deep waters where pipeline infrastructure is not cost effective. Pipeline infrastructure in many fields may increase disproportionately with water depths. In smaller or marginal fields, oil and gas may be

produced for a relatively short duration of time. Laying fixed bed platforms or pipelines can prove extremely capital intensive in such fields. FPSOPs are considered more economical systems for oil and gas drilling and storage especially in the marginal fields. Ierapetritou (2003) divided their continuous time formulation into a number of subproblems, the first comprising crude oil operations, second dealing with refinery processes and intermediate tanks, and the third related to finished products and blending operations. Jetlund and Karimi (2003) worked on efficient routing and scheduling of chemical tankers for improving logistics in global chemical supply chains. They developed MILP formulation for a single ship and extended it to maximize profit for a fleet of multi-parcel tankers engaged in shipping liquid bulk chemicals. Reddy and Karimi (2004) built discrete and continuous time models for optimizing refinery operations while taking different tanks as intermediate storage units before transferring crude oil to several Crude Distillation Units (CDUs).

2. Problem Definition

In this work, we consider our FPSOP as an FSOP owned and managed by a single centralized authority. The system consists of several vessels like VLCCs/ULCCs (v = 1, 2, …, V) connecting through an SBM pipeline to the FSOP. A number of small Service Ships (n = 1, 2, …, N) (SSs, Cheng & Karimi, 2007a, b) or pipelines are employed for offshore delivery to a set of refineries (r = 1, 2, …, R). The FSOP consisting of tanks (s = 1, 2, …, S) of different capacities acts as a central procurement unit between the oil market and refineries making the main purchasing decision of crude from the oil market. The refineries work in collaboration with each other for mutual benefit (Collaborative Logistics). The FSOP acts as a trans-shipment hub for crude supply to the refineries based on orders (o = 1, 2, …, O) from the refineries and charters out vessels, bringing a variety of crude parcels (p = 1, 2, …, P) into the tanks. We assume that each vessel scheduled to supply crude to FSOP has an estimated Earliest Docking Time (EDT_v). Furthermore, since only one vessel can dock at any time, we assume that no two EDTs are identical. Thus, we can sequence the EDTs in the chronologically increasing order. A vessel is stipulated to depart anytime at or before its Stipulated Time of Departure (STD_v). The tanks serve as storage, blending and offloading units. The refineries' orders contain multiple crude products (c = 1, 2, …, C) of fixed composition according to refineries' specifications. Each order is related to a delivery time window $[DD_o^L, DD_o^U]$. Refineries most commonly use 15 crude properties m which form a basis for specifying the crude quality (Li et al., 2007). These blending indices are either volume based or weight based.

Given:

- A scheduling horizon $[0, H]$.
- V vessels, their estimated arrival and departure times.
- P parcels, their total volumes and flow rates.
- S tanks, their capacities, outgoing flow rates, initial inventories and limits on their hold ups.
- C crudes, their property indices and specific limits on their property indices.
- O orders, their constituent crudes and required amounts.
- Economic data such as demurrage or delay costs (vessels and orders), inventory costs, changeover costs, set up costs (tanks).

Determine:
- A detailed schedule for parcel unloading from vessels and their amounts.
- A detailed schedule for order delivery from tanks and their amounts.

Operating policies:
- Owing to SBM pipeline, only one vessel can dock at any time and atmost one parcel can be unloaded from it at a time.
- A tank cannot receive and deliver simultaneously.

Assuming:
- There is no hold-up in the SBM pipeline connecting the vessel to the FSOP.
- Changeover times and set up times in tanks are negligible.
- All costs are constants and any price volatility with respect to costs is ignored.
- FSOP is physically stable and any stability fluctuations from turbulence in weather conditions are ignored.

3. Mathematical Formulation

Each vessel may carry multiple single-crude parcels. Since the unloading sequence of parcels from each vessel is usually pre-fixed (Reddy et al., 2004), and it takes non-zero time to transfer each parcel, we can also define a fixed sequence for unloading all crude parcels from the vessels. We now define an Earliest Time of Arrival (ETA_p) for each parcel p as the unique time at which the parcel may become available for transfer. We compute $ETA_p = EDT_v + \sum_{p=2}^{P} \tau_{p-1}$ for all $p \in \boldsymbol{P}_v$ where $\boldsymbol{P}_v$ = {p | vessel v carries a parcel p of some specific crude}. Note that, the Earliest Time of Arrival (ETA_p) of any parcel p is different from its actual time of transfer, which is decided by the optimizer based on the unloading sequence of parcels. Time representation is of prime importance in any scheduling problem. We propose an asynchronous slot based approach to model time on the SBM pipeline and tank s independently in terms of a series of K chronologically ordered contiguous slots (k = 0, 1, .., K) of variable lengths. Slot k (k = 1, 2, .., K) on the SBM pipeline starts at $T_{(k-1)}$, ends at T_k and has a length of $[T_k - T_{(k-1)}]$. Similarly, slot k on tank s starts at $T_{s(k-1)}$, ends at T_{sk} and has a length of $[T_{sk} - T_{s(k-1)}]$. Denoting the start of scheduling horizon as the start of slot 1, we have $T_0 = T_{s0} = 0$. By definition;

$$T_k \geq T_{(k-1)} \qquad T_k \leq H,\ 1 \leq k \leq K \qquad \text{(1a)}$$

$$T_{sk} \geq T_{s(k-1)} \qquad T_{sk} \leq H,\ 1 \leq k \leq K \qquad \text{(1b)}$$

We assume that a real operation starts only at the beginning of a slot and continues during the slot; however it may end at any time within the slot. Thus, the idle time if any, is always towards the end of the slot (Susarla et al, 2008). While parcel transfer is the only operation on the SBM pipeline; tank operations include parcel receipt, order delivery, crude storage, crude changeovers, brine settling as well as crude blending.

3.1 Parcel Transfer Operations

We model the parcel transfer operation on the slots of the SBM pipeline by defining one binary variable x_{pk}.

$$x_{pk} = \begin{cases} 1 & \text{if parcel } p \text{ is transferring during slot } k \\ 0 & \text{otherwise} \end{cases} \qquad 1 \leq p \leq NP,\ 1 \leq k \leq K$$

As noted earlier, atmost one parcel can transfer through the SBM pipeline during a slot. Therefore,

$$\sum_{p} x_{pk} \le 1 \qquad 1 \le p \le NP,\ 1 \le k \le K \qquad (2)$$

3.2 Parcel receipt into tanks

Now, if a parcel transfers during a slot of the SBM pipeline, it must also be received during a slot on its associated tank. Let us suppose that the parcel transferring in slot *k* of SBM pipeline is received by tank *s* during its slot k'. Now, we can have three possible scenarios such that: $k' < k$, $k' = k$, and $k' > k$. For the first and third scenario, we can simply add extra slots on the tank and SBM pipeline respectively, to make $k' = k$ (Susarla et. al, 2008). Therefore, we demand that if the SBM pipeline transfers a parcel into any tank *s*, then the slots corresponding to the parcel transfer on the SBM pipeline and parcel receipt on tank *s* should have the same index. With this explanation, we now denote parcel receipt into tank *s* by binary variable u_{sk} as;

$$u_{sk} = \begin{cases} 1 & \text{if tank } s \text{ is in a state of receipt during slot } k \\ 0 & \text{otherwise} \end{cases} \qquad 1 \le s \le S,\ 1 \le k \le K$$

We allow exactly one tank to receive a parcel during a slot. This means;

$$\sum_{s} u_{sk} = \sum_{p} x_{pk} \qquad 1 \le k \le K \qquad (3)$$

However, equation (6) does not indicate the transfer of a particular parcel *p* into its destination tank *s* during slot *k*. For this, we introduce a 0-1 continuous variable $xu_{psk} = x_{pk} u_{sk}$;

$$xu_{psk} = \begin{cases} 1 & \text{if parcel } p \text{ transfers into tank } s \text{ during slot } k \\ 0 & \text{otherwise} \end{cases} \qquad s \in \boldsymbol{S}_p,\ 1 \le k \le K$$

where $\boldsymbol{S}_p = \{s \mid \text{tank } s \text{ can receive parcel } p\}$

We use McCormick's linearization method to linearize xu_{psk}.

3.3 Material balance for parcels

Let volume of parcel *p* transferred into tank *s* during slot *k* be QI_{psk}. Total amount of each parcel is Q_p. Hence,

$$QI_{psk} \le Q_p xu_{psk} \qquad s \in \boldsymbol{S}_p,\ 1 \le k \le K \qquad (4a)$$

To ensure that a parcel transfers fully within the scheduling horizon;

$$\sum_{k} \sum_{s \in S_p} QI_{psk} = Q_p \qquad 1 \le k \le K \qquad (4b)$$

3.4 Timing constraints

Transfer of any parcel *p* through the SBM pipeline cannot start before its Earliest Time of Arrival (ETA_p). This means;

$$T_{k-1} \ge \sum_{p} ETA_p xs_{pk} \qquad 1 \le k \le K \qquad (5)$$

Now, the actual amount of parcel transferred into a tank during any slot is limited by its maximum pumping flow rate,

$$\sum_{s \in S_p} QI_{psk} \le FP_p^U (T_k - T_{k-1}) \qquad 1 \le k \le K \qquad (6)$$

Recall that each vessel has a Stipulated Departure Time (STD_v). If the last parcel from each vessel remains connected to the SBM pipeline beyond this time; demurrage penalty is imposed. Let d_v denote the delay in the departure of vessel *v*.

$$d_v \ge T_k - STD_v - H(1 - x_{pk}) \qquad p \in \boldsymbol{LP}_v,\ 1 \le k \le K \qquad (7)$$

where $\boldsymbol{LP}_v$ = {p | parcel p is the last parcel from vessel v}
Since a parcel transfers into one of its destination tanks at a time; we need to register the parcel transfer times on tanks as well. For this; we synchronize the start and end times of parcel transfer on SBM pipeline with that on the tank

3.2 Order Delivery Operations

Each tank can deliver several single-crude product orders within the scheduling horizon. We denote delivery of order o during slot k on any tank by a binary variable y_{ok} as follows;

$$y_{ok} = \begin{cases} 1 & \text{if order } o \text{ delivers during slot } k \\ 0 & \text{otherwise} \end{cases} \qquad o \in \boldsymbol{O}_s,\ 1 \le k \le K$$

where $\boldsymbol{O}_s$ = {o | tank s can deliver order o}
Let QO_{sok} be the volume of order o delivered by tank s during slot k. Total required quantity of each order is RQ_o. Now, if a tank does not deliver an order during slot k, then its quantity is zero. Hence,

$$QO_{sok} \le RQ_o y_{ok} \qquad o \in \boldsymbol{O}_s,\ 1 \le k \le K \qquad (8)$$

To ensure full delivery of each order o;

$$\sum_k \sum_s QO_{sok} = RQ_o \qquad o \in \boldsymbol{O}_s,\ 1 \le k \le K \qquad (9)$$

Amount of each order delivered from a tank is limited by maximum delivery rate FS_s^U of each tank such that,

$$\sum_{o \in Os} QO_{sok} \le FS_s^U (T_{sk} - T_{s(k-1)}) \qquad o \in \boldsymbol{O}_s,\ 1 \le k \le K \qquad (10)$$

Each order has a certain delivery time window as mentioned earlier $[DD_o^L, DD_o^U]$. This means a tank s cannot start delivering an order o before DD_o^L;

$$T_{s(k-1)} \ge DD_o^L y_{ok} \qquad o \in \boldsymbol{O}_s,\ 1 \le k \le K \qquad (11)$$

Also, if a tank delivers an order beyond DD_o^U, penalty is imposed under order delivery delay. We consider d_o as the delivery delay for each order.

$$d_o \ge T_{sk} - DD_o^U - H(1 - y_{ok}) \qquad o \in \boldsymbol{O}_s,\ 1 \le k \le K \qquad (12)$$

3.4 Inventory Balance in Tanks

Let I_{sk} denote the current inventory level of tank s during slot k. The initial inventory level of tanks s before time $T_{s0} = 0$ is known to be I_{s0}. Thus, inventory balance in any tank s is given as:

$$I_{sk} = I_{s(k-1)} + \sum_p QI_{psk} - \sum_o QO_{sok} \qquad s \in \boldsymbol{S}_p,\ o \in \boldsymbol{O}_s,\ 1 \le k \le K \qquad (13)$$

Also, the total inventory I_{sk} in tank s during any slot cannot exceed its maximum holding capacity Cap_s^U.

$$I_{sk} \le Cap_s^U \qquad 1 \le s \le S,\ 1 \le k \le K \qquad (14)$$

3.1. Objective

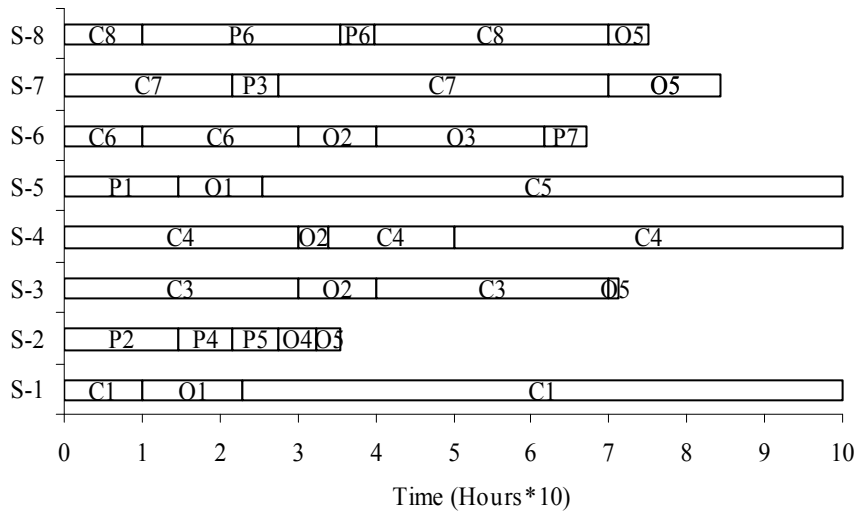
Minimizing Total Costs of FSOP is our scheduling objective. The costs include demurrage costs of VLCCs/ULCCs per hour given by DC_v; costs associated with order delivery delay DCO_o, inventory costs of FSOP-tanks per hour IC_{uk}. Here, d_v represents delay beyond stipulated time of departure of VLCC/ULCC v at the FSOP, d_o represents delivery delay of order o, and $I_{sk=K}$ represents inventory level of tank s during the last slot K.

$$Total\ Costs = \sum_{v} DC_v.d_v + \sum_{k=K} \sum_{s} ICs.I_{sk} + \sum_{o} DCO_o.d_o \quad (15)$$

4. Case Study

We apply our model to 3 VLCC/ULCCs carrying 3, 2 and 2 parcels and arriving at 5, 10 and 15 hours respectively. The FSOP comprises 8 different tanks which deliver 5 possible orders to refineries during the scheduling horizon of 100 hours. We summarize the key aspects of our solution in Fig 1.The solutions were obtained using CPLEX (10) in GAMS (Distribution 22.2) running on a Windows XP PC with AMD Ethlon 64 X2 Dual (3 GHz) processor and with relative optimality gap limit (OPTCR) of 0.0. The model involved 1,965 constraints, 177 binary variables and 7,461 nonzero elements. CPLEX solved the model in 137.95 seconds and gave Total Costs of $10.01.

Fig 1. Key results of Solutions



5. Conclusions

This paper introduces a novel concept of crude transshipment using FSOP. To the best of our knowledge, so far no literature exists on the management of such units. The use of asynchronous slot based formulation for crude transshipment is another new aspect it touches. Future work includes making our model more realistic by allowing crude changeovers within tanks.

References

H. Cheng and I.A. Karimi, 2006, Scheduling Transshipment Operations in Maritime Chemical Transportation, Ind. Eng. Chem. Res., 45, 1955-1973.

P.C. Reddy, I.A. Karimi, R. Srinivasan, 2004, A new continuous time formulation for scheduling crude oil operations, Chemical Engineering Science, 59, 1325-1341.

N. Susarla, J. Li, I.A. Karimi, 2008, A Novel Approach to Scheduling Multipurpose Batch Plants Using Unit Slots, Submitted to AIChE Journal.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

An ontological framework to support the implementation of waste minimisation strategies

Adriana P. Reyes-Córdoba[a] , Jorge A. Arizmendi-Sánchez[b]

[a] *School of Chemical Engineering and Analytical Science, The University of Manchester P.O. Box 88, M60 1QD, Manchester, UK.*
[b] *Advantica Ltd. Holywell Park, New Ashby Road, LE11 3BW, Loughborough, England*

Abstract

Modern understanding about waste management has provided a general hierarchy to deal with waste, with a priority in the reduction or minimisation of the amount of waste generated. Different methodologies have been devised to aid the process of implementing waste minimisation in the process industries. However, waste minimisation methodologies are still underestimated or disregarded due to several misconceptions about their applicability. One reason believed to cause this is the lack of an accurate way to identify suitable waste reduction techniques. This arises from a mismatch between the existing methodologies and the information available at the time of the study. An appropriate approach to overcome this situation must consider a more efficient manipulation and analysis of the available information, optimizing the cost and time required to obtain the critical missing data. The structured organisation of the knowledge about a process regarding its description, environmental concerns and available resources provides an effective way to identify the areas which need attention for the purpose of waste minimisation and facilitates the definition of activities that have to be performed to implement the devised solutions.

Keywords: waste minimisation, waste minimisation methodologies, knowledge management, ontological framework

1. Introduction

In process industries, waste refers to any material that cannot be further used in the production of goods. Such material has finished its usable life and has to be disposed properly. Waste minimisation strategies aim at three main goals. The most important of which is to reduce and, when possible, to prevent the generation of waste. However, generation of waste is not always avoidable; thus, the second goal is to improve the quality of any waste generated. This includes reducing the hazard of any waste material so that it can be safely disposed. Additionally, it is important to encourage the reuse, recycling and recovery of any waste material that is generated. The Environment Agency for England and Wales (EA) advises the efficient use of raw materials, energy and water as a way to achieve waste minimisation. In the European Union, it is also considered that processes and activities must be systematically understood and thus changed to prevent and reduce waste.

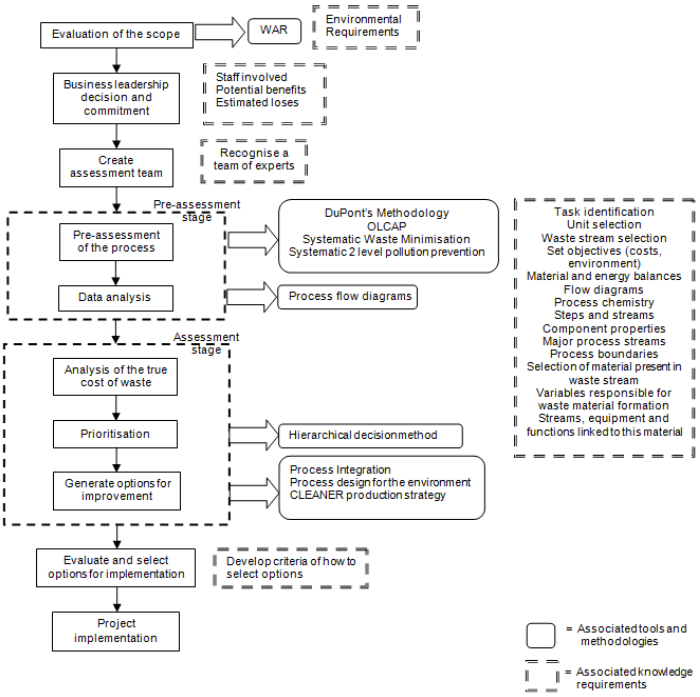


Figure-1 Main strategy for waste minimisation with associated knowledge requirements methodologies and tools

Several efforts have been made to tackle the complexity of the task that waste minimisation represents. These efforts have been in the form of technical recommendations for operative plants or for the design phase of new plants, however, a challenge is presented in the integration of all the information needed for these approaches to be practical. Figure-1 provides an overview of the general strategy for waste minimisation given by the EA. This overview includes a collection of predominant methodologies studied by different authors [1-4]. These have been integrated into the stage of the EA strategy in which they are most convenient. This figure shows the main concerns about the knowledge required for the effective performance of this strategy. The successful application of these approaches relies on the appropriate identification, gathering and classification of the relevant information required in each case.

Attempts to devise an integrated and optimal strategy to conduct waste minimisation studies usually include the development of automated systems. However, as waste minimisation problems are complex and multidisciplinary, efforts might end up not considering relevant information or overlapping it. This leads to problems like information being presented in incompatible terminology or being structured according to different patterns (i.e. databases, spreadsheets, etc). Knowledge management tools have been recognised as the most practical ways to identify, organize and share this kind of information. Various authors identify the suitability of ontologies for this purpose [5, 6]. Current ontologies in the field of chemical engineering focus on a broad domain and a challenge arises when trying to implement this into waste minimisation efforts. Knowledge is one of the most valuable assets in the effective implementation of waste minimisation programs. Most of the knowledge within a company is known

implicitly to the participating employees. These employees turn into experts on a specific function and become indispensable unless their knowledge is documented. To fully exploit this knowledge it is necessary to make it explicit and shared among the different participants of a specific process [5]. By definition, knowledge management tools are the adequate way to shape knowledge into an explicit form that can be used and shared within an organisation.

2. Knowledge management approach for waste minimisation

2.1. Identification of required domains

The approach presented in this contribution proposes a methodological framework based on ontological principles aimed to support the generation and implementation of waste minimisation strategies. Knowledge is classified in four different domains embedded in an ontological structure: process description, environmental requirements, available documentation and waste minimisation options. This approach enables the classification and retrieval of available information in a more efficient way, providing a more structured framework to support the implementation of waste minimisation strategies. The proposed ontological structure aims to provide answers to these questions: which environmental requirements are not being complied, what kind of information is needed to work towards compliance, how this information is obtained, what kinds of waste minimisation options are available and which knowledge resources are needed and available to implement them. The required domains are constructed by identifying and collecting expert knowledge. Following the general recommendations for ontology construction [7], the process knowledge domain includes a fraction of an existing ontology (OntoCAPE [8]) as the keystone on which further knowledge can be added. The ontological framework proposed bases the process description on the conceptual layer presented in the ontology as the module named *chemical process system*. The main focus is on the actual plant model and the *process*, *process step* and *process units* on which the plant depends. The legislation domain takes into account the conditions set by the IPPC directive for the European Union to promote the use of Best Available Techniques (BAT). These techniques provide the main issues to address for specific sectors within the process industry. The third domain classifies waste minimisation options and methodologies currently in use. These are drawn from the BAT reference documents and from the existing strategies mentioned previously. The knowledge resources domain is constructed by considering the experts required to conduct the waste minimisation strategies which constitute the explicit and tacit forms of knowledge available for an organisation. Additionally, explicit forms of knowledge are taken into account and divided into different tools such as documented information, specialized software, calculations, and external resources.

2.2. Development of the domains

The development of these domains is performed following the same principles of ontological construction. This approach was chosen due to the need for an integrated description of these domains which can be achieved by identifying their key concepts and relationships in the same manner as an ontology. The first stage is to select the scope of the domain and analyze its requirements. This is done by the use of guidance questions that involve queries which the domain should be able to answer. This has the intention of identifying the major requirements, purpose and scope for the domain. The second stage involves the collection of reusable resources such as existing ontological domains and conceptual data models relevant to the scope and to identify parts of

relevance that could be reused The third stage requires building the structure of the domain. This involves identifying important terms that describe the objects and concepts found in the domain, grouping these terms into classes and setting them to a hierarchy with defined relationships. The final purpose is to communicate the conceptualization of the four domains and to provide the backbone for the development of an ontology based on these domains.

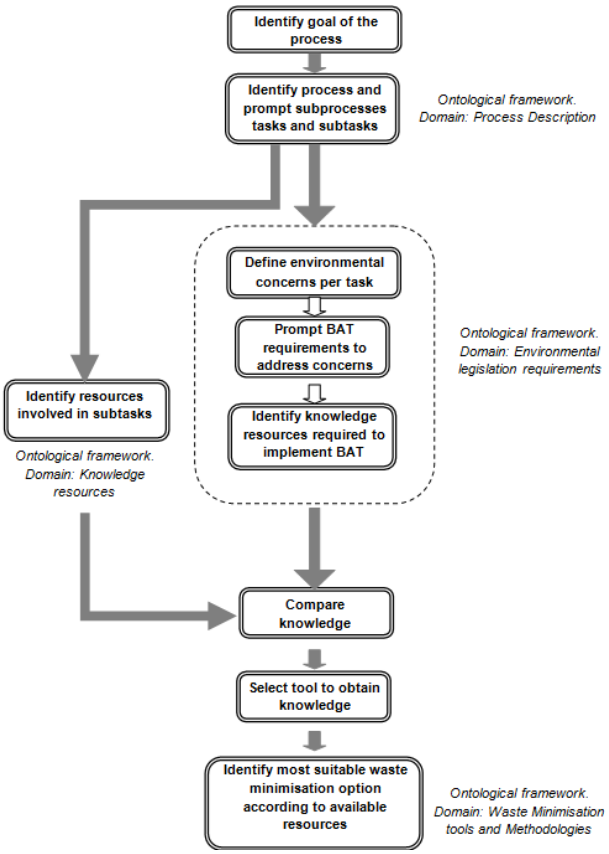


Figure-2 General methodology for the implementation of an ontological framework to aid in waste minimisation

3. Methodology to assess the suitability of waste minimisation actions

The methodology, presented in

Figure-2, starts with the identification of the goal of the process under study. This refers to the crucial knowledge element which dictates what the final purpose of the whole process is. According to this goal it is possible to identify to which industry the process belongs to and which subtasks are involved in it. Once this is done, the environmental requirements for each of the previously identified subprocesses are declared. These are drawn from the environmental legislation domain. Compliance of these environmental requirements is achieved by planning specific actions which use specific information resources. The next step in the methodology is to assess if these information resources

are present within the organisation. To achieve this, the knowledge resources domain is analysed. This domain lists the different information resources that should be available to the company in order to match them with the ones available in reality. In this way the missing knowledge is identified along with the effective sources to find it. Finally, once this knowledge has been gathered it is possible to suggest a waste minimisation option that is practical with the available resources.

4. Use of the framework

The framework has been tested on the case study of the reaction step of phthalic anhydride production. The previously discussed domains provide the framework to study the process as shown in Figure 3.

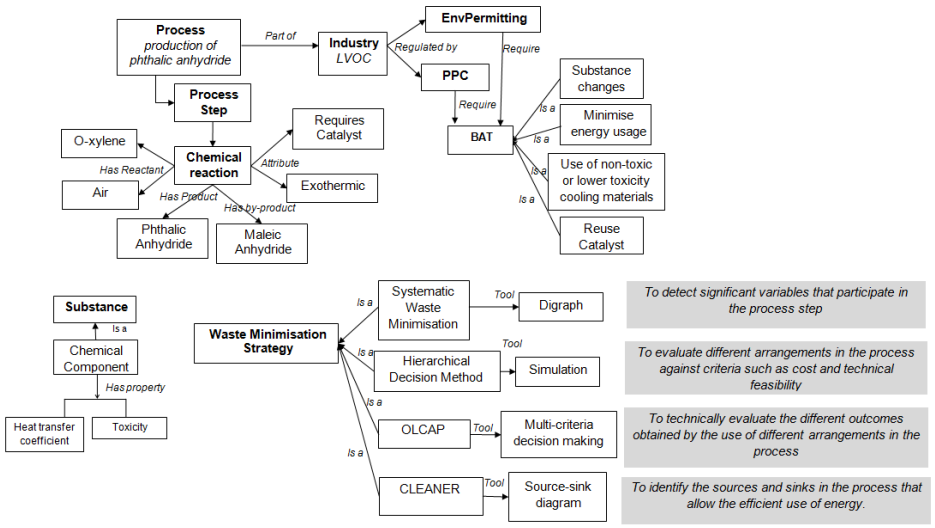


Figure 3 Study of the process using the proposed domains

The generic process steps defined by the process description domain include reaction, separation, and enthalpy change, amongst others. The focus in this case is the reaction process step. The reaction is defined by specific attributes which in this case include the need for a catalyst and the exothermic nature. The reaction also involves the use of specific materials namely o-xylene, air, and a catalyst. The overall process is classified in the large volume organic chemicals (LVOC) industry which is regulated by the Environmental Permitting regulations and the IPPC. Under these regulations the use of best available techniques (BAT) is required. Generic BAT recommendations for the LVOC industry include: minimization of energy usage, regeneration or reuse of catalysts, use of non-toxic or lower toxicity cooling materials amongst others. The exothermic nature of the reaction allows for the first BAT recommendation to be explored. The reaction can be analysed using the identification of process variables and their functionality studied in the waste minimisation strategies that form the waste minimisation domain. The variables identified for the reactor are temperature, pressure and composition. The need to control Temperature is evident and thus the first heuristic rule arises which identifies the action of adding a cooling agent as a solution. The cooling agent has to be evaluated against constraints such as cost, environmental impact

and technical feasibility as determined in the waste minimisation domain. Three different materials are available in this case, mercury, sodium potassium alloy and a heat transfer salt mixture. The material evaluation gives the third material as the best compromise solution between costs, technical feasibility and environmental impact. Furthermore the elaboration of source-sink diagrams (as suggested in the waste minimisation strategies) allows the identification of possible ways of using the heat gained by the salt as energy source for heating process water required for another stage of the process. Knowledge resources required for the evaluation include specialized simulation software, databases with material characteristics (including toxicity, heat transfer coefficient and costs), energy balances and personnel with experience in this kind of studies.

5. Discussion

Organizing the knowledge about a process regarding its description, environmental concerns and available information resources provides an effective way to identify the areas which need attention for the purpose of waste minimisation. This facilitates the process of defining the activities that have to be performed to implement the devised solutions. The assessment and pre-assessment stages of waste minimisation studies are crucial for the successful achievement of the goals set. These stages require vast amounts of information to be available and accurate. They also require multidisciplinary teams with experts and skilled staff from all the areas involved. These resources must be adequately managed in order for them to work towards the final goal of waste minimisation. Effective organisation of available information and knowledge is the key for the proper application of waste minimisation methodologies. Integration of waste minimisation into the field of knowledge management through the use of an ontological framework has been proposed. It has been recognised that ontologies can be used to organize knowledge and therefore tackle different areas without disregarding relevant issues. Waste minimisation methodologies need a tool that considers the complexity of the knowledge that has to be gathered for their correct performance. A systematic approach to organize this knowledge will therefore result in a more practical way to implement current waste minimization methodologies.

References

1. Azapagic, A. and R. Clift, *The application of life cycle assessment to process optimisation.* Computers and Chemical Engineering, 1999. **23**: p. 1509-1526.
2. Halim, I. and R. Srinivasan, *Systematic Waste Minimization in Chemical Processes. 1. Methodology.* Industrial and Engineering Chemical Research, 2002(41): p. 196-207.
3. Mulholland, K. and J. Dyer, *Follow this path to pollution prevention.* Chemical Engineering Progress, 1998. **94**(1): p. 34-42.
4. Dunn, R. and G. Bush, *Using process integration technology for CLEANER production.* Journal of Cleaner Production, 2000. **9**: p. 1-23.
5. Brandt, S.C., et al., *An ontology-based approach to knowledge management in design processes.* Computers and Chemical Engineering, 2008. **32**: p. 320-342.
6. Zhao, C., et al., *Toward intelligent decision support for pharmaceutical product development.* Journal of Pharmaceutical Innovation, 2006: p. 23-35.
7. Gómez-Pérez, A., M. Fernández-López, and O. Corcho, *Ontological Engineering*. 1st ed, ed. X. Wu and L. Jain. 2004: Springer-Verlag London Limited 2004. 403.
8. Morbach, J., A. Yang, and W. Marquardt, *OntoCAPE- A large-scale ontology for chemical process engineering.* Engineering applications of artificial intelligence, 2007. **20**: p. 147-161.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Abnormal Situation Management in a Refinery Supply Chain Supported by an Agent-based Simulation Model

Koen H. van Dam[a], Zofia Lukszo[a], Rajagopalan Srinivasan[b]

[a] *Department of Energy and Industry, Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, the Netherlands*
[b] *Department of Chemical and Biomolecular Engineering, National University of Singapore, 4 Engineering Drive 4, Singapore 117576, Singapore*

Abstract

Oil refineries are of high importance for global economic health and energy supply; any disruptions to their operations may have major worldwide impact. In this paper the application of an agent-based refinery supply chain model to abnormal situation management is described. Agents represent the various decision makers in the supply chain. They own, operate and manage the elements of the physical network of the supply chain. A disruption in ship arrival is used to illustrate the applicability of the decision support system. The decision support system derives a suitable course of action for a given situation based on the outcomes of a number of simulation runs according to the Nelder-Mead zero-order optimization method. This method is based on identification of the best, the worst, and the second worst outcomes in each iteration for the pre-defined experiment. It can be concluded that the decision support system can interact with multiple actors in the supply chain to diagnose and compensate for unanticipated disruptions, with a substantial impact on refinery productivity.

Keywords: refinery supply chain, agent-based model, decision support, disruption, Nelder-Mead

1. Introduction

Abnormal situations in supply chains (SCs) encompass a range of events outside the "normal" operating modes, including human error, fires, delays in ship arrival, (unplanned) maintenance and equipment failure. As a consequence, planned production targets may not be reached, unless swift response is taken to minimise the negative effects. Which response is the most efficient one, however, is not easy to determine and requires decision support tools. This paper describes how a decision support tool using an agent-based simulation model can help an oil refinery in dealing with disturbances and to ensure smooth operation at minimal costs.

Oil refineries are of high importance for global economic health and energy supply; any disruption to their operations may have major local and worldwide impact. A typical refinery SC comprises crude oil suppliers, 3rd party logistics providers, shippers, jetty operators, the refinery and customers. The refinery occupies a pivotal position in the SC with its functional departments initiating and controlling the interactions with the external entities for the various SC activities, such as procurement, storage, logistics and operations. A maze of complex interactions between the different entities and resulting decisions ensure the orderly and efficient functioning of the supply chain as described in detail by Julka *et al.* (2002) and Pitty *et al.* (2008). The combined

performance determines the economics via crude costs, product prices, operation costs, transportation, etc.

The multi-actor, distributed, complex and dynamic nature of a supply chain can be best evaluated using simulation models. There is a strong need for models that can help decision makers in the process industry to analyse the risks of abnormal situations in the supply chain and to assess possible solutions. The agent-based paradigm, with agents as problem solving entities characterised by autonomy, reactiveness, pro-activeness, goal orientation and the ability to communicate with other agents and the environment (Jennings, 2000), is suitable for supply chain modelling: flexible models can be developed that are easy to extend and which can capture and visualise the complex social interactions between the supply chain units, modelled as agents (van Dam *et al.*, in print). In this paper it is demonstrated how an agent-based model can be used to support decision makers dealing with abnormal situations. While each one of the actors (i.e. stakeholders) can be considered as a main problem owner, here the perspective of the oil refinery is chosen. The decision support tool described here is designed for one such decision maker in an industrial setting.

This paper contributes to the area of the abnormal situation management of a refinery supply chain and to the application of agent-based models to industrial decision making. Section 2 describes a generic decision problem for a refinery supply chain. In section 3 the search algorithm supporting the decision problem is discussed, followed by a specific example in Section 4. Finally, Section 5 gives concluding remarks on the use of agent-based models for decision support.

2. Decision problem

This section describes a decision problem for an oil refinery supply chain. First, the disturbances are defined, followed by the degrees of freedom and the performance criterion.

2.1. Disturbances

Although the approach proposed in this work is general, in this paper the scope is limited to disturbances dealing with the *supply* of crude oil to the crude distillation units. A disruption in the supply can be caused by a delay in the shipment of crudes from the supplier (at a large distance from the refinery) or problems in the tank farm, for example. In both cases the operations department runs the risk of not having access to enough crude to perform the scheduled operations.

These disturbances are then defined as follows:

$\underline{d} = (ShipDelay, StorageProblem)$

$ShipDelay \in \{0, 1, 2, ..., n\}$	in days, for 1 ship for 1 cycle
$StorageProblem \in \{0, 1\}^m$	for each of the $m = 5$ crude storage tanks

For instance, $\underline{d} = (2, 0,0,0,0,0)$ indicates that a ship at sea is delayed by 2 days, and there are no problems with the storage installation. Note that n for a ship delay can be larger than the time horizon used, effectively "sinking" the ship. For simplicity, we assume that the magnitude of the disturbance is known as soon as the disturbance occurs. In reality, this may involve uncertainty. Furthermore, the granularity could be adjusted so that a delay could be expressed in parts of a day (e.g. hours) in stead of full days.

2.2. Degrees of freedom

When faced with such a disturbance, the problem owner has to make a number of choices. Firstly, he has to determine if the disturbance has a significant effect on the operation of the supply chain. If the effect is deemed minor, no action may be necessary, but when not able to execute the previously planned schedules due to inadequate crude, corrective action may be required. This can be addressed by changing the operating mode, the throughput or by emergency crude procurement. Often a combination of these actions may be needed.

For the Emergency Procurement *EmPr* the procurement department can contact a local supplier to buy crude at a much higher price but with a short lead time. The degree of freedom for each of the five crudes is between 0 kbbl to the amount that could reasonably available on short notice, which is assumed 600 kbbl here. The procurement department has to ask logistics department for the expected delay to be able to make this decision. Note that when there is an error in a storage installation (*StorageInstallationError=1)*, emergency procurement will not solve a disturbance, as all crudes have to rest in storage to allow the brine to settle and cannot be transferred directly from the vessel to the crude distillation unit but have to pass storage tanks.

Furthermore, the operations department can choose to Change the Operational Configuration (*COC*), meaning that a different recipe (one out of four) is selected using crudes that are still in stock, but resulting in yields that are not ideal compared to the scheduled operation. Finally, the operations department can Change the Operational Scale (*COS*), to run the refinery at a lower throughput (from 40% to 100% of the maximum capacity, with a minimum throughput to keep plant running), producing less end products but avoiding having to shutdown the plant when crude runs out (or postponing plant shutdown, for example to allow emergency procurement crudes to arrive).

The degrees of freedom are defined as follows:

$\underline{x} = (EmPr_1, EmPr_2, EmPr_3, EmPr_4, EmPr_5, COC, COS)$

$0 \le EmPr_i \le 600$ in kbbl, for each of the $i = 5$ crudes

$COC \in \{R1, R2, R3, R4\}$ discrete choice between operating modes

$40 \le COS \le 100$ percentage of CDU throughput capacity

2.3. Performance criterion

There are different options for the criteria with which to choose the best alternative. As examples, one can look at the overall profit of the refinery (for a certain time frame), profit during the production cycle effected by the disturbance, other financial measures, but also non-economical criteria such as customer satisfaction. Here we choose profit *P* of refinery 14 days after a disruption took place. This means that the effect of a disturbance at $t = 32$ will be simulated over the next two cycles of operation during which new raw material deliveries and product dispatch. We expect that the impact of the disturbance would have worn off by then.

The objective function is defined as follows:

$$\max P(\underline{x},\underline{d}) = \sum_{t=1}^{50} (Income^{t}_{sales}(\underline{x},\underline{d}) - Cost^{t}_{procurement}(\underline{x},\underline{d}) - Cost^{t}_{transp}(\underline{x},\underline{d}) - Cost^{t}_{maint}(\underline{x},\underline{d}))$$

$$+ Value^{t=50}_{producstocka}(\underline{x},\underline{d}) + Value^{t=50}_{rawmaterials}(\underline{x},\underline{d})$$

In this equation the value of the product inventories and raw materials at the end of the simulation run have been included. The consequences of the disruption on future cycles are included in the cost function (e.g. if the response is to switch to another mode of operation without any emergency procurement, it is possible that during a later cycle the planned operation cannot be met) but no new decisions following to any such new disturbances are assumed; a single response is formulated.

The transportation costs, and therefore profit, are discontinuous because they are a function of the amount of crude procured, the capacity of the ships (either a VLCC for long distance shipping or a general purpose tanker with much smaller capacity for short haul in the case of emergency procurement) and the travel time. The transport cost is calculated per vessel and the cost for procurement plus transport therefore follows a saw-tooth pattern, making it more difficult to determine the right amount to buy (especially in combination with other measures such as switching to another recipe).

3. Search method for decision support

The decision support system presented here derives the advice for a given situation based on the outcomes of a number of simulation runs according to the Nelder-Mead zero-order optimization method (Nelder & Mead, 1965). The term zero-order refers to the fact that the search of the optimum is carried out without calculating any derivatives of the performance criterion but directly by measuring (without help from a process model) or simulating the state of the system. Zero-order search is recommended when the process has one or more of the following properties (Wright, 1995):

- Process model is difficult or expensive to obtain;
- Process exhibits discontinuities;
- Measurement data are contaminated by significant noise.

The search with a Nelder–Mead simplex algorithm is applicable to solve decision problem formulated in this paper. The method is based on identification of the best, the worst, and the second worst outcomes in each iteration for the pre-defined simplex (a set of experiments). An initial simplex S is defined as a convex hull with $n + 1$ vertices $\{\underline{x}_j\}^{n+1}_{j=1}$ in an n-dimensional space $\mathbb{R}^n$ (in this case $n=7$). These vertices satisfy the non-degeneracy condition, meaning that the volume of the simplex hull is non-zero. For every next iteration $j+1$, the values for $\{\underline{x}_j\}^{n+1}_{j+1}$ are determined by comparing the objective-function values followed by replacement of the worst vertex by another point. The simplex adapts itself to the local landscape and contracts finally on the optimum.

Compared with traditional iterative search algorithms, the proposed examples in this paper show that the pattern search algorithm is robust, easily implemented and easy to communicate to decision makers. Moreover, the search algorithm is more likely to give a global minimum in dealing with nonlinear problems, non-linear constraints and multi-criteria problems (Verwater-Lukszo, 1996; Lu, 2008; Ouria, 2009; Ye 2008).

4. Abnormal situation management supported by an agent-based model

The system model of an oil refinery SC used here for simulation and experimentation to support the decision-making process is an agent-based model developed in Repast, as described in detail in van Dam *et al.*, (2008). Agents model all stakeholders, each with their own unique behavioural rules and own goals. They communicate with each other to make decisions about the trade of (raw) materials or the operation of the refinery, for example. The model has been benchmarked with an equation based model of the same

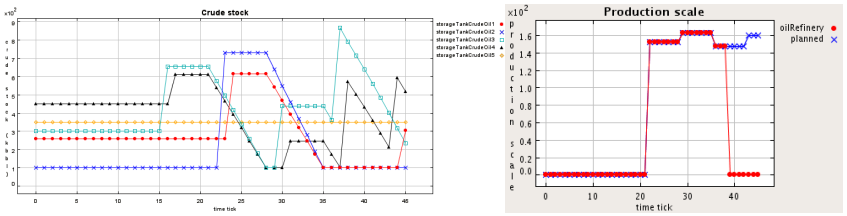


Figure 1. Crude stocks under normal operation (left) and the gap between planned and actual throughput after a disturbance causing a shut-down of the refinery (right)

Table 1. Outcome of the Nelder-Mead simplex optimization algorithm with the agent-based model

	$EmPr_1$	$EmPr_2$	$EmPr_3$	$EmPr_4$	$EmPr_5$	*COS*	$P(\underline{x},\underline{d})$
1	0	0	0	0	0	60	-5.69 E8
2	0	0	100	100	0	60	-3.65 E8
3	0	0	200	300	0	60	-1.61 E8
4	0	0	300	250	0	55	-1.64 E8
5	0	0	300	250	0	50	-6.19 E7
6	100	100	100	100	100	50	-3.70 E8
7	10	10	500	600	10	50	1.39 E8
Final	4	4	349	382	4	54	1.45 E8

system through replication (van Dam *et al.*, in print). It has been used for the assessment of various policies, including those for procurement and storage.

To determine *nominal* process conditions the model performs optimisation for normal operation by choosing which crudes to buy, how much crude is needed and from which supplier to order. The mode of operation is scheduled based on predicted demands and the throughput for operation of the refinery is set based on actual demand from the consumer. When a disturbance is manifesting itself this normal approach is not adequate any more because the complexity is too big and there are too many inter-depended variables. A model-based decision-support tool is therefore called for.

Figure 1 illustrates the refinery behaviour under normal operation and after disturbance $\underline{d}$ = (30,0,0,0,0,0) which occurs on day t = 22, resulting in a loss of 24 m$ because of loss of production. Next, the agent-based simulation model supports the choice which response is the most appropriate given the many degrees of freedom. Preliminary results for the decision on emergency procurement and the change of operational scale are shown in Table 1. After 20 iterations no further improvement is made so the algorithm is stopped. The proposed solution prevents a shut-down of the refinery by buying emergency crudes to make up for the delayed ship and by slightly reducing the throughput. The loss caused by the disruption is reduced by 14.7 m$ (not keeping in mind penalties to be paid by the shipper for delays).

The change in operational configuration was not included here. For each possible $COC \in \{R1, R2, R3, R4\}$ an optimisation for $EmPr_i$ and *COS* has to be performed because choosing a different recipe will influence the criterion surface. CPU time is not an obstacle to running the optimisation through multiple cycles, after which the best solution can finally be selected.

5. Final remarks

The Nelder-Mead optimisation method is a commonly used method in process systems engineering using mathematical models or samples from experiments in a real system.

Here we demonstrated that it can be a powerful approach when combined with an agent-based model, too. The model is developed in a bottom-up fashion, making it relatively simple to change the configuration: It is easy to include new actors (e.g. more suppliers with different lead times) or to adjust the physical configuration (e.g. of the storage tanks). ABM are particularly suitable to experiment with different scenarios and to answer "what if" questions, critical for decision support under disruptions.

Determining the right combination of options (e.g. a switch of operational configuration and emergency procurement for the crudes used for that recipe instead of the ones in the delayed tanker) is difficult, and becomes extremely hard when responses at different times are allowed (e.g. a small emergency procurement now, switching recipe a few days later, switching back when the delayed ship arrives, etc) and when responses to new disturbances are also included (e.g. the long term effects of not always following planned operation). The full power of the approach presented here becomes visible then.

An illustrative case study using an agent-based simulation model of a supply chain with the Nelder-Mead optimisation method was presented and it was demonstrated how agent-based models can be applied as decision support tool. The approach presented can provide valuable support choosing the right response to abnormal situations in such a highly complex system.

Acknowledgment

The authors would like to thank Arief Adhitya for his support in the development of the agent-based model of the oil refinery. This project is supported by the Next Generation Infrastructures Foundation http://www.nginfra.nl/.

References

K. H. van Dam, A. Adhitya, R. Srinivasan, Z. Lukszo, 2008, Benchmarking numerical and agent-based models of an oil refinery supply chain, ESCAPE-18, France.

K. H. van Dam, A. Adhitya, R. Srinivasan, Z. Lukszo, in print, Critical evaluation of paradigms for modelling integrated supply chains, accepted for J. Computers and Chemical Engineering

N. Jennings (2000). On agent based software engineering, Artificial Intelligence 117: 277-296.

N. Julka, I. Karimi, R. Srinivasan, 2002, Agent-based supply chain management – 2: a refinery application, J. Computers and Chemical Engineering, 26(12), pp. 1771–1781.

L. Lu, D.J. Murray-Smith, D.G. Thomson, 2008, Issues of numerical accuracy and stability in inverse simulation, Simulation Modelling Practice and Theory 16, pp.1350–1364.

J.A. Nelder, R. Mead, 1965, A Simplex Method for Function Minimization, Comput. J. 7, pp.308–313.

A. Ouria, M. M. Toufigh, 2009, Application of Nelder-Mead simplex method for unconfined seepage problems, Applied Mathematical Modelling (In Press).

S. Pitty, W. Li, A. Adhitya, R. Srinivasan, and I.A. Karimi, 2008, Decision Support for Integrated Refinery Supply Chains. 1. Dynamic Simulation, J. Computers and Chemical Engineering, Vol 32, pp 2767–2786.

Z. Verwater-Lukszo, 1996, A practical approach to recipe improvement and optimization in the batch processing industry, PhD Thesis, Eindhoven University of Technology

M. H. Wright, 1995, Direct search methods: Once scorned, now respectable, In: D. F. Griffiths, & G. A. Watson (Eds.), Numerical analysis, Harlow, UK: Addison-Wesley, Longman.

Q. Xiong, A. Jutan, 2003, Continuous optimization using a dynamic simplex method, Chemical Engineering Science 58, pp. 3817–3828.

T. Ye, C.H. Xiong, 2008, Geometric parameter optimization in multi-axis machining, Computer-Aided Design 40, pp. 879–890.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Detailed Supply Chain Design considering Production Campaigns

Yanina Fumero,[a] Gabriela Corsano,[a] Jorge M. Montagna[a,b]

[a]*INGAR, Avellaneda 3657, (3000) Santa Fe, Argentina*
[b]*CIDISI – Facultad Regional Santa Fe, UTN, Santa Fe, Argentina*

Abstract

Until now, supply chain (SC) design models have been mainly focused on SC integration, where nodes allocation and links among them are selected in order to allow an efficient operation of the whole system. Usually, the detailed configuration and operation of the plants have not been taken into account. In this work, a heuristic strategy is presented in order to design the SC, including the structure and the operation of plants with mixed product campaigns. The incorporation of plant design and mixed product campaign in the SC design model leads to a non linear formulation. Hence, a two stages approach is addressed in order to solve this problem through linear models. In the first stage, a SC design standard model is solved in order to obtain the network design with minimum logistic cost and the production of each selected plant. These results are used to estimate the possible campaigns composition for each plant. In the second stage, specific scheduling constraints are incorporated in the plant design model to determine the optimal mixed product campaign configuration and the structure of each plant, minimizing the investment cost. This methodology allows obtaining the SC logistic configuration and, for each selected plant, the optimal mixed production campaign simultaneously with the plant design, in order to meet a specified economic criterion fulfilling demand requirements.

Keywords: supply chain, multiproduct batch plants design, mixed product campaign, optimization.

1. Introduction

A Supply Chain (SC) is a network of organizations that performs a set of activities in different phases of the production process. Determining an optimal configuration and efficient operation of the SC is a complex problem, taking into account the different criteria involved (Amaro and Barbosa-Póvoa, 2008; Guillén et al. 2006). Following Shah (2005), there are no published works dealing with the connection between plant and SC design and operation. Usually, plants are modeled as black boxes and nodes allocation and flows among them are determined in the solution.

In a previous work (Corsano and Montagna, 2008), a new approach has been presented in order to include plant design in SC models for multiproduct batch plants. This work shows that SC costs, design and capacity are severely affected by the plant structure. However, that work assumes single product campaigns (SPC) in order to simplify the formulation. This assumption is usual in these models but it is very limiting from the commercial point of view where a more steady supply of products is required.

At plant level, Birewar and Grossmann (1989) consider that the sequencing of batches of different products in mixed product campaign (MPC) can reduce idle times to increase the utilization of equipment. Nevertheless, their models incorporate MPC in the

design of multiproduct batch plants considering only one unit per processing stage. On the other hand, parallel units can reduce the cycle time of the campaign. However, there are few works considering MPC taking into account the formulation is more complex.
In this work, a heuristic approach is presented for SC and plants design optimization considering MPC. The proposed methodology is structured in two stages: first a MILP model determines the SC design and then, according to the production amounts of each product in each plant, a MILP model is used to select the MPC and the structure of each plant with minimum investment cost. The model assumes that units are provided using discrete sizes following the commercial procurement policy and that the transfer policy between stages is zero-wait (ZW). Unit duplication out of phase for each stage is taken into account in order to achieve an efficient plant design.

2. Problem Description and Mathematical Model

In this work, the SC comprises four echelons: raw materials sites, plants, warehouses and customers zones. At each raw material site, several types of raw materials are available to be delivered to plants. Each product can be produced at several plants. The warehouses can be supplied from more than one plant. Each customer zone demands for one or more products. The problem is to determine the allocation of raw materials sites, plants and warehouses, the flows among them, the amounts to be produced in each plant, and the design and the optimal MPC of each selected plant in order to meet a specified economic criterion fulfilling demand requirements during a time horizon.
A decomposition strategy is addressed through two MILP models where, in the first stage, a general model is formulated for the SC design to determine: i) the allocation of raw material sites, plants and warehouses; ii) the production of each product in each plant; and iii) the flows among SC nodes; in order to minimize the logistic cost (transport cost from raw materials sites to plants, from plants to warehouses and from warehouses to customer zones) and inventory costs. From the optimal solution, relations among the number of batches of each product for selected plants can be established according to the obtained production amounts. Therefore, the possible campaign compositions can be estimated.
Then, for each selected plant and each proposed composition, a MILP model is formulated. The production sequence of the MPC and the plant structure that minimize the investment cost are attained. Several compositions of the production campaign can be proposed for each plant, and the MPC with smallest objective function value is selected as the optimal one.

2.1. Logistic Model

Given the product demands, the set of available raw materials and their availability at each raw material sites, and the transport and inventory costs, the MILP model determines the production of each product in each plant, the allocation of plants, raw material sites and warehouses, and the flows among them in order to achieve the most economical design. Basically, the model considers mass balances between the SC nodes and allocation constraints for plants and warehouses. Similar constraints are presented by Corsano and Montagna (2008).

2.2. Plant Design and Mixed Product Campaign Model

This model introduces the MPC sequencing in the batch plant design. Previous models assumed SPC that simplified the formulation. From the logistic model solution, for each installed plant, the relation among the number of batches of the different products can be estimated. Therefore, the campaign composition can be proposed for each plant specifying the number of batches of each product in the campaign as data. Batches are

sequenced using the concept of production slot (Birewar and Grossmann, 1989), which requires the assignment of batches to these slots.

A batch plant includes N stages, K_j available units in parallel in stage j to produce N_p products. N_b represents the number of batches of the proposed MPC for that plant, NBC_i is the number of batches of product i in the campaign and N_s is the number of production slots where the N_b batches will be assigned. Since the parallel units in each stage j are assumed to be identical, a batch can be processed on any unit and the processing time will be the same. Thus, a unique processing time t_{ij} is defined for product i in the stage j.

2.2.1. Allocation Constraints

Binary variables Y_{bi} are used to indicate that a batch b is of product i. The following constraints must be satisfied for these variables:

Each batch b corresponds exactly to one product i: $\sum_i Y_{bi} = 1, \quad \forall b$ (1)

Product i is assigned to NBC_i batches: $\sum_b Y_{bi} = NBC_i, \quad \forall i$ (2)

In order to determine the processing of batch b in stage j, the binary variable YY_{bjkl} is defined and it takes value 1 if batch b is assigned to slot l and is processed in unit k in stage j, and 0 otherwise.

The slot l must be processed in a unit of stage j: $\sum_b \sum_k YY_{bjkl} = 1, \quad \forall j, l$ (3)

Moreover, in order to eliminate alternative solutions, for each stage j, the slot l_1 is arranged in unit k_1: $\sum_b YY_{bjkl_1} = 0, \quad \forall j,\ k,\ k \geq k_2$ (4)

Each batch b is processed in only one slot l of only one unit k of stage j:

$$\sum_l \sum_k YY_{bjkl} = 1, \quad \forall b, j \tag{5}$$

In addition, each batch b on different stages must be assigned to the same slot l:

$$\sum_l \sum_k l\, YY_{bjkl} = \sum_l \sum_k l\, YY_{bj'kl} \quad \forall b, j, j',\ j < j' \tag{6}$$

Let Z_{jk} be the binary variable equal to 1 if unit k of stage j is employed. In order to reduce the search space, units will be utilized in ascending order, i.e. unit $k+1$ is used only if unit k has been already allocated: $Z_{jk} \geq Z_{jk+1}, \quad \forall j, k$ (7)

When unit k of stage j is utilized, then at least a batch in a slot must be processed in that unit. Otherwise, this unit is not utilized to process any batch:

$$\sum_l \sum_b YY_{bjkl} \geq Z_{jk}, \quad \forall j, k \tag{8}$$

$$YY_{bjkl} \leq Z_{jk}, \quad \forall j, k, l, b \tag{9}$$

2.2.2. Timing constraints

Let T_{bj} be the processing time for batch b in stage j, TI_{jkl} and TF_{jkl} the starting and finishing times, respectively, to process the slot l in unit k of stage j. Equation (10) establishes the relation between these variables and YY_{bjkl}:

$$TF_{jkl} = TI_{jkl} + \sum_b \sum_i t_{ij}\, Y_{bi}\, YY_{bjkl}, \quad \forall j, k, l \tag{10}$$

When no batch is assigned to slot l of unit k in stage j, $YY_{bjkl} = 0$, and TI_{jkl} and TF_{jkl} are equal. This constraint can be linearized introducing an extra nonnegative variable W_{bijkl}

that takes value 1 if $Y_{bi}=1$ and $YY_{bjkl}=1$, and 0 otherwise, so (10) is replaced by (11):

$$TF_{jkl}=TI_{jkl}+\sum_b\sum_i t_{ij}\,W_{bijkl}, \qquad \forall j,k,l \tag{11}$$

where $TI_{j,k,l_1}=0$ and $W_{bijkl}\geq Y_{bi}+YY_{bjkl}-1, \qquad \forall b,i,j,k,l$. (12)

The new variable must satisfy the following constraint:

$$\sum_k\sum_b\sum_l W_{bijkl}=NBC_i, \qquad \forall i,j \tag{13}$$

In order to avoid the overlapping between different batches in a unit, the following constraint is added: $TF_{jkl}\leq TI_{jkl'}, \qquad \forall j,\ k,\ l,\ l',\ \ l<l'$ (14)

If no batch is processed in slot $l+1$, then the starting time of this slot is equal to final time of slot l: $TF_{jkl}-TI_{jkl+1}\geq -M\sum_b YY_{bjkl+1}, \qquad \forall j,k,l$ (15)

where scalar M is large enough.

When the batch b utilizes units k in stage j and k' in stage $j+1$, the ZW policy is imposed by: $TF_{jkl}=TI_{j+1k'l}, \qquad \forall j,\ l,\ k,\ k'$ (16)

In order to obtain the cycle time of the campaign, CT, the cycle time of unit k of stage j, CT_{jk}, must be determined, where l_1 and l_n are the first and the last slots of stage j:

$$CT_{jk}=TF_{jkl_n}-TI_{jkl_1}, \qquad \forall j,k=k_1 \tag{17}$$

$$CT_{jk}=TF_{jkl_n}-TI_{jk\tilde{l}}, \quad \forall j,k,k\geq k_2 \quad \text{where } \tilde{l}=\min_l\left\{\sum_b YY_{bjkl}=1\right\}. \tag{18}$$

Thus, CT is given by: $CT\geq CT_{jk}, \quad \forall j,\ k$. (19)

2.2.3. Design constraint

Let $SV_j=\{VF_{j1},.....,VF_{jP}\}$ be the set of available discrete sizes for stage j. vv_{jp} is a binary variable equal to 1 if units of stage j have size p, with $p=1,...,P$. Then, the size of units of stage j, V_j, is given by: $V_j=\sum_p vv_{jp}VF_{jp}, \qquad \forall j$ (20)

where $\sum_p vv_{jp}=1, \qquad \forall j$. (21)

SF_{ij} is the size factor of product i in stage j, Q_i the demand of product i, α_j, β_j cost coefficients of units in stage j, and H the time horizon. Let NN be a nonnegative variable that indicates the number of times that the campaign will be repeated in H. The following equations must be satisfied:

$$NBC_i\ NN\ B_i=Q_i, \qquad \forall i \tag{22}$$

$$V_j\geq SF_{ij}B_i, \qquad \forall i,j \tag{23}$$

Substituting equations (20) and (22) into inequality (23), yields:

$$NN\geq\sum_p\frac{SF_{ij}Q_i}{NBC_iVF_{jp}}vv_{jp}, \qquad \forall i,j \tag{24}$$

Defining the time horizon constraint by: $CT\ NN\leq H$ (25)

and replacing with (24), the following inequality must be satisfied:

$$\sum_p\frac{SF_{ij}Q_i}{NBC_iVF_{jp}}vv_{jp}CT\leq H, \qquad \forall i,j \tag{26}$$

In order to avoid non linear constraints, a new variable is defined as $X_{jp}=vv_{jp}CT$, and replaced in (26). The following constraints must be also satisfied:

$$\sum_{p} X_{jp} = CT, \qquad \forall j \tag{27}$$

$$X_{jp} \le CT^{UP}\, vv_{jp}, \qquad \forall j,p \tag{28}$$

2.2.4. Objective Function

The objective function minimizes the investment annual cost given by:

$$CI = \sum_{j}\sum_{k} Z_{jk}\alpha_j V_j^{\beta_j} \tag{29}$$

Considering equations (20) and (21), and defining the variable XX_{jkp}, which take value 1 if $Z_{jk}=1$ and $vv_{jp}=1$, and 0 otherwise, the equation (29) becomes linear:

$$CI = \sum_{j}\sum_{k}\sum_{p} \alpha_j VF_{jp}^{\beta_j} XX_{jkp} \tag{30}$$

This objective function subject to the previous constraints defines the MILP model to obtain simultaneously the optimal campaign sequencing and the plant design.

3. Example

Consider a network with 3 sites which supply 3 raw materials, 3 possible plants with 4 stages which can produce 3 products, 3 possible warehouses and 3 customer zones. The available sizes for units are SV = {100, 250, 400, 650, 900} and their cost coefficients are: α_{J1}=5000, α_{J2}=8000, α_{J3}=7000, α_{J4}=6000, $\beta_{J1}=\beta_{J2}=\beta_{J4}$=0.6 and β_{J3}=0.7. Table 1 shows demands in the time horizon H = 7000 hrs, and, Table 2 relevant data for plants.

The optimal solution for the logistic model is illustrated in Fig. 1. Two plants are installed: L1 that produces all products and L3 only two products.

For each plant, the relation among the number of batches of the different products can be estimated. For instance, for the plant L1 the following campaign compositions can be proposed: 2 batches of I1, 1 of I2 and 2 of I3; or 2 of I1, 1 of I2 and 1 of I3; or 4 of I1, 2 of I2 and 3 of I3; among others. The optimal solutions of plant design and MPC sequencing model for the first two compositions are shown in Tables 3 and 4. Up to two parallel units are allowed for every batch stages.

The annual investment cost is \$416400 and the optimal sequence is I3-I2-I3-I1-I1 for the first proposed campaign composition; and \$518000 and I3-I2-I1-I1 for the second one. From the proposed compositions and the reported results, the first MPC campaign is selected as the optimal one, and Fig. 2 shows its Gantt chart. The cycle time of the campaign is equal to 33.33 hrs and the campaign is repeated 210 times.

Table 1.
Product demands (x10^3kg)

Product	C1	C2	C3
I1	140	90	100
I2	70	45	50
I3	60	80	60

Table 2. Processing times and Size factors for each plant

	t_{ij} (h)				SF_{ij}			
	J1	J2	J3	J4	J1	J2	J3	J4
I1	6	12	4	10	0.7	0.6	0.5	0.7
I2	4	9	6	20	0.6	0.55	0.65	0.6
I3	4	10	3	5	0.7	0.6	0.5	0.6

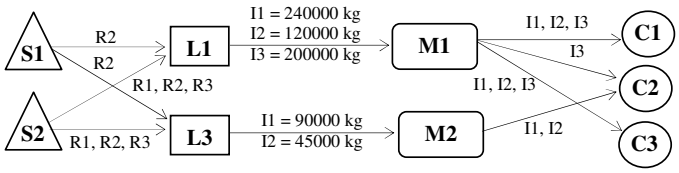


Figure 1. SC optimal design for logistic model

For plant L3, the composition campaign with 2 batches of I1 and 1 of I2 was proposed. For the optimal solution, the campaign configuration is I1-I2-I1 and the plant design consists of only one unit per stage of 250 l. The cycle time of the campaign is equal to 23 hrs and the campaign is repeated 126 times.
The models were solved using GAMS via CPLEX solver, with a 0% optimality gap.

Table 3. Optimal solution for L1 with campaign composition 2 I1 – 1 I2 – 2 I3

Stages	J1	J2	J3	J4
Units Number	1	2	1	2
Units Sizes (l)	400	400	400	400

$B_{I1} = 571.43$ $B_{I2} = 571.43$ $B_{I3} = 476.19$
Optimal Schedule: I3 - I2 - I3 - I1 - I1
Constraints: 695
Binary variables: 228
Continuous variables: 379
CPU time (s): 47.578

Table 4. Optimal solution for L1 with campaign composition 2 I1 – 1 I2 – 1 I3

Stages	J1	J2	J3	J4
Units Number	1	2	1	2
Units Sizes (l)	650	650	400	650

$B_{I1} = 480$ $B_{I2} = 480$ $B_{I3} = 800$
Optimal Schedule: I3 - I2 - I1 - I1
Constraints: 529
Binary variables: 156
Continuous variables: 291
CPU time (s): 8.734

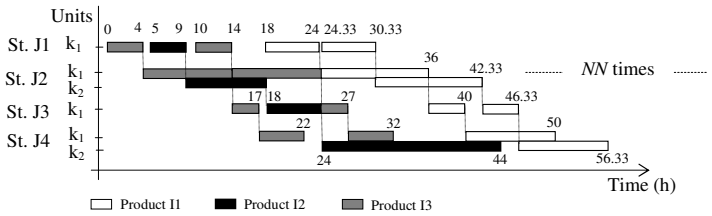


Figure 2. Gantt chart for campaign I3-I2-I3-I1-I1

4. Conclusions

A decomposition strategy for SC design optimization taking into account plants configuration and MPC was presented. In first stage a MILP model determines the SC design and then, according to the production amounts of each product in each plant, a new MILP model is used to determine the structure of each plant with minimum investment cost taking into account the MPC sequencing. Usually MPC sequencing is not considered in multiproduct batch plant design models.
As future work, the authors are exploring a methodology to solve both models in an overall MILP formulation, handling all the requirements.
Due to space inconvenient, some problem data was not reported and they are available for everyone who requests them.

References

A.C.S. Amaro and A.P.F.D. Barbosa-Póvoa, 2008, Supply chin management with optimal scheduling, Ind. Eng. Chem. Res., 47, 116-132.
D. Birewar and I. Grossmann, 1989, Incorporating scheduling in the optimal design of multiproduct batch plants, Comp. & Chem. Engn., 13, 141 – 161.
G. Corsano, and J. M. Montagna, 2008, Mathematical modeling for simultaneous design of plant and supply chain in batch process industry, Under review in Ind. Chem. Eng. Res.
G. Guillén et al., 2006, Addressing the design of chemical supply chain under demand uncertainty. Ind. Eng. Chem. Res., 45, 7566 – 7581.
N. Shah, 2005, Process industry supply chains: Advances and challenges. Comp. & Chem. Engn., 29, 1225 – 1235.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Multi-site Scheduling/Batching and Production Planning for Batch Process Industries

Georgios M. Kopanos, Luis Puigjaner

Universitat Politècnica de Catalunya-ETSEIB, Diagonal 647, 08028, Barcelona, Spain

Abstract

In the current work, a general mixed-integer linear programming (MILP) model is presented that contemplates most of the key aspects found in the scheduling/batching and production planning for multi-site batch process industries. Batch sizes and processing times are not fixed; instead they are optimized simultaneously with scheduling decisions that also consider products' changeovers and multiple production plants. This way a short-term enterprise-wide optimization entity is built up. Backlogs and period crossover costs/times are also explicitly considered. An illustrative case study is presented to highlight the characteristics of the proposed model.

Keywords: multi-site, scheduling, lot-sizing, planning, MILP.

1. Introduction

Nowadays, manufacturing facilities participating in the same supply chain should intensively share the necessary information and coordinate for the overall planning and scheduling tasks to achieve a holistic optimization performance (Alvarez, 2007). The multi-site production scheduling problem seeks the optimum use and allocation of available resources among several production sites. Scheduling and lot-sizing (batching) are a vital part of operational production planning and control. Moreover, simultaneous lot-sizing and scheduling is essential if sequence-dependent setup costs/times occur during the production process. Several attempts to cope with the scheduling/planning problem can be found in the literature. These contributions are mainly based on discrete or slot-based time representations. By nature, these approaches may provide myopic optimal solutions once implemented; especially, when processing times are lower than the predefined time intervals, since they usually restrict the allocation to only one product per unit and time interval. Thus, the production capacity is underestimated. One way to mitigate this effect, it is to define shorter time intervals or slots, at the expense of increased computational effort that may cause the problem to become intractable. Furthermore, changeovers are not properly handled and high unit idle times may appear. Note that the production scheduling problem has received little attention (Floudas & Lin, 2004). In a multi-site production environment, planning each plant separately will not, in general, lead to optimal solutions. Thus, a multi-site optimization over all production facilities is required (Bassett et. al, 1998; Shah, 1998).

2. Problem statement

The multi-site scheduling/batching and planning problem in batch plants with different processing units in parallel is addressed. Batch-product to unit-plant assignment and batch sequencing meeting a production goal constitute the problem under study. The main problem characteristics and proposed model assumptions include:

- A unit cannot process more than one batch at a time.
- Batch sizes B_{pijn} and processing times pt_{pijn} are variables.
- The unsatisfied demand of a period is not lost. It can be met in next periods generating a backlog cost as long as it is not satisfied.
- Each production facility has a maximum storage capacity.
- Raw materials' flows among plants are allowed PP_p (Fig.1).
- Products' demands, dem_{in}, are realized in specific time intervals, which constitute the production time intervals n.

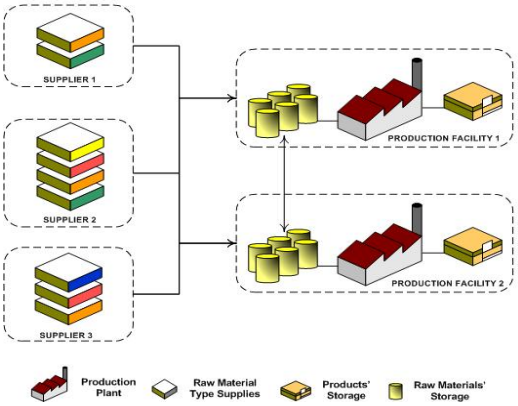


Figure 1. Multi-site production scheme

- Different units j may be installed to each plant p, PJ_j; plants may not be identical.
- Supply of raw materials availability is considered, SR_r; each supplier may provide a different set of raw materials, and with different maximum capacity.
- Final products recipe may need more than one raw material type, IR_r.
- Total sequence-dependent setup times $sd_{ii'j}$ are less than twice the time horizon n.

3. Mathematical model

In the proposed mathematical formulation, the problem constraints have been grouped according to the type of decision (such as lot-sizing, sequencing) to which they are associated. The basic set of equations are stated and briefly explained next.

3.1. Lot-sizing constraints

Parameters ρ_{ij}^{min} and ρ_{ij}^{max} denote the minimum and maximum production rate for producing product i at unit type j, respectively. Set JI_i refers to units of type j that can process product i, while JP_p corresponds to the set of plants p that have installed unit j.

$$\rho_{ij}^{min} pt_{pijn} \leq B_{pijn} \leq \rho_{ij}^{max} pt_{pijn} \qquad \forall p,i,j \in (JP_p \cap JI_i), n \tag{1}$$

$$pt_{ij}^{min} Y_{pijn} \leq pt_{pijn} \leq pt_{ij}^{max} Y_{pijn} \qquad \forall p,i,j \in (JP_p \cap JI_i), n \tag{2}$$

Y_{pijn} corresponds to the assignment binary variable of a batch of product i to unit j of plant p while pt_{pijn} and B_{pijn} variables denote the batch size and the processing time of every product-batch i at unit j of plant p at planning period n, respectively.

3.2. Sequencing-allocation constraints

Eqs.(3)-(5) apply the unit-specific immediate precedence sequential notion for the sequencing and allocation decisions. $X_{pii'jn}$ constitutes the sequencing binary variable, which is active, i.e. $X_{pii'jn}$=1, whenever a product-batch i' directly follows a product-batch i in the same unit j at production plant p at planning period n. Eq.(3) denotes that the number of the total number of active sequencing variables in a specific processing unit j is equal to the total number of the assigned products in this unit j minus 1.

$$\sum_{i \in IJ_j} \sum_{i' \neq i, i' \in IJ_j} X_{pi'ijn} = \sum_{i \in IJ_j} Y_{pijn} - 1 \qquad \forall p,j \in JP_p, n \tag{3}$$

$$\sum_{i' \neq i, i' \in IJ_j} X_{pi'ijn} + WF_{pijn} = Y_{pijn} \qquad \forall p, i, j \in (JP_p \cap JI_i), n \tag{4}$$

$$\sum_{i' \neq i, i' \in IJ_j} X_{pii'jn} + WL_{pijn} = Y_{pijn} \qquad \forall p, i, j \in (JP_p \cap JI_i), n \tag{5}$$

To continue with, WF_{pijn} and WL_{pijn} binary variables represent the first and the last assigned product-batch *i* at each unit *j* of every plant *p* at each planning period *n*. Set IJ_j contains all final products *i* that can be processed in unit type *j*. Eq.(4) states that when a batch *i* is allocated to a unit *j*, it may be assigned in the first place, i.e. WF_{pijn}=1, or there exists a batch *i'* that is processed before it, i.e. $X_{pi'ijn}$=1. Following the same logic, eq.(5) expresses that when a batch *i* is allocated to a unit *j*, it may be assigned in the last place, i.e. WL_{pijn}=1, otherwise it exists a batch *i'* that is processed after it, i.e. $X_{pii'jn}$=1.

3.3. Timing constraints

It is worth mentioning that eqs.(3)-(5) are not sufficient to prevent the generation of sequence sub-cycles. Hence, the timing constraints of eqs.(6)-(7) are included into the formulation in order to thwart the occurrence of any possible sequence sub-cycle.

$$C_{pi'jn} \geq C_{pijn} + pt_{pi'jn} + su_{i'j} WF_{pi'jn} + sd_{ii'j} - M(1 - X_{pii'jn}) \qquad \forall p, i, i' \neq i, j \in (JP_p \cap JI_i \cap JI_{i'}), n \tag{6}$$

$$su_{ij} WF_{pijn} + pt_{pijn} + \sum_{i' \neq i, i' \in IJ_j} sd_{i'ij} X_{pi'ijn} \leq C_{pijn} \leq Hor_{pjn} Y_{pijn} \qquad \forall p, i, j \in (JP_p \cap JI_i), n \tag{7}$$

Eq.(6) forces the timing between a pair of two different product-batches *i* and *i'* that can be allocated to the same unit *j*. C_{pijn} variable denotes the completion time of a product-batch *i* at unit *j,* and su_{ij} is the sequence-dependent setup time. Finally, eq.(7) provides the upper and the lower bound for the completion time variable C_{pijn}.

3.4. Period crossovers timing constraints

In order to take into account the sequence-dependent costs and times between two consecutive batches across the borders of two adjacent production periods, the following set of equations is needed. By using eq.(8), sequence-dependent ($ssd_{ii'j}$) and –independent ($su_{i'j}$) times are treated in tandem (see Fig.2).

$$sd_{ii'j} = ssd_{ii'j} + su_{i'j} \qquad \forall i, i' \neq i, j \in (JI_i \cap JI_{i'}) \tag{8}$$

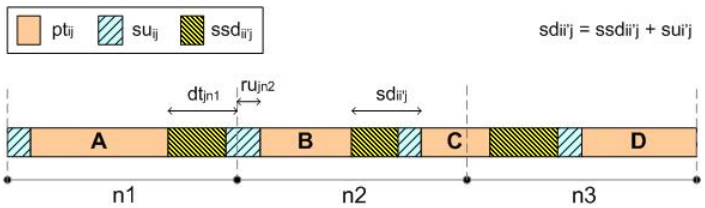


Figure 2. Illustrative scheme for variables dt_{pjn} and ru_{pjn}, and parameter $sd_{ii'j}$.

The product-batch *i'* that is assigned first at unit *j* of plant *p* at period *n* ($WF_{pi'jn}$=1) follows the last processed product-batch *i* at unit *j* of plant *p* at previous period *n-1* (WL_{pijn-1}=1), as eq.(9) states. Binary variable $X_{pii'jn}^{cros.}$ corresponds to the sequencing variable across the border of two successive planning periods; period *n-1* and period *n*.

$$X^{cros.}_{pii'jn} \geq WL_{pijn-1} + WF_{pi'jn} - 1 \quad \forall p, i, i' \neq i, j \in (JP_p \cap JI_i \cap JI_{i'}), n > 1 \quad (9)$$

Additional constraints, eqs.(10)-(11) indicate that the total operating time plus the total changeovers time plus the 'ready' unit time, ru_{pjn}, and the 'dead' unit time, dt_{pjn}, is less than the available unit and plant dependent scheduling horizon Hor_{pjn}. Fig.2 delineates a representative scheme of the continuous variables ru_{pjn} and dt_{pjn}. These variables have been introduced in order to take into account effectively the sequence-(in)dependent setup times between two consecutive batches across periods' borders. By eq.(12), the sequence-(in)dependent setup times can be split in two adjacent planning periods, thus improving capacity utilization.

$$ru_{pjn} + dt_{pjn} + \sum_{i \in IJ_j} pt_{pijn} + \sum_{i \in IJ_j} \sum_{i' \neq i, i' \in IJ_j} sd_{ii'j} X_{pii'jn} \leq Hor_{pjn} \quad \forall p, j \in JP_p, n > 1 \quad (10)$$

$$ru^0_{pjn} + dt_{pjn} + \sum_{i \in IJ_j} (pt_{pijn} + su_{ij} WF_{pijn}) + \sum_{i \in IJ_j} \sum_{i' \neq i, i' \in IJ_j} sd_{ii'j} X_{pii'jn} \leq Hor_{pjn} \quad \forall p, j \in JP_p, n = 1 \quad (11)$$

$$ru_{pjn} + dt_{pjn-1} \geq \sum_{i \in IJ_j} \sum_{i' \neq i, i' \in IJ_j} sd_{ii'j} X^{cros.}_{pii'jn} \quad \forall p, j \in JP_p, n > 1 \quad (12)$$

The aforementioned set of equations can be viewed as the extended formulation of a recent contribution (Sung & Maravelias, 2008) where only sequence-independent setups were considered.

3.5. Mass balance constraints

Eq.(13) reflects the total production of a product i at each period n, Pr_{in}.

$$Pr_{in} = \sum_{p} \sum_{j \in (JP_p \cap JI_i)} B_{pijn} \quad \forall i, n \quad (13)$$

Eq.(14) corresponds to the products mass balance at each period n. The unsatisfied demand of a period n, dem_{in}, is represented as backlog, $Back_{in}$. This backlog is added to the demand of the following period; it also generates an additional cost. St_{pin} stands for the stocked amount of product i at plant p at period n.

$$\sum_{p} St_{pin} - Back_{in} = \sum_{p} St_{pin-1} - Back_{in-1} + Pr_{in} - dem_{in} \quad \forall i, n \quad (14)$$

The raw materials mass balance is forced by eq.(15). The variable $Q_{pp'rn}$ denotes the quantity of raw material r transferred from plant p to plant p' at period n, while TR_{sprn} represents the amount of raw material r transferred from supplier s to plant p at period n. Parameter φ_{ri} is the stoichiometry coefficient of raw material r for producing final product i. St^{RM}_{pin} stands for the stocked amount of raw material r at plant p at period n.

$$St^{RM}_{prn} = St^{RM}_{prn-1} + \sum_{s \in (SR_r \cap SP_p)} TR^{RM}_{sprn} + \sum_{p' \in PP_p} Q_{p'prn} - \sum_{p' \in PP_p} Q_{pp'rn} - \sum_{i \in IR_r} \sum_{j \in (JP_p \cap JI_i)} \phi_{ri} B_{pijn} \quad \forall p, r, n \quad (15)$$

3.6. Capacity constraints

Equations considering maximum storage capacity of raw materials and maximum storage capacity of products for every production facility, *p*, are also included. Moreover, supplier capacity restrictions, regarding the amount of each raw material type that can provide, are considered. Finally, transport capacity of raw materials between each pair of production facilities is also taken into account.

3.7. Objective function

The minimization of backlogs, inventory, sequence-dependent, setup, operating, raw materials' purchase/transportation and raw materials' interplant transportation costs are the optimization goal.

4. Case study

It follows a brief description of the main special features of this case study in order to demonstrate its complexity. Afterwards, the results obtained are thoroughly discussed.

4.1. Case study description

In this section, it is addressed the multi-site batching/scheduling and planning problem of 20 final products (*i1-i20*) which can be produced in 5 different processing unit types (*J1-J5*) of 2 different production plants (*P1-P2*). 6 different raw materials' types (*r1-r6*), which are provided from 4 different suppliers (S1-S4), are necessary to produce the set of final products. Products demand arrives in a weekly basis. 4 planning periods (168 hours each) have been considered. Production plants are not identical; production plant *P1* has unit types *J1* to *J5*, while production plant *P2* has installed units *J1* to *J3*. Suppliers *S1*, *S2* and *S4* can provide raw materials to *P1*, and suppliers *S1-S3* can provide raw materials to *P2*. Supplier *S1* provides *r1-r3*, S2 provides r2-r5, *S3* provides *r1* and *r6*, and *S4* provides *r4-r6*. 1 final product (*i20*) requires 3 different raw material types. 13 final products consume 2 different raw material types while 6 final products need 1 raw material type. 1 product (*i8*) can be done in 4 different units. 6 products can be produced into 3 different unit's types. 12 products can be processed in 2 different unit's types and 1 product (*i10*) can be produced just to 1 unit's type. Raw materials can be transferred between production facilities. The minimum production time is equal to 10 hours for all units but unit *J5* wherein it is equal to 5 hours.

4.2. Case study results

The model consists of 6,744 equations, 1,444 continuous variables and 5,256 binary variables. The problem has been solved to optimality in GAMS 22.8 (CPLEX 11.0), in a Dell Inspiron 1520 (2 GHz, 2 GB RAM), after 430 CPU seconds.

Fig.3 shows the contribution of each production facility to the total inventory profile for every planning period. There is no inventory at the last planning period. There are no backlogs. Fig.4 shows the multi-site Gantt chart for both plants at all planning periods. Raw materials flows are observed between the two production plants.

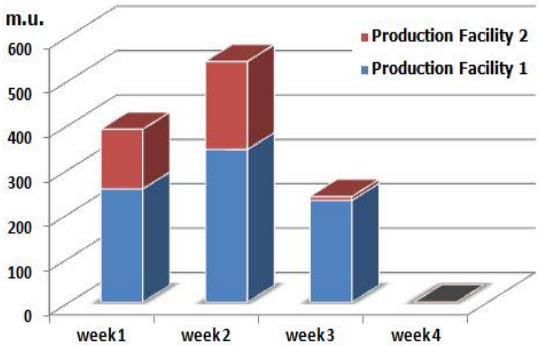


Figure 3. Inventory profile for the case study

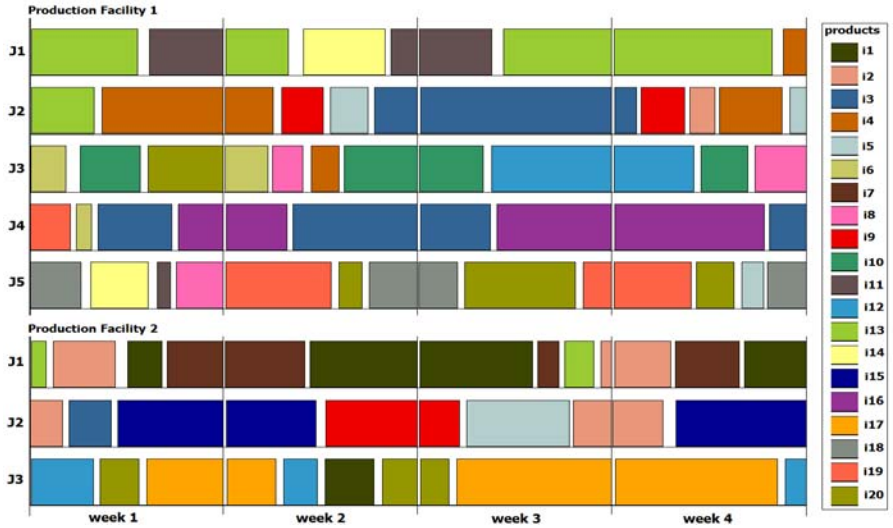


Figure 4. Gantt chart for the case study

5. Final considerations

In this work, a new multi-site batching/scheduling and planning MILP framework has been proposed. Raw material flows are allowed among production facilities; exploring the flexibility of these operations. Sequence-dependent costs are explicitly calculated avoiding the generation of ostensible optimal solutions. The ability of our approach to split sequence-(in)dependent times across the border of adjacent planning periods results to production capacity utilization improvement. Our current research activity is focused on the expansion of the current model to cope with more complicated product recipes, where intermediate products are present. All materials flows among the nodes of the supply chain entity will also be considered in order to take into advantage the flexibility of logistics operations. Finally, decomposition techniques will be revised and built up with the purpose of tackling large-scale industrial problems.

Acknowledgements

Financial support received from the Spanish Ministry of Education and Innovation (FPU grants) is fully appreciated. Authors thank Christos T. Maravelias for the enlighten discussions during the research visit of Georgios M. Kopanos in the University of Wisconsin-Madison.

References

E. Alvarez, 2007, Multi-plant production scheduling in SMEs, Robotics and Computer-Integrated Manufacturing, 23, 608-613

M.H. Bassett, P. Dave, F.J. Doyle III, G.K. Kudva, J.F. Pekny, G.V. Reklaitis, S. Subrahmanyam, D.L Miller, M.G. Zentner, 1996, Perspectives on model based integration of process operations. Computers & Chemical Engineering, 20, 821-844

C.A., Floudas, X. Lin, 2004, Continuous-time versus discrete-time approaches for scheduling of chemical processes: a review. Computers & Chemical Engineering, 28, 11, 2109–2129

N. Shah, 1998, Single and multisite planning and scheduling: current status and future challenges, Proceedings of the 3rd International Conference on FOCAPO, 75-90

C. Sung, C.T. Maravelias, 2008, A mixed-integer programming formulation for the general capacitated lot-sizing problem, Computers & Chemical Engineering, 32, 1-2, 244-259

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Optimal Location Planning for Self-Storage Enterprises

Richard Lackes[a], Markus Siepermann[a]

[a]*Department of Business Information Management, Technische Universität Dortmund, Vogelpothsweg 87, 44227 Dortmund, Germany*

Abstract

The location planning problem is a long-term problem that aims for the determination of storehouse sites and the point in time when they should be put on stream. The decision criterion consists in the net present value. The paper introduces a decision model based on binary decision variables representing the sites and the points in time of investment with respect to a given budget and the constraints of the market situation and volume.

Keywords: Location planning, self-storage, decision model, genetic algorithm

1. Objective

In the past decade, an innovative concept for storehouses evolved with its origin in the USA (see Self Storage Association 2007). For outsourcing purposes, service providers offer storage capacities for individuals as well as for business users. The market is promising because the investment in storehouses, the operating costs and the market penetration are relatively low while the potential demand is high (see Duffy/Kliebenstein 2005). The basic idea is as follows: The service provider procures standardized storage room for a short period of time, the storage equipment (roller shutter, fork lift etc.) and administers the storehouse, but stockpiling and stock removal have to be done by the customers themselves. Due to an electronic entry system customers can access their rented storage capacities at every time independent of the presence of warehouse employees. For renters, this concept allows to reduce fixed storage costs that can now be replaced by usage dependent variable costs (see Mark 2005).
Because of the innovative character of the service and the developmentally chances many new sites will emerge in the next few years (see Duffy/Kliebenstein 2005). Therefore, it is very important in this stage to choose appropriate locations. Densely populated areas are attractive because of the restricted catchment area of a storehouse and the closeness to the potential target group. But the competitive situation and the investment costs in these areas are normally inauspicious. Thus, we are facing a complex long term site planning problem: How to choose the site that is the economically favorable one for the next years? (see Fleischmann/Klose 2004) In the following we present a multi periodic optimization model for this problem and show how this problem can be solved.

2. Development of the Optimization Model

2.1. Characteristics and Goal

The characteristics of self-storage storehouses (SSS) are (see Duffy/Kliebenstein 2005): (1) Construction and equipment of SSS are considered as a medium-/long-term investment. (2) Once a decision on investment and location is made, a revision can't be taken without greater loss. (3) SSS provide a certain capacity of storage space. (4) The offered

product »storage possibility« isn't affected by usage concerning its quality and life expectancy. (5) Operating costs of an SSS aren't constrained by use and load. They just ensure the disposability (availability fees). Thus, the marginal costs of an additional contract, if it lies within the capacity limits, matches 0. Non-use of storage capacity doesn't diminish the operating costs. (6) The sales market of an SSS is locally bounded to the location of choice. Main target group are individuals and craftsmen. (7) Even if the rental contracts allow flexible durations, most of the contracts are on a long term basis. To convince a customer once is important for the »natural« customer loyalty.

The location planning for self-storage enterprises is a multi-periodic dynamic decision problem. The planning horizon amounts to T years, scaled in t=1,2,...,T periods. Opening of storehouses takes place at period begin. The present point in time is t = 0. The goal is the maximization of the net present value that is determined by all site decisions made within the planning horizon. The decision concerns the expansion strategy. That is if and when storehouses should be built at a location within the planning horizon.

2.2. Optimization Model

2.2.1. Site Alternatives and their Characteristics

The investigation area shall be split into equal grid boxes (e.g. 5*5 km), each regarded as an »atomic« location element. A location will be determined by its x- and y-coordinates (x,y). Let x=1,…,X and y=1,…,Y be the relevant coordinates of all locations, so that the complete surface of the investigation area can be covered. Irrelevant locations that lie outside the grid because of the irregular shape of the investigation area are initially kept for an easier, formal description even if they fall apart later on. Each location is characterized by a set of attributes. Depending on their values a location rating can be computed. The relevant attributes of a location (x,y) for period t = 0,…,T are:

1. Outpayment: Land prices $GP_{x,y,t}$ [$]; storehouse equipment costs $LEQ_{x,y,t}$ [$]; labor costs index $LNIV_{x,y,t}$; annual outpayment-effective operating costs $K_{x,y,t}$ [$/period].
2. In-payment: Population $POP_{x,y,t}$ measured in 1.000; purchasing power of population, measured by purchasing power index $KKI_{x,y,t}$; market range of coverage, attainable price per square meter $P_{x,y,t}$ [$/sqm]; average rented storage space per contract [sqm]; life cycle curve of »salable« storage space (contracts or rather rented space) conditioned by age of the storehouse; competitive situation (foreign as well as one's own SSS in catchment area); economic trend

2.2.2. Prerequisites and Decision Variables

For the optimization, the following conditions shall hold: (V1) At each location should be built a maximum of one storehouse. Locations with an already existing storehouse aren't considered any further (see further V2). (V2) Shutting down of storehouses won't be allowed. (V3) There exist competitors on the market. (V4) Due to financial shortage or other bottlenecks only Bt storage houses can be built in one period t. (V5) Each storehouse provides a certain maximum capacity of storage space KAPx,y (e.g. 4200 sqm). (V6) The periods aren't subdivided any further. All payments, except acquisition payments, occur at the period-end.

The decision variables consist of binary variables differenced after the locations and the construction periods (see Vahrenkamp 2007). Because of condition V1, only the values 0 (no construction) and 1 (construction of an storehouse) can occur so that the optimization model is a binary decision problem with the decision variables $S_{x,y,t}$:

$$S_{x,y,t} = \begin{cases} 1 & \text{if a storehouse is built at (x,y) in t} \\ 0 & \text{else} \end{cases} \quad \text{with } x \in \{1,...,X\},\ y \in \{1,...,Y\},\ t \in \{1,...,T\}$$

At the start of planning, already existing storehouse locations are such $(x,y) \in X \times Y$ with $S_{x,y,0} = 1$. Because of condition V1 and V4 it applies formula (1) and (2):

$$(1) \qquad \sum_{t=0}^{T} S_{x,y,t} \leq 1 \text{ for all } (x,y) \in X \times Y$$

$$(2) \qquad \sum_{x=1}^{X} \sum_{y=1}^{Y} S_{x,y,t} \leq B_t \text{ for all } t = 1,...,T$$

2.2.3. Determination of Outpayments for Equipment and Operation of a Storehouse

With regard to the outpayment, we have to consider site specific land prices $GP_{x,y,t}$ and site neutral payments for storage equipment $LEQ_{x,y,t}$ (e.g. 2 Mio $). Furthermore, there occur operating costs which are almost fixed costs. With approximately 50%, labor costs are the biggest cost pool as surveys are showing. That means the annual site neutral costs affecting payments are K_t [$/year] and the site specific costs – affected by the labor costs index $LNIV_{x,y,t}$ – are represented by $K_t \cdot LNIV_{x,y,t}$. The labor costs index $LNIV_{x,y,t}$ indicates the multiplier referring to a base salary (e.g. 1.07). From this, the annual costs affecting payments for the location (x,y) result in: $K_t \cdot (1 + LNIV_{x,y,t})$.

2.2.4. Determination of In-Payments

2.2.4.1. Calculation of Market Potential

A storehouse's market range of coverage is determined by its catchment area. It may reaches beyond its own location (x,y) and can also contain the ones nearby. Therefore, we define a degree of proximity $1 \geq Ng((x_1,y_1),(x_2,y_2)) \geq 0$ for all pairs of locations in such way, that they decrease with increasing distance from the observed location. It indicates which share of the population in (x_2,y_2) can be reached by a storehouse in (x_1,y_1) due to distance and transportation infrastructure. The degree of proximity of one's own location obviously is 1. Symmetry shall always apply. All locations in the neighbourhood with a positive degree of proximity are relevant for the site decision. With the help of this environment information the potential reachable customers $KUZ_{x,y,t}$ of a location $(x,y) \in X \times Y$ in period $t=1,\ldots T$ is determined as:

$$(3) \qquad KUZ_{x,y,t} = \sum_{i=1}^{X} \sum_{j=1}^{Y} POP_{i,j,t} \cdot Ng\big((x,y),(i,j)\big)$$

with $POP_{i,j,t}$ is the population of location grid box (i,j) in period t. If there are competing storehouses (own or foreign) that have access to the same market potential then the market potential has to be split. It has to be noted that not only storehouses of competitors but also own storehouses may reduce the market potential of a location (cannibalization effects). Let $L_{i,j,t}$ be the number of storehouses at the beginning of period t in location (i,j) without differencing of own and foreign storehouses. Thus, the starting situation is described by $L_{i,j,1}$ with $L_{i,j,1} \geq S_{i,j,0}$ because of the competing storehouses. Then, considering the starting situation and the site decisions within the planning horizon the value of $L_{i,j,t}$ (t>1) can be computed as:

$$(4) \qquad L_{i,j,t} = L_{i,j,1} + \sum_{t'=1}^{t-1} S_{i,j,t'} \text{ with } (i,j) \in X \times Y;\ t=2,3,\ldots,T$$

The »access intensity« $ZUG_{i,j,t}$ that describes how many customers can be reached in period t by the storehouse in location (i,j) can be defined as:

$$(5) \qquad ZUG_{i,j,t} = \sum_{k=1}^{X} \sum_{l=1}^{Y} Ng\big((k,l),(i,j)\big) \cdot L_{k,l,t} \text{ with } (i,j) \in X \times Y;\ t=1,2,\ldots,T$$

Consequently, in period t the relevant market potential (in thousand inhabitants) of a storehouse to be built in location (x,y) is determined by:

$$(6) \qquad MP_{x,y,t} = \sum_{i=1}^{X}\sum_{j=1}^{Y} POP_{i,j,t} \cdot \frac{Ng\big((x,y),(i,j)\big)}{\max\left\{Ng\big((x,y),(i,j)\big) + ZUG_{i,j,t};1\right\}}$$

In a bottleneck situation, the market potential is distributed proportionally in accordance to the degree of proximity.

2.2.4.2. Attainable Price and Quantity of sales

Empirical studies have shown that the price per sqm $P_{x,y,t}$ correlates positively with the purchasing power index $KKI_{x,y,t}$. The capacity utilization depends except for the market potential $MP_{x,y,t}$ on the the age of a storehouse. Thus, there is a »life cycle curve« that can be described with the age dependent success rate $SUCCESS_s$ ($s=1,\ldots,T_L$) measured in contracts per 1.000 reachable customers. T_L is the lifetime of a storehouse. In its beginning a storehouse becomes known and gets used until the capacity limit is reached. The typical curve is first ascending continuously up to a certain absorption point and then stagnating. Practical experiences have shown that this point usually is reached after six years. The reason for this phenomenon is that many customers are storing goods during a long period of time. Once a storehouse has gained a customer he most likely will rent his storage box over the next years. This effect is enhanced by relatively high costs for stock transfer if another storehouse will be chosen for rental. Therefore, the number of contracts $AK_{x,y,t}$ is determined by (t is the storehouse's building time):

$$(7) \qquad AK_{x,y,\tau} \le SUCCESS_{\tau-t+1} \cdot MP_{x,y,\tau} \qquad \tau = t, t+1, \ldots T$$

Additionally, the number of contracts is limited by the capacity of a storehouse. Let $LF_{x,y,t} = f_L(KKI_{x,y,t})$ with $dLF/dKKI_{x,y,t} \ge 0$ be the averaged storage space per contract. Then, the demand is determined by the number of contracts multiplied with the averaged storage space per contract.

2.2.4.3. The Investment's Residual Value

Let $RW_{x,y,t}$ be the residual value of a storehouse built in period t at location (x,y). This value represents a storehouse's value at the end of the planning horizon T. It is needed because the revenue of such an investment takes place after a certain period of time that might lie beyond the planning horizon. If $RW_{x,y,t}$ wouldn't be taken into account investments at the end of the planning horizon would be monetarily misinterpreted.

2.2.5. Objective Function

The acquisition value $AW_{x,y,t}$ of a storehouse at location (x,y) at the beginning of period t has two components: The site specific land price $GP_{x,y,t}$ as well as the site neutral payments for the storehouse equipment $LEQ_{x,y,t}$:

$$(8) \qquad AW_{x,y,t} = GP_{x,y,t} + LEQ_{x,y,t}$$

Then, the net present value $CV_{x,y,t}$ of a storehouse in location (x,y) built at the beginning of period t ($x \in X$, $y \in Y$, und $t=1,\ldots,T$) will be:

$$(9) \qquad CV_{x,y,t} = -AW_{x,y,t} \cdot (1+i)^{-(t-1)} + RW_{x,y,t} \cdot (1+i)^{-T} + \sum_{\tau=t}^{T}\left(AK_{x,y,\tau} \cdot LF_{x,y,\tau} \cdot P_{x,y,\tau} - K_{x,y,\tau} \cdot \left(1+LNIV_{x,y,\tau}\right)\right) \cdot (1+i)^{-\tau}$$

The discounting always is for the planning horizon begin, i.e. t=0. The acquisition payments accrue at period begin, all other payments at period-end. Naturally, the following condition must hold:

$$(10) \qquad CV_{x,y,t} \ge 0 \text{ with } (x,y) \in X \times Y;\ t=1,2,\ldots,T$$

Now, the objective function consists in the maximization of the total net present value:

$$(Z) \qquad \text{Max} \sum_{t=1}^{T}\sum_{k=1}^{X}\sum_{l=1}^{Y} CV_{x,y,t} \cdot S_{x,y,t}$$

Because $S_{x,y,t}$ are binary variables we are facing a binary decision model. The number of decision variables is $X \cdot Y \cdot T$. In order to compute the access intensity $ZUG_{i,j,t}$ all existent storehouses in t plus those storehouses to be built (represented by $S_{x,y,t}$) including the proximity index have to be considered. Additionally, the access intensity is a determination factor of the market potential. Thus, the decision model is NP-hard and cannot be solved in an acceptable calculation period. Even commonly known optimization algorithms like e.g. branch and bound have to compute all solutions in order to find the optimal solution. Only with proximity index 0 for all neighbour grid boxes a classical binary optimization algorithm could succeed because then, the access intensity only depends on the known starting situation at the beginning of the planning horizon.

3. Example

Let us now apply the presented model to Germany. Figure 1 shows a map of Germany on the left side. On the right side we can see the grid boxes the investigation area is split into. The grey grid boxes indicate the six areas with the highest population density: Hamburg, Berlin, Munich, Frankfurt, Cologne and the Ruhr. These are the regions the model is coping firstly. Due to the relatively low outpayments in Berlin and the Ruhr, storehouses are first of all built in respectively beside these regions. The chosen locations depend on the values we are using for each grid box and the access intensity. If the outpayments of the white boxes are low in comparison to the grey boxes and the access intensity is not zero the chosen locations lie outside the six areas. Otherwise, if each region consists homogenously of boxes with equal values the chosen locations lie in the center of each region.

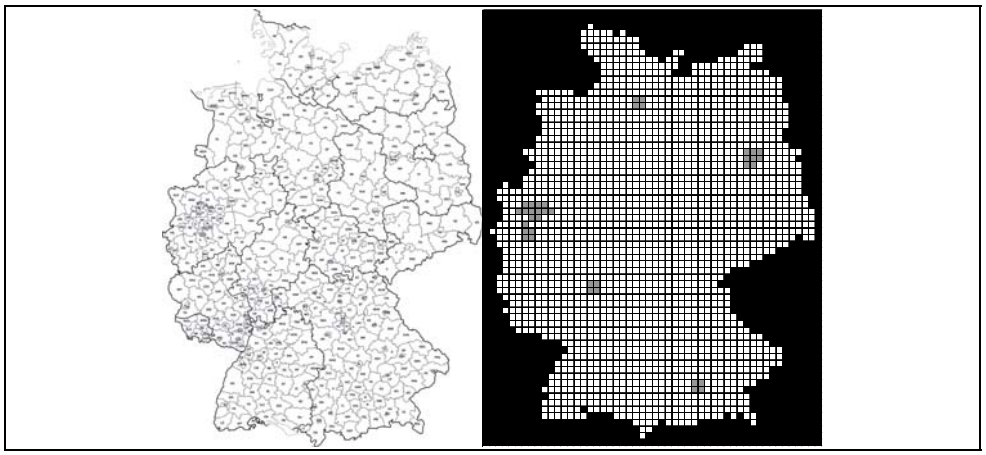


Figure 1: Investigation area Germany

The calculation for Germany could only be done because we reduced the number of grid boxes without empty values and the number of periods so that the number of possible solutions was decreased enormously. If we use a totally filled grid the calculation period explodes: With five periods and a 100×100 grid the solution space consists of 9,5 billion elements that have to be computed. Therefore, we will use a genetic algorithm in order to find good solutions in an acceptable time.

4. A Genetic Algorithm as Approach

4.1. Individuals

In the following we assume $B_t = 1$ for simplification. A solution can be described via a 3D-cube with the location coordinates and the periods as dimensions (in the following see Mitchell 1998, Vose 1999). Then, an individual of the genetic algorithm is one solution alternative that can be defined as follows: The individuals can be represented as a X×Y-matrix M. The matrix contains a maximum of T values between 1 and T whereas no value occurs twice. The other values of the matrix are 0. A value $M_{x,y} > 0$ in the matrix indicates that a new storehouse is built at site (x,y) in period $M_{x,y}$. The value 0 indicates that no storehouse is built at site (x,y). The net present value is used as fitness to choose the individuals for crossover and the selection of the next generation.

4.2. Mutation

The mutation operator can be defined as follows: Randomly choose two cells of an individual M with different values and swap the two values. Because of the crossover we are discussing later on some values might be lost during the calculation. In order to reproduce those missing values between 1 and T we can insert them instead of swapping two values.

4.3. Crossover

Let M and N be individuals of the genetic algorithm. Then, the crossover operator exchanges parts of the two individuals as follows: Randomly choose the bounds of a rectangle in the matrix. Then the rectangles are cut out of M and N and implanted into the other solution. This operation may lead to invalid individuals: The number of values between 1 and T may now be greater than T (case 1). Values between 1 and T may occur twice but there are only T values greater than 0 (case 2). Therefore, a repair mechanism has to be installed. In the first case, we randomly choose a cell with a value that occurs twice and set the value of the cell to 0 until there are only T values greater than 0. In the second case there are doublets as well as missing values. Therefore, we randomly choose a cell with a doublet value and set the cell's value to one of the missing value until no value occurs twice.

5. Conclusion and further Enhancements

In this paper we presented an optimization model for the location planning of self-storage enterprises concerning the expansion strategy. As this problem is a binary decision problem with many decision variables it can hardly be solved with deterministic algorithms. As genetic algorithms rapidly find good solutions (see Koza 1993) we designed a genetic algorithm that finds a good solution in an acceptable computation time.

References

Duffy, S., Kliebenstein, R.K., 2005, How to invest in Self-Storage, MiniCo Publishing.

Fleischmann, B., Klose, A. (eds.) 2004, Distribution Logistics: Advanced Solutions to Practical Problems, Springer.

Koza, J.R., 1993, Genetic Programming, MIT Press.

Mark, K., 2005, Leveraging Wal-Mart, eBay and USPS, in Morse, E. A., Mitchell, R. K., Cases in Entrepreneurship: The Venture Creation Process, SAGE.

Mitchell, M., 1998, An Introduction to Genetic Algorithms, MIT Press.

Self Storage Association, 2007, US Self Storage Industry 1985 to 2005: Facilities, Square Meters, Revenue etc., Self-Storage-Conference 2007, Paris.

Vahrenkamp, R., 2007, Logistik, 6. Auflage, Oldenbourg.

Vose, M.D., 1999, The Simple Genetic Algorithm: Foundations and Theory, MIT Press.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

What, if anything, is a Chemical Engineer?

Laureano Jiménez-Esteller, Gonzalo Guillén-Gosálbez, Dieter T. Boer

Department of Chemical Engineering, School of Chemical Engineering, University Rovira i Virgili, Av. dels Països Catalans, 26, Tarragona, 43007, Spain. E-mail: {Laureano.Jimenez,Gonzalo.Guillen, Dieter.Boer}@urv.cat

Abstract

The core domains of Chemical Engineering are in crisis as corresponds to a mature discipline that has been going on for more than 100 years. However, over the last two decades, some concerns have been raised about the identity and future of Chemical Engineering and our ability to adapt to new challenges: (1) Chemical engineering (ChE) graduates accept positions in an extremely vast a diverse range of sector, all having specific knowledge requirements; (2) globalization of the chemical industry forces professionals to work with people with different values and principles; (3) the research scope of ChE has expanded, and now the *hot areas* are at the interfaces with other disciplines, more focused in products than in processes, and covering the spectrum from the macroscopic to the molecular level. To adapt to this new scenario, an educational model was designed and implemented at the School of Chemical Engineering (ETSEQ) in Tarragona (Spain) to enable ChE students to acquire and integrate technical and scientific knowledge through the simultaneous and gradual development of competencies encompassing social and management skills. This model is based on the large-scale deployment of a project-based cooperative learning approach throughout the curriculum.

Keywords: problem-based solving, non-technical skills, project based learning.

1. The realm of ChE is large and diverse

I would like to begin this communication by explaining its title. I have borrowed it from the naturalist Stephen Jay Gould [Gould, 1984]. Gould wrote an essay called '*What, if anything, is a zebra?*' which objective was to answer if zebras are white animals with black stripes or black animals with white stripes. The beauty of this title (originally from A. E. Wood, 1957) is that it allows to discuss something that is clear in all our minds (identity of Chemical Engineering), but which role and functions are continuously evolving to solve old problems, to answer new questions, and preparing the topics that will have to elucidate in the (uncertain) future.

Coming back to chemical engineering, it is commonly accepted that ChE has progressed through three different paradigms that promoted an equal number of metamorphosis:

1. The concept of unit operations, developed by Arthur Little (≈ 1910) as a notion to understand the underlying processes common to all chemical products, reactions, and machinery.
2. The inclusion of transport phenomena, thermodynamics and kinetics in the design of unit operations and the unified mathematical description of mass, energy, and momentum transfer [Bird *et al.*, 1960].
3. The extension of the domain to the interfaces with other disciplines, like biotechnology, medicine, food technology or environmental sciences. As a proof of

its impact, during the last two decades an increasing number of Chemical Engineering Departments have changed their names and now their affiliations include other areas, like "*environmental*" or "*biotechnology*".

1.1. Is chemical engineering in crisis?

Following with the evolutionary analogy of the title of this work, the transition of each one of the paradigms was a time for crisis, where different, and sometimes conflicting approaches, were explored (similar to the extinction periods of the evolution). Right after each new paradigms was established, a period of great richness and exciting findings in the new directions flourished (like the explosion of biological patterns and *phyla* that followed each extinction periods).

As time progresses, the technology we have been developing during the last decades give the impression to be incompatible with the problems that seems to arise (*e.g.*, global warming, water use, energy production…). In this context, the spatial scale of chemical engineering studies could range from nanotechnology (polymers, composites, electronics, coatings, catalysis, drug discovery, diagnostics…) to huge macrosystems (*e.g.* the boundaries of the system must include all upstream and downstream activities to study the environmental and sustainability problems). On the other hand, the temporal scale can cover from nanoseconds (nuclear reactions, particle interactions…), to centuries (*e.g.* stability of radioactive waste disposals, long-time effects of chemicals in the environmental and health…).

The need for change is evident and we must face those challenges. Sustainability and green technologies probably will be the cornerstone of chemical engineering of the future. The energy demand is rapidly increasing and the society will look at engineers to find a solution. One of our tasks is to prepare the technology to deal with the important issues that will be derived from the peak-oil, a very sensitive topic in geopolitics. At a given time, we will be forced to change from carbon-based energies to more sustainable sources. Technologies that now represent a small percentage of the power generation (wind, solar, hydro electrical, oceans, bio-fuels, geothermal, nuclear…) are expected to growth and should reach their maximum potential very quickly with a robustness similar to the actual resources. We can have a key role in the design, implementation of large scale facilities and improvements of these technologies.

If we consider the '*Big Four*' engineering fields (civil, mechanical, electrical, and chemical), our domain is numerically the smallest. However, on average, chemical engineers consistently top the list of the higher starting salary with a bachelor degree in the USA [NAS, 2007]. This issue highlights that, at least from the demand side, there is not such a crisis.

2. Benchmarking of chemical engineering

The work by Marrero [Marrero, 2008] focuses on obtaining data about the number of ChE graduates and their present and future employment on a world-wide basis. Table 1 indicates the increase of places to study Chemical Engineering. The study indicates imbalances between supply and demand sides. In addition the flow of professionals from one nation to another is affected by political constraints.

2.1. Comparison of chemical engineering undergraduate programs by multivariate statistical analysis

This study compares ChE undergraduate programs from all over the world (≈120 Universities) to obtain patterns based on the location and curricula. The work is based on the information available on the internet, and thus the bottleneck was the lack of specific course descriptions.

Table 1. Number of nations and schools with Chemical Engineering [Marrero, 2008].

	North America	South America	Africa	Asia	Europe	Mid-East
			# of nations			
1985	3	7	4	12	21	2
1995	3	5	4	11	21	5
2005	3	5	1	12	20	6
			# of schools			
1985	177	24	10	93	110	4
1995	181	20	10	79	116	9
2005	179	14	5	64	88	9

The ChE curricula were classified in 24 categories, divided in practical and theoretical classes. Multivariate statistical techniques (cluster analysis, linear principal component analysys and linear discriminant analysis) were applied to study the similarities between undergraduate programs. Data were pre-processed to use the percentages dedicated to each category, thus comparing the structure of the curriculum, not the duration [Jiménez and Mateo, 2003].

The average percentage of theoretical classes is around 55%, an average value for disciplines as ChE. Results from the principal component analysis show that most of the 24 categories do not group according to any predefined pattern, except mass and energy unit operations and mathematics and physics. When practical hours are considered, the cluster analysis shows sub-clusters, and there seems to be a certain similarity between Spanish speaking universities. In fact, most of the variance of the curricula depends on the practical classes, while theory does not have such a rich diversity.

The discriminant analysis shows that universities are not randomly distributed. Results show that over 80% of them were correctly allocated by the model to the geographical region they belong to (if the allocation were completely random, we would only expect a 14% correct allocation). The authors can not state if the similarities are due to cultural reasons rather than idiomatic (*e. g.*, USA, Australia). Also, the homogenization effect of accreditation programs (*e. g.*, ABET or Bologna process) may have some impact.

3. Approach at the School of Chemical Engineering

An educational model has been implemented at the School of Chemical Engineering (ETSEQ) at the University Rovira i Virgili (Tarragona, Spain), to enable ChE students to acquire and integrate technical knowledge encompassing social skills (teamwork, cooperation, planning, decision making, problem-solving abilities, communication skills…). In all courses, classical teaching, experimental and virtual laboratories and computer aided process engineering tools had to be coordinated to solve an open-ended project (*i. e.* with many valid approaches and different proper solutions).

3.1. Does faculty research improve undergraduate teaching?

Faculty members and educational researchers have been arguing for decades whether research supports undergraduate training. Those who agree with this statement cite many ways in which research can enrich teaching *a priori*, while those on the other side state that all studies consistently fail to show any measurable linkage between both activities [Prince *et al.*, 2007].

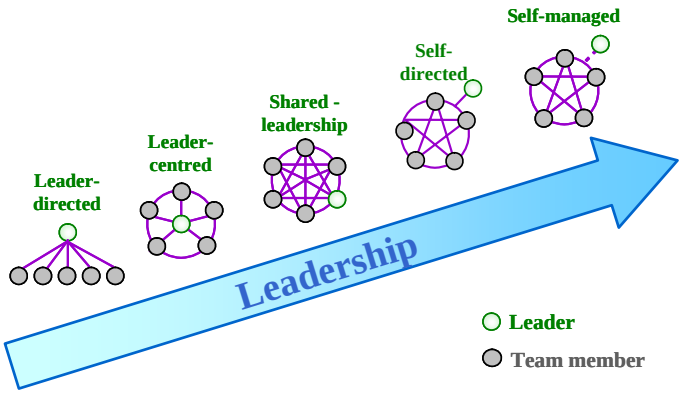


Figure 1. Team-work development [Witt *et al.*, 2006].

At the School of Chemical Engineering (ETSEQ), the research-teaching *nexus* have been exploited through two specific strategies: bringing research into the classroom (and to the society, through popular science) and involving undergraduates in research projects (‘*Research laboratory*’, a 75 hours compulsory subject in the 9th or 10th semester). The first topic is not covered by faculty members with research projects bringing their results into the classroom, as research contents is well over the level of most undergraduate students. To better realize the potential synergies between faculty research and undergraduate education, students are introduced into a research group to develop a specific task (typically around 40 students per year course this subject, a third of them in companies). During this period, students are exposed to solve routinely open-ended, ill-posed and ill-defined problems, figure out what they need to know, search scientific sources to obtain the missing information, present an intermediate report with the hypothesis, propose some test to find possible solutions, obtain the final results, explain them in a written report and defend them in an oral presentation.
Faculty members act as mere advisors and consultants so students are not subjected to passive teaching [J. Ziemlewski, 2009]. Undergraduate students are actively involved, enhance understanding, improve retention, become proficient in problem solving, self-directed learning, build decision criteria and team participation (cooperate rather than compete). Obviously, we do not expect students magically develop their entire individual potential within this subject but we see some progress (Figure 1).

3.2. 2. Project-based learning and teaching methodology

The educational model followed at the ETSEQ is based on the large-scale deployment of a project-based cooperative learning approach throughout the ChE curriculum: 1st, 2nd, 3rd, 4th and 5th courses [Witt *et al.*, 2006; Giralt *et al.*, 1994a, 1994b; Jiménez *et al.*, 2002, 2003].
The expertise in team development and change management methodologies, gained from nearly a decade of implementation work in Tarragona, has complemented the practice of the School of Chemical Engineering with empirical and pragmatic learning methodologies. A set of external training interventions has been designed to support the development of competencies by students. In this educational approach student teams grow from leader-directed teams in the first semester of the first academic year to self-directed (or empowered) teams in the fifth year. In this empowerment journey, fourth-year students play a key role as they act as facilitative leaders of first-year and second-year project teams, adjusting their leadership role according to the team development stage (Figure 1).

For example, two eight hours modules are taught during the first two weeks of each term to first-year students. The modules are structured to take into account the factors needed for students to understand and commit themselves to change. The first-year students' evaluation shows that the module helps them:

- To identify what they need to accomplish to gain future employment as Chemical Engineers, thus settling the patterns to become life-long learners.
- To understand, from a practical point of view, what an integrated design project consists of and what the benefits of teamwork are.
- To realize that the integrated project and the related teamwork are great opportunities to acquire competencies that are essential in today's workplace.

In this way, hands-on teamwork training is implemented through specific seminars. These modules deal with different teamwork-related issues such as, leadership, team capabilities, common purpose, team norms, communication, conflicts, team operating procedures, member integration, team evaluation… These activities are distributed across the curricula, considering a long-term deployment. The activities are organized with optional attendance; nevertheless our experience is that more than 80% of students follow these modules, which can be validated as official credits (*i. e.* elective courses). In each session, two instructors that play different roles are available: one leads the session while the other acts as facilitator.

The core of the competency-based educational model is client orientation. The need to satisfy clients and to adapt to their changing needs triggers the development of competencies related to the transformation of the students (versatility, entrepreneurship and innovation...), of the organization (facilitative leadership, teamwork, cooperation…), and of the institution (organizational development, organizational performance, organizational leadership…). For a detailed distribution and allocation of each capability in the undergraduate ChE program, check the team development matrix that is explained in detail in Table 2.

Table 2. How to assess the team-work capabilities: team development matrix deployment over the ChE curriculum [Witt *et al.*, 2006].

Responsibility for work activity Leader / Shared / Team members	Leader directed	Leader centered	Shared leadership	Self-directed	Self-managed
1. Formulate team goals					
2. Communicate outside the team					
3. Decide on team rules					
4. Resolve conflicts within the team					
5. Solve problems					
6. Interact with customers					
7. Planning work, meetings, tasks, deliverables, etc.					
8. Provide feedback to other team members					
9. Improve/define work processes					
10. Allocate work assignments					
11. Plan training and development for team and members					
12. Allocate resources within the budget					
13. Determine recognition and rewards					
14. Measure team's progress towards the goals					
15. Decide on leadership within the team					
16. Monitor expenses to be within a budget					
17. Select new team members					
18. Evaluate performance					
19. Distribute and allocate scores					
20. Integrate and train new members					

4. Conclusions

The training in non-technical capabilities helps our students to set the pattern to become successful life-long learners. Student's main objections to the project-based cooperative learning approach were the excessive time devoted to the project. In addition they demand more effort in supervising, in particular in the initial steps of the project development. We realize that, as students are not used to this kind of teaching, at the beginning of the project more continuous help and guidance is needed than in traditional teaching methodologies. However, the lack of information forces students to use their own initiative to solve the open-ended problems (*i. e.* there are many valid approaches and solutions). For example, the inclusion of computer aided process engineering tools (*i. e.* virtual laboratories) help students to acquire the insights of the unit operations, since mathematical models are not useful for a qualitative interpretation of how design variables influences the unit operation performance.

Preliminary results show that student attendance has increased, that drop out has decreased, that more professors act as facilitators, and that active-oriented and student-centered educational methodologies are increasingly applied. The academic staff also needs to improve our performance as a teamwork model role for students, as sometimes we are surpassed by everyday research tasks and teaching is not in the top list of priorities. To sum-up, our opinion is that the benefits of applying this methodology largely exceed the tremendous effort required.

References

R. B. Bird, W. E. Stewart and E. N. Lightfoot, 1960, Transport phenomena. J. Wiley and Sons, New York, USA.

F. Giralt, M. Medir, H. Thier and F. X. Grau, 1994a, A holistic approach to the ChE education. Part 1. Professional and issue-oriented approach. Chemical Engineering Education, 28, 122-127.

F. Giralt, A. Fabregat, X. Farriol, F. X. Grau, J. Giralt and M. Medir, 1994b, A holistic approach to the ChE education. Part 2. approach at the introductory level. Chem. Eng. Ed., 28, 204-213.

S. J. Gould, 1984, What, if anything, is a zebra? Essay in Hen's teeth and horse's toes. Norton Paperback, New York, USA.

L. Jiménez, J. Font and X. Farriol, 2003, Unit operations laboratory using ill-posed problems, International Journal of Engineering Education, 19 (5), 717-720.

L. Jiménez and J. M. Mateo, 2003, Comparison and characterization of chemical engineering undergraduate programs by multivariate analysis, European Congress on Chemical Engineering, Granada, Spain.

L. Jiménez, J. Font, J. Bonet and X. Farriol, 2002, A holistic unit operations laboratory, Chemical Engineering Education, 36, 150-155.

National Academy of Sciences, 2007, International benchmarking of U.S. chemical engineering research competitiveness in brief. Availble at (April 2009): http://dels.nas.edu/dels/rpt_briefs/benchmarking_chem_eng_brief_final.pdf.

T. R. Marrero, 2008, Global chemical engineering workforce statistics, AIChE anual meeting, Philadelphia, USA.

M. J. Prince, R. M. Felder, R. Brent, 2007, Does faculty research improve undergraduate teaching? An analysis of existing and potential synergies, Journal of Engineering Education, 96 (4), 283-294.

H. J. Witt, J. R. Alabart, F. Giralt, J. Herrero, L. Vernis and M. Medir, 2006, A competency-based educational model in a chemical engineering school. Internationl Journal of Engineering Education, 22 (2), 218-235.

A. E. Wood, 1957, What, if anything, is a rabbit? Evolution, 11, 417-425.

J. Ziemlewski, 2009, Designing the new global chemical engineer, Chemical Engineering Progress, 105 (2), 6-10.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Information Modelling: Industrial Standards for Integrated Plant Management

Zofia Lukszo

Department of Energy and Industry, Faculty of Technology, Policy and Management, Delft University of Technology, Jaffalaan 5, 2628 BX Delft, the Netherlands

Abstract

The paper describes the Master course "Integrated Plant Management" developed for the students of the Faculty Technology, Policy and Management at the Delft University of Technology. During the regular classes students learn advanced analytical methods for problem solving of long and short-term planning, forecasting and scheduling problems as well as process optimization. During the instructions they make an IDEF0 model of an industrial plant according to the standards ISA88 and ISA95. In this model the same activities as discussed during regular classes are placed in the context of manufacturing execution systems. This helps students in understanding the management complexity of an industrial plant and in contributing adequately to implement many challenging simulation and/or optimization solutions in real-life cases.

Keywords: Integrated Plant Management, Information Modelling, IDEF0, Industrial standards

1. Introduction

Process system engineering research into methods to support an (eco-) efficient process design and operation is gaining momentum since many years. Many contributions to the PSE research area concentrate on a specific aspect of the process design and operation: modelling and optimizing a process, plant or even an enterprise, defining more effective planning and scheduling algorithms, designing more advanced controllers etc. Some of these contributions have a strong theoretical character and therefore their practical applicability is rather limited.

Other contributions, often related to real-world industrial problems, also seldom find a way to be implemented in the industry and to contribute to the improvement of the overall plant performance. It is striking that the organizational aspects and the applicability of the valuable theoretical results are hardly explored. Academic courses aimed at providing knowledge and understanding of a plant operation in such a way that the challenges imposed by the economic, environmental and social sustainability are made more transparent, may change this situation. We strongly believe that scientific methods for first-principle modelling, rigorous optimisation and advanced control to be applicable in the industrial practice should be supported by a clear picture of the whole enterprise and working processes taking place there.

This paper contributes to the area of the integrated plant management. It describes the Master course "Integrated Plant Management" developed for the students of the Faculty Technology, Policy and Management at the Delft University of Technology. During the regular classes students learn advanced analytical methods for problem solving of long and short-term planning, forecasting and scheduling problems as well as process optimization (Edgar, 2001; Pinedo 2001; Grossmann, 2002; Biegler 2004). During the instructions they make an IDEF0 (Integration DEFinition of function modelling) model

of an industrial plant according to the industrial standards ISA88 and ISA95 (ISA, 1995; ANSI 2000, 2001, 2005). The activities as discussed during regular classes, e.g. scheduling, recipe optimization etc are placed during the modelling assignment in the context of manufacturing execution systems. At the end of the course an activity model "run a plant" is made and the students know who is involved in operational activities, where the necessary information is coming from, what is the interaction between these activities and to what extent changes at one system level contribute to the overall plant performance. This helps students in understanding the management complexity of an industrial plant and in contributing adequately to implementation of many challenging simulation and/or optimization solutions in real-life cases.
Section 2 introduces industrial standards ISA88 and ISA95 aimed at supporting operations management. In section 3, the activity modelling approach IDEF0 is described, followed by an illustrative example of activity decomposition for planning, scheduling and control functions. Finally, Section 4 gives some concluding remarks.

2. Industrial Standards for Integrated Plant Management

In the process industries there are very large financial incentives for fundamental business processes, e.g. planning, scheduling and control activities, to function in a coordinated and integrated fashion (Shobrys, 2002). Nevertheless, many industrial companies have not achieved desired integration in spite of multiple initiatives. Operating in a rapidly changing world of global economy and more short-term dynamics they continuously search for opportunities to improve their competitive position: the production processes produce products more efficiently and the internal methods of operation enable to be more effective. The awareness of the importance of environmental affairs and objectives are also set to the agenda of many enterprises. Mostly, during the improvement projects the companies concentrate on improving one task without taking sufficiently into account interactions with other activities and with the surroundings. Moreover, the formulated company objectives and policy are very often communicated inside the organization insufficiently. The language the people talk is different: top managers talk about profit and continuity, process people about quality and quantity and the environmental department about emissions and waste.
To contribute to the integrated plant management the modern PSE concepts should be properly placed in the total manufacturing execution system, so that it will be clear who is responsible in the plant for the particular activity, what information and from whom is needed to perform the activity, which other activity and to what extent depends on the generated results, and what is a possible contribution of the activity to the overall plant performance.
The ISA88 and ISA95 standards, published by the Instrumentation, Systems, and Automation Society, provide standard terminology and models for design as well as operation of (batch) control systems and operations management activities (ISA, 1995; ANSI 2001, ANSI 2005). Figure 1 presents the scope of both standards. The aim of the ISA88 standard is to provide a modular framework for recipe development and management. The aim of the ISA95 standard is to improve information exchange by providing a framework for the integration of the enterprise functions as strategic and tactical forecasting and planning, scheduling, quality assurance, maintenance management, procurement etc with the control and process execution layer as described by ISA88.

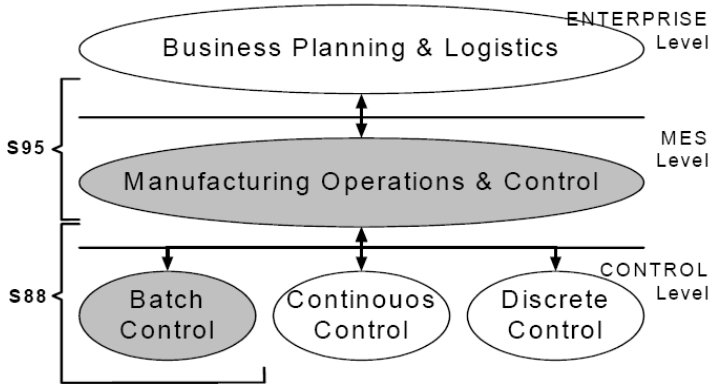


Figure 1. Scope of the ISA-88 and ISA-95 standards (grey areas)

The activities of manufacturing execution system (MES) layer are those activities of a production facility that coordinate the personnel, equipment, material, and energy in the conversion of raw materials and/or parts into products. Operations management including activities that may be performed by physical equipment, human effort, and information systems, corresponds to the activity model defined in the Part 1 of the ISA 95 standard. It defines 10 generic functions in an enterprise-control model and two additional functions as presented in Figure 2. Further, it describes point by point the tasks of each function. It should be stressed, that the model does not say anything about the importance and complexity of the functions. For example, material control for automotive industry is a very complex matter comparing to a fairly simple task in a water distribution company (Scholten, 2007).

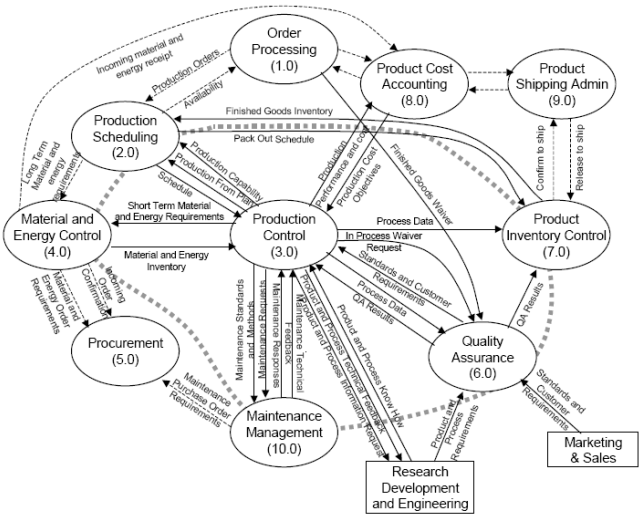


Figure 2. Functional ISA – 95 enterprise-control model

The bold dotted line in Figure 2 illustrates the boundary of the enterprise-control interface between the Enterprise Level and the MES Level (Manufacturing Execution System). The function *Production Scheduling* is placed on this boundary. At the Enterprise Level it generates production plans with a time horizon of months or weeks: master production schedule. On MES Level, detailed scheduling takes place based on days or hours, taking into account product demand pattern, available equipment, storage capacity and personnel, recipes, the relationship between cleaning time and product combination, available utility profile etc. The activity *Production Control* is a pure Control Level function. Every function exchanges information with other functions. ISA-95 groups this information into information flows; each with its own name and content, such as Schedule, Production Performance and Cost, QA results etc.
For the students is even important to understand how to perform scheduling activities by optimizing one or more adequate criteria (Mendez, 2006) resulting in an optimal schedule as to know all relevant information flows to create their own view on integrated plant management.

3. Key Business Processes and Their Mutual Relationships

Plant modelling according to the ISA-S88 and ISA-S95 standards by using the so called SADT (Structured Analysis and Design Technique) techniques and especially IDEF0 diagrams is very useful to support the understanding of integrated plant management. IDEF0 (Integration DEFinition of function modelling) offers a well-structured method for activity and information-flow analysis, focussing on the hierarchical decomposition of activities as well as on the interaction flows between activities (IDEF0, 1993). The primary objectives of the IDEF0 modelling technique are (Heijnen, 2006, Fuchnino, 2008):

- To provide means for complete and consistent hierarchical modelling of activities that take place in a system, and the data and objects that inter-relate those activities.
- To provide a modelling language that is generic (applicable for analysing systems and subject areas of varying purpose, scope and complexity); rigorous and precise (for producing correct, usable models) and concise (to facilitate understanding, communication, consensus and validation).

The whole system – in this case an industrial plant - is modelled in a clear way (there are only two symbols used: boxes and arrows) as a set of interrelated and nested activities. The first activity to be designed is the activity that describes the system itself, i.e. run a plant, and is called the context activity, see Figure 3. In IDEF0 not only the inputs, which are transformed into the outputs, are modelled, but also the controls and mechanisms. Controls (are fed to the top of an activity) are the objects that govern the manner in which inputs are transformed; yet they are not themselves transformed by the activity, e.g. in Figure 3 Market Information. Mechanisms (are fed to the bottom of an activity) are those objects that actually perform the transformation of inputs to outputs, yet they are not themselves transformed by the activity, e.g. software tools and methods, equipment and people. ICOM is an acronym for the categories of information presented in IDEF0 diagrams and represents four types of arrows: Inputs, Controls, Outputs and Mechanisms. The inputs are defined as the representation of the material or information that is consumed or transformed by the activity to produce outputs. Controls influence the activity without being transformed or consumed, they are necessary to perform the activity; they are often in the form of policies or procedures, and we extend them with the objectives. Outputs are the materials or information produced by an activity. Mechanisms are the resources that help to perform it. An output from one activity can

be an input to other activity. In this way, activities can be combined in a chain or a network.

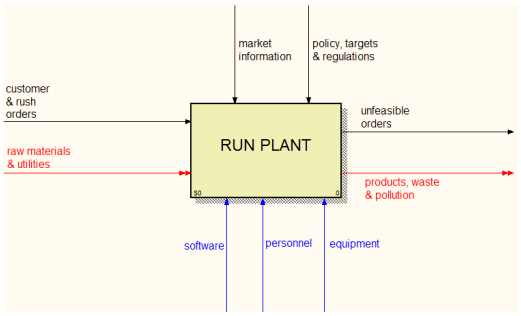


Figure 3. Context diagram 'Run a Plant". Software: AllFusion Process Modeller from Computer Associates International

The context activity, as presented in Figure 3, defined at the highest level can be decomposed in a number of activities, which collectively should achieve the main objective of the plant. In this case the top activity "Run a Plant" can be decomposed further, e.g. into "Do Order Processing and Procurement", "Do Production Planning, Scheduling and Control, "Do Inventory Management", "Perform Product Cost Accounting", "Do Product Shipping Administration", "Do Quality Assurance". Next, each activity can be decomposed further, e.g. "Do Production Planning, Scheduling and Control". In the decomposed model an output from one activity can be an input to another activity. The feedback loops are easily understood. In this way activities can be combined in a chain or network.

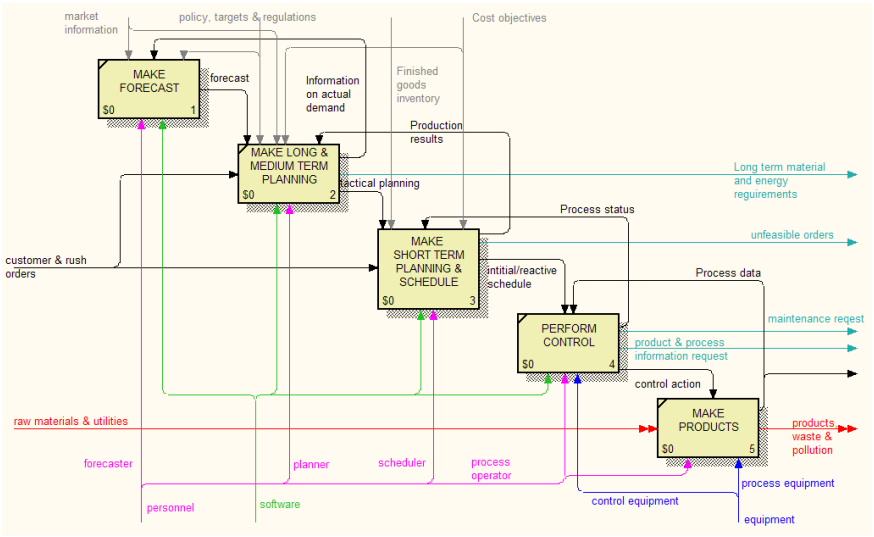


Figure 4. Decomposition of "Do Planning Scheduling and Control"

4. Final Remarks

Modelling enterprise activities and their interdependencies, and relating them to the operational objectives makes it possible to visualize how improvements or changes in

one activity interact with other activities and which results could be expected. This visualization supports the decision, which activities at the operational level contribute mostly to the overall objective of the plant, in other words which ones are most effective. Improvements in the efficiency of those activities will be most effective.
In an ongoing process together with partners from the industry, the students apply this method successfully to industrial projects (Salverda, 2006; Klompenhouwer, 2007, Stekelenburg, 2008). Through many industrial projects, effectiveness of the proposed procedure is confirmed. Relaying on these results, the activity modelling approach is shown promising in the implementation phase of PSE solutions to real world applications in the process industry.

Acknowledgment

The work described in this paper is supported by the Next Generation Infrastructures Foundation http://www.nginfra.nl/.

References

ANSI/ISA–S95.00.01–2000, *Enterprise-Control System Integration Part 1: Models and Terminology*, ISBN: 1-55617-727-5, 2000, ISA–The Instrumentation, Systems, and Automation Society.

ANSI/ISA–95.00.02–2001, *Enterprise-Control System Integration Part 2: Object Model Attributes*, ISBN: 1-55617-773-9, 2001, ISA–The Instrumentation, Systems, and Automation Society.

ANSI/ISA-95.00.03-2005, *Enterprise-Control System Integration Part3: Activity Models of Manufacturing Operations Management*, ISBN: 1-55617-955-3, 2005, ISA–The Instrumentation, Systems, and Automation Society.

L.T. Biegler, I.E. Grossmann, 2004, *Retrospective on Optimization*, Comp. and Chem. Eng., 28, pp. 1169-1192.

T.F. Edgar, D.M. Himmelblau, L.S. Lasdon, 2001,*Optimisation of Chemical Processes*, McGraw-Hill Companies, Inc.: New York.

T. Fuchino, Y. Shimada, M. Miyazawa, 2008, *Business Process Model for Knowledge management in Plant Maintenance*, ESCAPE 18.

I.E. Grossmann, 2002, *Review of Nonlinear Mixed-Integer and Disjunctive Programming Techniques*, Optimization and Eng, vol 3.

P. Heijnen, Z. Lukszo, 2006, *Continuous improvement of batch wise operation—a decision support framework*, Production Planning & Control, 17 (4), pp. 355–366.

C.A. Méndez, *et al*., 2006, *State-of-the-art Review of Optimization Methods for Short-Term Scheduling of Batch Processes*, Computers and Chemical Engineering, vol 30.

M. Pinedo, 2001, *Scheduling: Theory, Algorithms, and Systems*, Prentice Hall.

B. Scholten, 2007, *The road to integration*, ISA, Reserarch Triangle Park.

D. E. Shobrys, D. C. White, 2002, *Planning, scheduling and control systems: why cannot they work together*, Computers and Chemical Engineering 26, pp. 149–160.

ISA Standard ISA-S88.01, 1995, *Batch Control. Models and Terminology*, ISA, Standard.

Integration Definition for Function Modelling (IDEF0), 1993, Publication 183, FIPS PUBS, 1993

M. Salverda, Z. Lukszo, P. Bosman, 2006, *A decision support tool for process optimization of sulphur free diesel production*, ESCAPE 16.

A.M. Klompenhouwer, Z. Lukszo, F. Janssen, 2007, *Analyzing the relationship between manufacturing lead-times and line flexibility – the Line Flexibility Model*, ESCAPE 17.

R.Ytsma, Z.Lukszo, R. Maliepaard, 2009, *Sustainable reduction of dredging fleet emission*, ESCAPE 19.

D. Stekelenburg, Z. Lukszo, J. Lowe, 2008, *Improvement of the production process of leached optical fibers in a technological and organizational context*, ESCAPE 18.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

An experimental approach to complement Process Systems Engineering learning

Roger J. Zemp, Renata Waki and Flávio V. da Silva

School of Chemical Engineering
University of Campinas, UNICAMP
P.O. Box 6066, 13083-970, Campinas-SP, Brazil

Abstract

This paper describes a proposal for complementing Process Systems Engineering (PSE) undergraduate learning. Traditional teaching of PSE is normally restricted to modeling and computer simulations, with very little contact with actual processes. In this paper we describe how laboratory experiments can be combined with modeling and simulation tools to enhance the overall understanding of Process Systems Engineering.

Keywords: experimental process setup, modeling and simulation, undergraduate teaching

1. Introduction

Traditional teaching of PSE is typically focused on modeling and computer simulations, with very little contact with actual processes. We felt that applying theoretical concepts to practical problems would be very welcome by the students and improve their understanding of PSE. On the other hand, experimental work done by chemical engineering students in labs is more than often restricted to collecting data, doing some data processing and comparing the results to those expected by applying the appropriate theory.

We were looking for a way to combine experimental work with modeling and simulation tools. However, this should be done keeping low cost and low complexity in mind. This ruled-out experiments requiring measurement of composition, at least for this first tentative step. And since flow and temperature measurements are quite easy to be achieved, the heat exchanger experiment was chosen for our proposal.

2. The proposed approach

The heat exchanger experiment of our undergraduate lab suffered a major change a few years ago when the old double-pipe heat exchanger was scraped and a new small plate heat exchanger installed. Additionally, a hot-water storage heater was also installed to provide the hot fluid, while as cold fluid tap water was used. Flow measurements were done by variable area flow meters, and temperatures by thermocouples connected to a digital display. However, despite the changes in equipment the experimental work carried-out by the students was still restricted to the measurement of heat balance and heat transfer coefficient, and comparison with expected values. For our purpose this setup was far from ideal, and a few modifications were introduced.

2.1. The equipment

The new experimental setup consists of a process system (a heat exchanger) and a utility system (a cooling tower), which provides the cold water for the heat exchanger.

The heat exchanger is of the plate-type, with 19 plates and approx. 0.4 m^2 of area. Hot water (up to 65°C) is provided by a water storage heater.

The cooling tower is of the countercurrent type, made of three stacked PVC cylinders (40 cm diameter), using plastic mesh as filling in the middle cylinder. The bottom tank stores the cooled water, while the top cylinder contains the fan. Heated cooling water from the heat exchanger can be bypassed through a finned heat exchanger to pre-cool the water (as would be the case in an actual plant using air-coolers).

Pumps and valves are used to control the flow of water through the system.

2.2. Instrumentation

A set of industrial-standard instrumentation was used for measuring process variables: flows of hot and cold water (magnetic flow-sensors), fluid temperatures at inlet and outlet of exchanger, air-cooler and cooling tower (thermocouples), air humidity at inlet and outlet of cooling tower (humidity sensors), and pressure drop of cold water across the heat exchanger (differential pressure sensor). All instruments are connected to an industrial data-logging device using using 4-20 mA standards. Air flow through the tower was measured using a hot-wire anemometer.

Pictures of the actual equipments are shown in Figure 1.

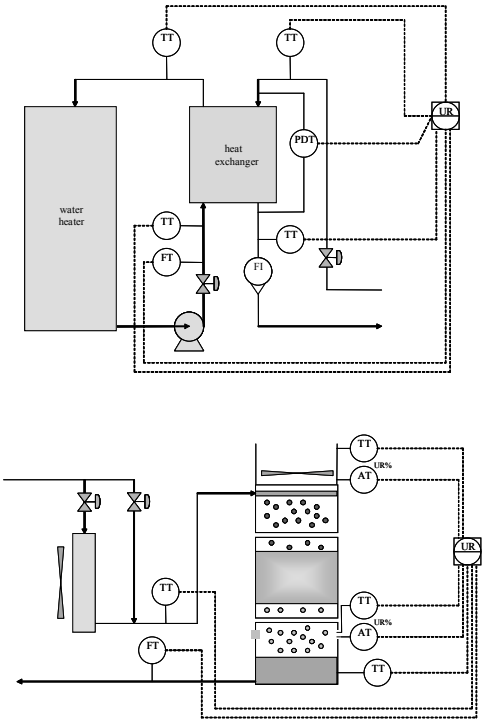


Figure 1 – The process (heat exchanger) and the utility system (cooling tower), with full instrumentation.

2.3. Data Monitoring

The measured process parameters are sent from the datalogger to a computer using the MODBUS/RS485 standard, and displayed to the user as process flowsheets. The flowsheets and additional screens were developed using the industrial-standard supervisory control and data acquisition system InduSoft Web Studio (Figure 2). Data can be stored in files for further off-line processing.

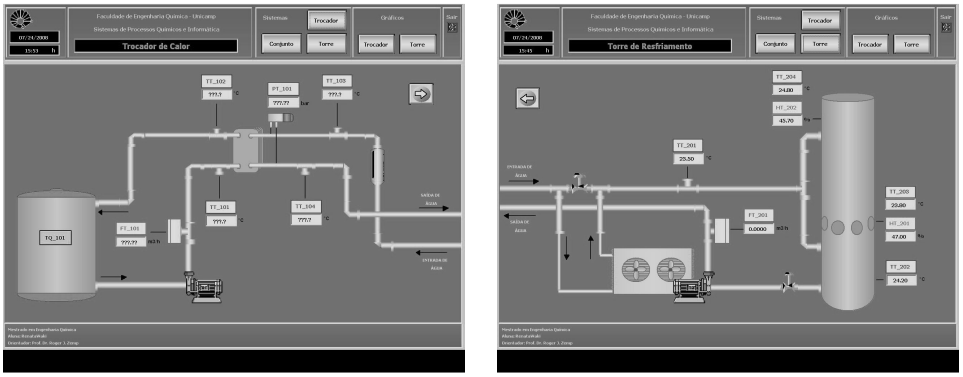


Figure 2 – Process flowsheets for the heat exchanger and the cooling tower

The overall benefits of the chosen experimental set-up are:

- The process is a small-scale example of a typical processing plant (a process and a utility system).
- Instrumentation and monitoring is done using industrial standard devices.
- Process parameter can be continuously monitored and stored.

3. Experiments

Once the experimental setup was finished, the practices to be used by the students were developed and implemented. As stated in the introduction, our goal was to combine experimental work with process systems engineering concepts, so we had to "set a scene" for each experiment. This was done by proposing an industrial problem to be solved using modeling and simulation techniques, but where some of the data required was not available and had to be determined experimentally. This was a major change in direction away from our traditional approach to lab teaching, where the experiment was the main issue. In our new approach, the problem to be solved was the main objective, and the lab providing one or more steps of the solution procedure.

Some of the proposed experiments are now presented, with the discussion of the relevant PSE techniques used.

3.1. Industrial heat transfer

The problem proposed to the students in one of designing a plate heat exchanger for a given process, with minimum cost. The process data (fluids, flows, temperatures) are known, as are the available pressure drops. The problem is that no correlation is available for estimating heat transfer coefficients, but they are to assume that the small-scale exchanger in the lab can be used to provide a suitable correlation.

As the first step the students have to implement a mathematical model of a plate heat exchanger, where for a given set of temperatures and flows the lab exchanger is simulated. This involves computing the energy balance and the actual overall heat transfer coefficient. Once their simulation is working, the students use the lab setup to perform a series of experimental runs to gather temperature, flow and pressure drop data. The next step is to propose a procedure to obtain the set of parameters that be fits the model to the experimental data. For this task a proper correlation for the heat transfer coefficient is chosen, normally of the form (Hewitt et al., 1994; Hewitt, G., 1993):

$$Nu = aRe^{b}Pr^{c} \qquad \textit{Eq. 1}$$

Due to the restriction of only one fluid being available (water), it is suggested to use c=0.4. The students than have to implement their procedure to bet fit the proposed correlation to the experimental data:

$$\min \sum \left(U_{\text{experimental}} - U_{\text{predicted}}\right)^2 \qquad \textit{Eq. 2}$$

A similar approach is used to obtain a correlation for the friction factor of the pressure drop equation. After a proper correlation for the heat transfer is obtained, the final step is to design a heat exchanger for the given problem. Here, different types of problems can be proposed: restrictions on exchanger area or number of plates, available pressure drop, required flow layout and number of passes. Normally the students have to resort to non-linear programming procedures to obtain a feasible result (Biegler et al., 1997; Edgar et al., 2001).
This type of task complements very well the concepts learned in the modeling, simulation and unit operation courses. Figure 3 shows a typical result from the correlation fitting procedure, with the points located close to the 45° diagonal.

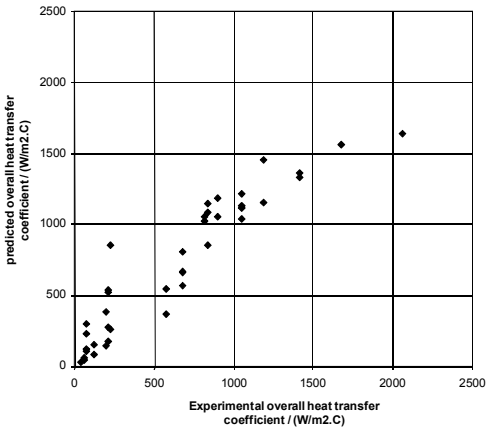


Figure 3 – comparison between measured and estimated global heat transfer coefficients

3.2. Process simulation and identification

In this problem the students are faced with the task to provide the Company's Engineering Department with a mathematical model of an existing heat exchanger. This model is supposed to be used to design a control system for a network of similar exchangers.

The actual work required is to build a mathematical model for the transient behavior of a countercurrent plate exchanger, and program it on a computer using an appropriate modeling system, MatLab for example (Silebi and Schiesser,1992). This requires the solution of distributed-parameter differential equations, and use of numerical methods. The students are then asked to gather experimental transient data in the lab, for several different process conditions, and compare these to the simulations results. Adjustments of the model are then done to match the model to actual data as closely as possible. The students have to deal with uncertainties in measured data, and be able to propose a method of fitting the transient model to the experimental data. Figure 4 shows a typical result from a simulation program, showing the behavior of the exchanger (exit temperature) for different process conditions and solution parameters.

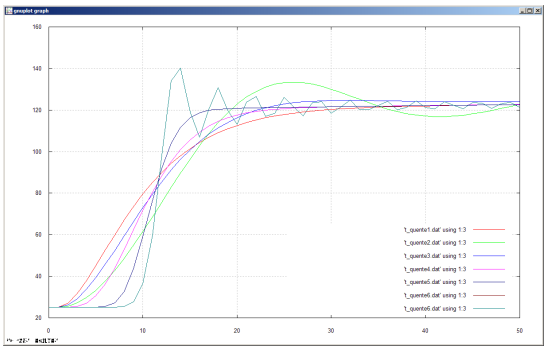


Figure 4 – transient behavior of heat exchanger

3.3. Interaction between process (heat exchanger) and utility system (cooling tower)

While the previous experiments dealt with a single unit operation, in this one the students are faced with the interaction between two unit operations: a process (heat exchanger) and a utility system (cooling tower).

The students are given the task to analyze the behavior of an industrial cooling tower under changes of process heat load. The details of the tower are all given, if an existing design, or need to be determined (if a new design). In both cases, however, the mass transfer coefficient is not known, and as in previous cases, a lab setup with similar tower filling is available for tests.

The students are initially required to implement a program that simulates the counter-current cooling tower. This requires the use of thermodynamics (psychrometry), and modeling techniques for countercurrent unit operations, with equilibrium and operating lines (Figure 5). The next step is to use the experimental setup to analyze the behavior of the cooling tower under different heat loads and cooling water return temperatures. The experimental data is the used to fit a very simple correlation for the overall mass transfer coefficient for the tower filling (Hensley, 2006).

In the second part of the experiment the students use the simulations and data obtained in the first part to solve the proposed industrial problem. Again, different problems can be proposed, with different process constraints or cost functions, so that each group of students has a different task to work on. In this experiment the students have to resort to modeling techniques involving differential equations / integration and optimization.

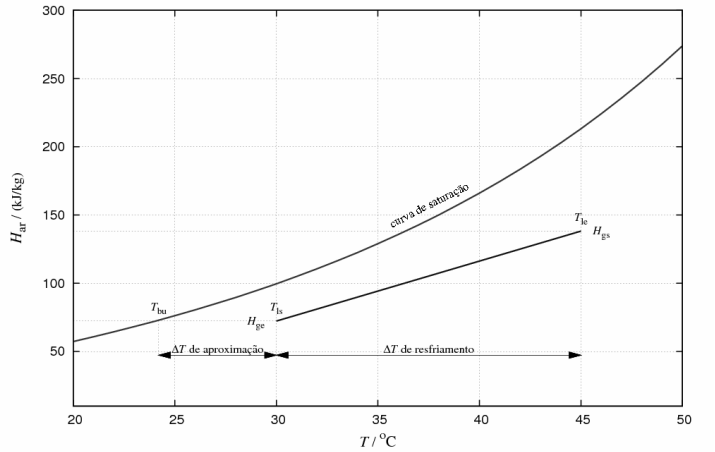


Figure 5 – Cooling tower operating line and saturation line

4. Conclusions

This paper describes the implementation of a new approach to combine experimental chemical engineering lab work with process systems engineering. Overall, this new experimental setup allows for the students to apply their knowledge of PSE in a more practice-oriented environment, dealing with actual process equipments. PSE tools like modeling and optimization are fitted into a problem-solving environment. Numerical tools like integration, differential equations and non-linear optimization are use to analyze and model actual processes. Instrumentation and data monitoring are done using industry-standard equipment.

Based on our assessment of the students that have been working under the new methodology, the proposed sequence of modeling a process, fitting the model parameters using experimental data, and using the model to optimize or design a given process has been shown to be very efficient in helping the students understand and apply process systems methodologies.

References

Biegler, L., Grossmann, I., Westerberg, A. 1997, *Systematic Methods of Chemical Process Design*. Upper Saddle River: Prentice Hall PTR.

Edgar, T., Himmelblau, D., Lasdon, L., 2001, *Optimization of Chemical Processes*. 2nd Edition, New York: McGraw-Hill Publishing Co.

Hensley, J. (editor), 2006, *Cooling Tower Fundamentals*, SPX Cooling Technologies, Inc., Overland Park, Kansas USA

Hewitt, G., 1993, Heat Exchanger of Design Handbook. S.L.: Begell House.

Hewitt, G., Shires, G., Bott, T. 1994, Process Heat Transfer. Boca Raton: CRC Press.

Silebi, C., Schiesser, W. ,1992, Dynamic Modeling of Transport Process Systems. Boston: Academic Press.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Cooperative Weblab: A Tool for Cooperative Learning in Chemical Engineering in a Global Environment

Galo A.C., Le Roux[a], Giann B. Reis[b], Charles D.F. de Jesus[b], Roberto C. Giordano[b], Antonio J.G. Cruz[b], Paulo F. Moreira Jr.[a], Claudio, A.O., Nascimento[a], Luiz V. Loureiro[b]

[a] *Chemical Engineering Department, Polytechnic School of the University of São Paulo, Av. Prof. Luciano Gualberto, trav. 3, 380, 05508-900, São Paulo, SP, Brazil*
[b] *Chemical Engineering Department, Federal University of São Carlos, Rodovia Washington Luis (SP 310), km 235 – 135 5 905, São Carlos, SP, Brazil*

Abstract

Weblabs are defined as a set of web-based physical laboratories that allows remote access and control in real time. Since the last decade, Weblabs are being more and more employed in many educational institutions around the world. In São Paulo state (Brazil) a "Cluster of Weblabs in Chemical and Biochemical Process Engineering" was implemented. This project has received a grant from FAPESP (State of São Paulo Research Agency). The experimental setups are physically placed in laboratories in São Paulo, São Carlos, Ribeirão Preto and Campinas, that are up to 250 km apart. This paper presents the implementation of two Cooperative Weblab (CW) experiments. The main concept behind the CW is to develop a set of experiments for undergraduate students that should be performed by "mixed teams" located in different institutions at the same time. Collaboration is achieved by gathering participants into working groups that are asked to simultaneously solve a technical problem, for which a weblab experiment is available. This procedure emulates challenges that will frequently take place in their future professional lives. Each group is supervised by a local instructor. Video conference software is used along the experiments (the institutions are connected through a high speed network, named KyaTera). All long the experiments, the students are encouraged to interact, exchange information and opinions on the phenomena that take place during the experimental practice. At the end of the session, the data collected are shared by the students and they produce a unique report for both groups. The use of this approach fosters interest in students. It is clear that the CW offer real benefits to chemical engineering education. It is now necessary to spread the dimension of the cluster and to increase the interaction with groups from other countries.

Keywords: Internet, experiment, teaching, intercultural

1. Introduction

In the year of 2004, the Foundation for the Support of Research in São Paulo State (FAPESP) established a program named Information Technology in the Development of Advanced Internet (TIDIA program). One of the projects of this program was named KyaTera, an acronym of Kya (net in "Tupi-Guarani", the language of the aboriginal people in Brazil) and Tera (from Terabits per seconds). KyaTera is a cooperative project

consisting in an optical fiber network designed for the research and development of high speed connections, linking various laboratories focusing on the study, development and demonstration technology and application on Advanced Internet. One of the projects of KyaTera is the "Cluster of Weblabs for Chemical and Biochemical Process Engineering". This project intends to build a network of undergraduate chemical engineering experiments that should be performed by "mixed teams", integrating students of four different Universities of the São Paulo state (EPUSP, Escola Politécnica da Universidade de São Paulo – the hub of the Weblabs network, UFSCar, Universidade Federal de São Carlos, USP-RP, Universidade de São Paulo, campus Ribeirão Preto and UNICAMP, Universidade de Campinas). Further information can be found at www.kyatera.fapesp.br. Figure 1 presents a picture of the network dark fibers implemented in the São Paulo state by KyaTera project.

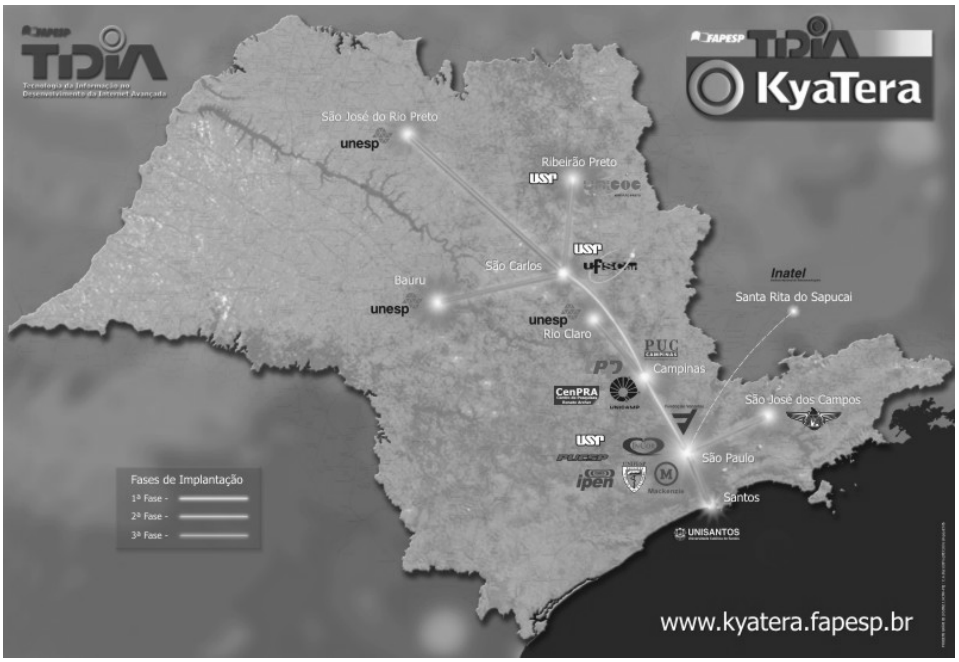


Figure 1. Dark fiber connecting researchers laboratories in São Paulo state, Brazil.

Weblabs are defined as a set of web-based physical laboratories that permits remote and control access in real time (see Figure 2). During the experiments users do not need to be in front of the experimental apparatus to set-up experimental conditions. The changes are remotely made from computers with a graphical interface that permits access to the main variables of the process. At the end of the experiment, data can be transferred to the user.

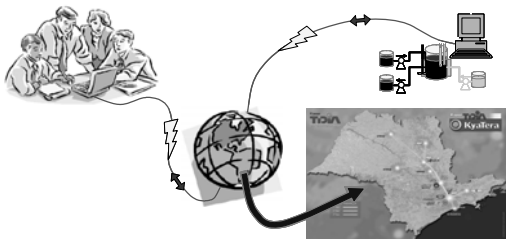


Figure 2. A sketch of the remote Weblab operation.

This paper presents the idea of implementing Cooperative Weblab (CW) experiments. The main concept behind the CW is to develop a set of experiments for undergraduate students that should be performed by "mixed teams" located in different institutions at the same time. The main idea behind this tool is to promote an increasing cooperative learning. The implementation of this tool encompasses the agreement of two or more institutions. A great effort is necessary in order to reach this kind of activity in the curricula of Chemical Engineering courses.

2. Review and Previous Works

The application of weblabs experiments in chemical engineering education is quite recent. The concept was introduced by Prof. Clark Colton at the heat transfer course of the undergraduate program at the Massachusetts Institute of Technology (MIT). The weblab experimental setup consists of a plate heat exchange with all necessary instrumentation and control devices. This weblab was shared with students from the chemical engineering from the University of Cambridge, UK, during their process control courses (Selmer et al. 2005). Another set of weblabs was built in University Leipzig by Ralf Moros under supervision of Helmut Rapp for training courses (personal communication).

SHIN et al (2002) were amidst the first authors to report a virtual web-based lab for unit operations course as a chemical engineering education tools. A collaborative web-based experiment was presented by GILLET et al. (2005) that provides more flexibility for students performing experiments in automatic control. They integrate learning environment that enables the recollection and sharing of preparatory notes and experimental results.

Two international workshops were organized one at MIT, in January 2005, and the other one at University of Cambridge, UK, in July 2005, with the participation of representatives from many institutions from various countries, e.g. Monterey Institute of Technology, Mexico, ENSIACET, France, University of Leipzig, Germany, discussed weblabs as an educational tool, its potential developments and outcomes.

3. Cooperative Weblab in Chemical Engineering

In this work we is introduce a unique approach for weblabs: Cooperative Weblab (CW). This new format for weblabs promotes intercultural experiences to students also allowing them to integrate the advantages of cooperative learning. Cooperative learning is a more formal kind of activity where students work in teams that stay together for extended periods of time under conditions that involve five criteria: positive interdependence, individual accountability, face-to-face interaction, development of

interpersonal skills, regular self-assessment of group functioning (Kaufman et al., 2000).
CW experiments must be performed by at least two groups of undergraduate students, in two different locations. The cooperation is achieved by asking the participants of the group to simultaneously solve a technical problem, for which a Weblab experiment setup is available.

4. Case Study: CW between USP and UFSCar

The CW presented here was selected among two weblabs of chemical engineering departments, one at USP and the other one at UFSCar, that are 255 km distant.

4.1. Mass Transfer Weblab at UFSCar

It is important for the education of chemical engineers, to handle the fundamentals of mass transfer involved in this experiment and to get familiar with techniques that assess rates of oxygen transfer from the gas phase into the liquid culture medium as well. It is known that the control of oxygen level during aerobic cultivation of microorganisms is very important, and this experiment is suited for this purpose. The oxygen can be removed from the system by bubbling nitrogen, or added by bubbling oxygen. The dissolved oxygen is monitored by a specific electrode. After reaching zero oxygen concentration, the nitrogen flow stops and air starts flowing. The electrode probe measures the dissolved oxygen (DO) in the liquid phase. The experiment aim is to calculate kLa values at different operating conditions of air flow rate and stirrer speed employing the gassing-out method (Shuler and Kargi, 2002; Blanch and Clark, 1997). The main screen of the weblab for mass transfer was built employing National Instruments hardware for data acquisition and LabVIEW software as the supervisory system. The users can choose the experiment operating conditions (air flow rate and stirrer speed) through the main screen. The users can choose the experiment operating conditions either (air flow rate and stirrer speed) through this interface.

4.2. Reactor Temperature Control Weblab at USP

The reactor is a continuous stirred tank reactor (CSTR) and was assembled specially for teaching purposes. The CSTR is a 4.5 L capacity, jacketed glass vessel. At the inner part of the reactor there is an immersion heater (2.500 W), a stainless steel coil, a temperature sensor, a level sensor, a pH sensor, and a double helix impeller. The reactor temperature can be controlled by manipulating the hot water flow through the coil or the cold water flow through the jacket by two electro pneumatic valves. An electrical heater, immersed in the reactor, can simulate different exothermic reaction behaviors. The interface is very friendly and includes a webcam in order to remember the users that they are dealing with an actual setup and not with a simulator. During the CW experiment, only one user at a time can change the system settings, monitored by the other users. The users can also communicate with each other with video conference facilities, and, if necessary, they download the data log and evaluate the results during or after the experiment.
During the experiments, the students must perform calculations together in order to apply the Cohen-Coon (or other equivalent) procedure (Seborg et al. 1989) to obtain the PID tuning parameters. The set of PID tuning parameters is implemented by closing the control loop and is tested by different methods, e.g., set point changes and disturbance rejection tests.

5. Performing the CW Experiments

The main goal of CW experiments is to make students from the different institutions work together. Groups are formed with students from UFSCar and USP in equal number. The idea is to create a link between these students. Firstly they are forced to work together in order to prepare the experiment. They must review the theory and talk in order to decide what procedure will be used during experiments. The idea is not to give them just a recipe but to only give them the guidelines, so that they are forced to decide what will be the sequence of the experiment by themselves. If necessary a previous video conference is performed between the students and at least one instructor.

The experimental session is performed with a lot of video streaming and conversation. There is one instructor at each side, and the main function is to serve as a consultant and to assist them with practical issues that could arise during the experiment. The students exchange information and opinions about the phenomena that take place. In the mass transfer experiment, some concerns arise about the quality of the mixing, the size of the bubbles and the various problems that arise during the experiment (bubbles blocking the tip of the probe, high coalescence phenomena, and conditions of inefficient mixing). In the reactor temperature control experiment, a step change on the coil flow is performed in order to obtain the response curve, and the students are frequently concerned about if the system has reached the new steady state. In general they are surprised by the great time constants of the system.

After the experiments, the students are invited to freely interact with the equipment. They are left free to explore extreme conditions and to test different perturbations. In the case of the reactor temperature control the student can set different unmeasured perturbations by means of a resistance that is immerse in the reactor and verify if the controller they set up is able to cope with them.

The interaction that takes place between the students during the experiment is very strong. As the students shall process the data and prepare the final group report together, the interaction is still present and they are invited to use the network to communicate and prepare the report.

A closing video conference session is performed where the students get a feedback from the instructors and discuss the main issues of the experiments.

6. Conclusions

CW experiments, as any other cooperative learning experiment is sought to allow students the development of important skills, as team working and a significant learning in some relevant chemical engineering topics.

The opinion of the students is in general extremely positive and encouraging. The aforementioned cluster already performed CW experiments with other institutions like ENSIACET in Toulouse, France and TH Karlsruhe, Germany. The experience has shown that this is a valuable educational tool in a world with an increasing global interaction between professionals.

A special effort is necessary in order to fit this kind of activity in the curricula of chemical engineering programs, because it involves two or more institutions. This is the reason why it is fundamental to make this experience known to a broad public: to congregate a great diversity of centers that could be open to perform this kind of experiment.

Some practical issues must be overcome for CWs in different countries such as the difference in time zones, calendars and language. Nevertheless these issues are

unimportant in comparison with the worst of the intercultural experience CW can bring to the chemical engineering education.

References

Blanch, H. W., Clark, D. S., Biochemical Engineering. Marcel Dekker, Inc., New York, 1997.

Gillet, D., Ngoc, A.V., Rekik, Y., Collaborative Web-Based Experimentation in Flexible Engineering Education, IEEE Transactions on Education, 48(4), November 2005.

Kaufman, D.B.,Felder, R.M.,Fuller, H., Accounting for individual effort in cooperative learning teams, Journal of Engineering Education, 89 (2), 33-140, 2000

Kyatera Program Annual Reports, www.kyatera.fapesp.br

Seborg, D.A., Edgar, T.H., Mellichamp, D.A., Process dynamics and control, John Willey & Sons, New York, 1989.

Selmer, A., Goodson, M., Kraft, M., Sen, S., Faye Mcneill, V., Johnston, B.S., Colton, C.K., Performing Process Control Experiments Across the Atlantic, Chemical Engineering Education, 39 (3), Summer 2005.

Shin, D., Yoon, E.S., Lee, K.Y., Lee, E.S., A web-based, interactive virtual laboratory system for unit operations and process system engineering education: issues, design and implementation, Computer and Chemical Engineering, 26, 319-330, 2002.

Shuler, M. L., Kargi, F., Bioprocess Engineering: Basic Concepts. 2nd edition, Prentice Hall International Series, Upper Saddle River, NJ, 2002.

Acknowledgement

Authors would like to acknowledge FAPESP for the financial support though the Kyatera program.

10th International Symposium on Process Systems Engineering - PSE2009
Rita Maria de Brito Alves, Claudio Augusto Oller do Nascimento and Evaristo Chalbaud Biscaia Jr. (Editors)
© 2009 Elsevier B.V. All rights reserved.

Steps for a Multidisciplinary Engineering Competition

Fernanda E. S. Duarte[a], Caio V. P. Delgaudio[a], Flavia N. David[a], Flavio Waltz[a] and Valter C. Souza[a]

[a]*CHEMTECH – A Siemens Company, Rua da Quitanda, n. 50, Rio de Janeiro, CEP 20011-030 - Brasil.*

Abstract

The worldwide number of engineering projects is growing as never before and so is the demand for trained workers. These facts motivate CHEMTECH to train engineering students in crucial areas from process industry, contributing to their development, improving their acceptance in the labor market, and guarantying the availability of better professionals to the industry. In this context, CHEMTECH promoted a multidisciplinary competition between engineering undergraduate students from the most prestigious universities of Brazil. In 2008 the National Engineering Marathon included Computational Fluid Dynamics and Process Control Theory, two important engineering areas. The application of CFD (Computational Fluid Dynamics) to engineering has been growing extremely fast during the last years, especially in design and process optimization as well as to prevent and to mitigate potential process risks. CFD greatest appeal is that one makes these kinds of studies without interrupting the normal operation of industrial plants. Process control analysis is as important as CFD because an efficient control system guarantees stable and confident processes, improving the quality of the final product. This challenge was organized in three steps: call for the contest, training of the attendees and the contest itself, which took place during the Rio Oil & Gas Expo & Conference. During the first step, CHEMTECH invited the best Brazilian universities, each one represented by their best students and an advisor professor. In the next step they were all trained through the internet using the methodology of Distance Learning. The students were trained in PHOENICS (Parabolic Hyperbolic or Elliptic Numerical Integration Code Series), a CFD software developed by the British company CHAM (Concentration, Heat and Momentum Limited) and Scilab, open software for Control Process. This training also gives the students an opportunity to be in contact with real industrial applications. In the end of this step two students from each university, as a team, were qualified for the final contest, in which they were presented to real industrial problems that they needed to solve using the learned tools. Their solutions were evaluated by professionals from CHEMTECH and from some of its remarkable clients and the best group was the champion. Beyond the knowledge and recognition, the students and professors of the winning group also receive prizes and their university receives a donation of computers. Moreover, CHEMTECH is contributing to improve the national engineering and the quality of projects in world process industry.

Keywords: Multidisciplinary Engineering Competition; Computational Fluid Dynamic; Process Control Theory; and Phoenics.

1. Introduction

The worldwide number of engineering projects is growing as never before and so is the demand for trained workers. These facts motivate CHEMTECH to train engineering students in crucial areas from process industry. This multidisciplinary competition improves their acceptance in the labor market, as training the students is the main objective of this project.
CHEMTECH has been promoting and sponsoring events involving Brazilian university students since 2004. This kind of competition develops the technical knowledge of the students and helps them to apply this knowledge in real problems and common industrial situations.
The first edition was promoted in 2004. A regional competition was conduced as a first step, in three Brazilian cities: Belo Horizonte (Minas Gerais state group), Rio de Janeiro (Rio de Janeiro state group), and São Paulo (São Paulo state group). The competition was a really success, fact that encouraged the organization of a national competition.
In 2006, Chemtech organized the first National CFD challenge. The event counted with the participation of thirteen Brazilian universities and more than thirty students. The competition took place in the Rio Oil & Gas Expo of that year. This well-succeeded experience was expanded for other areas of the knowledge, as process control analysis. The first process control challenge happened in 2007. In 2008, a innovation was proposed: Organize a challenge mixing both areas, the National Engineering Marathon, involving eighteen universities and more than two hundred students.

2. Competition

The National Engineering Marathon included Computational Fluid Dynamics and Process Control Theory, two important engineering areas. The application of CFD (Computational Fluid Dynamics) to engineering has been growing (due to advancements in computer's process capacity), especially in design and process optimization as well as to prevent and to mitigate potential process risks. CFD greatest appeal is that one makes these kinds of studies without interrupting the normal operation of industrial plants. Process control analysis is as important as CFD because an efficient control system guarantees stable and confident processes, improving the quality of the final product. This multidisciplinary competition contributes to student's development, guarantying the availability of better professionals to the industry. This challenge was organized in three steps: call for the contest, training of the attendees and the contest itself, which took place during the Rio Oil & Gas Expo & Conference 2008.

2.1. Call for the contest

During the first step, CHEMTECH invited eighteen Brazilian universities: UFRGS - Federal University of Rio Grande do Sul, UFSC - Federal University of Santa Cartarina, USP- São Paulo University, Unicamp - University of Campinas, UFSCar- Federal University of São Carlos, ITA - Technological Institute of Aeronautical, UFBA - Federal University of Bahia, UFMG - Federal University of Minas Gerais, UFU - Federal University of Uberlandia, UFCG - Federal University of Campina Grande, UFPE - Federal University of Pernambuco, UFRN - Federal University of Rio Grande do Norte, UFPR - Federal University of Paraná, UFRJ - Federal University of Rio de Janeiro, UFF - Fluminense Federal University, IME - Engineering Military Institute,

UERJ - State University of Rio de Janeiro, UFES - Federal University of Espírito Santo, each one represented by an advisor professor and their best students, which means 250 graduation students. The students were from different engineering areas. The distribution is showed in the Figure 1.

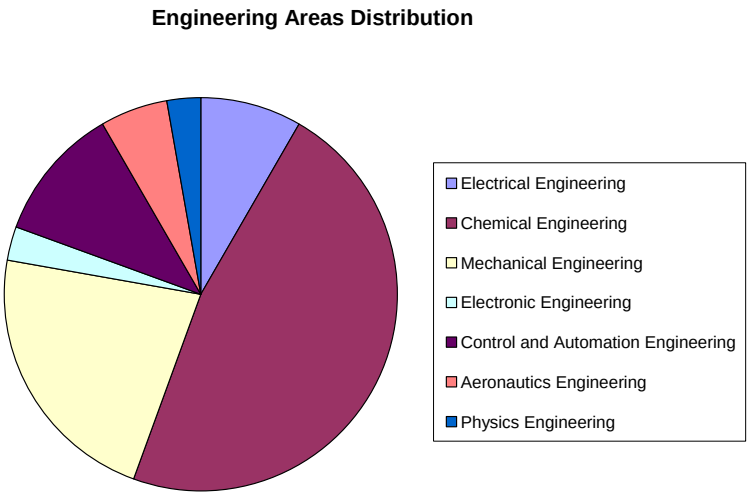


Figure 1: Distribution of the students trough engineering areas

2.2. Training

In the next step, the students were all trained through the internet using the methodology of Distance Learning. The course material was composed by electronic books, divided in modules, available for download to all students. Each module corresponded to a week of course. During this time, chats among the students and the monitors (CHEMTECH professionals, with experience in the CFD and Control fields) were scheduled. Moreover, the student's doubts and questions could be discussed through forums on the internet or trough email, directed to the monitors.

The most difficulty task in this step, and in this teaching methodology itself, is to keep the attendance and participation levels. In order to try to keep the levels, the proposed exercises were all based in real industrial problems and special chat sessions with remarkable specialists and professors in those areas were scheduled. Although all of this efforts the participation level lowered along the course. In the first week, the attendance was 92% (average over the eighteen universities), while in the last week it was 48%. The tendency described above can be observed in Figure 2, which shows the participation along the course weeks for four selected universities.

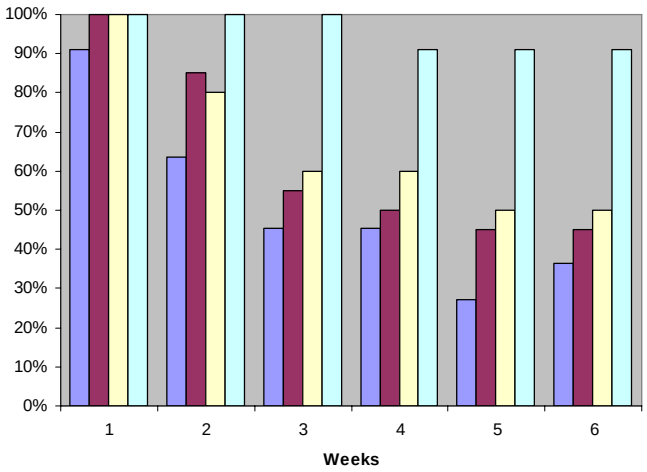


Figure 2: Participation in the training along the weeks. Each color indicates a different university and the participation level is indicated through the bars.

The whole training took six weeks. Besides the course material available about CFD and Control theory, two software were sent to the students: PHOENICS (Parabolic Hyperbolic or Elliptic Numerical Integration Code Series), CFD software developed by the British company CHAM (Concentration, Heat and Momentum Limited) and Scilab, open software for Control Process. In order to make easy the learning process with the software, some tutorials were written and sent to the students, concerning the main functions of the software.

An exercise list was also sent every week to the students, with problems involving the topics learned until that point. The solution of the list was used to rank the students of each university in order to select the better grades to the next step of the event: the contest itself. Doubts and questions about the exercises could also be clarified during the weekly chats with CHEMTECH's professionals.

This training gave the students an opportunity to be in contact with industrial applications, something that occurs rarely during the under graduation courses. In the end of this step, two students from each university, as a team, were qualified for the final contest, during which the students were presented to real industrial problems that they needed to solve using the learned tools.

2.3. Final Contest

The final contest took place in the Rio Oil & Gas 2008, the biggest event of the Oil and Gas industry in Latin America. The selected students and their professors were brought to Rio de Janeiro from many places in Brazil during the exposition week.

This final step was divided in two parts. The first one was constituted by eliminatory tests, which were done by all the eighteen teams. The six better teams, ranked by the grade in the tests, were classified to the second step. A new chance was given to the twelve teams that did not get their place in the second step: another test was applied and the two best teams were also classified. With the best sixteen from the initial two

hundred and fifty students the second step was conducted in the last day of the Rio Oil&Gas. The final test lasted four hours and a problem about a solar energy collector and its functioning coupling CFD and control was proposed to the students.
Their solutions were evaluated by professionals from CHEMTECH and from some of its remarkable clients and the best group was the champion. Beyond the knowledge and recognition, the students and professors of the winning group also received prizes and their university received a donation of computers.

3. Conclusion

The National Engineering Marathon promoted an exchange among universities, their students, and the engineer companies, identifying new talents. Moreover, CHEMTECH is contributing to improve the national engineering and quality of projects in world process industry.
Some recognition has already resulted from this educational initiative and the company was awarded the 2004 Social Responsibility Prize, granted by CREA (National Engineering and Architecture Council).
The competition has also awaked the interest of a lot of undergraduate students in the most different engineering areas and has revealed some great potential future engineers. This successful experience is now being expanded to other countries, in the International Engineering Marathon that will happen in this year.

2150

10th International Symposium on Process Systems Engineering - PSE2009
João B. M. Santos, George Acioli Júnior, Henrique C. Barroso and Péricles R. Barros
© 2009 Elsevier B.V. All rights reserved.

A Flexible Laboratory-Scale Quadruple-Tank Coupled System for Control Education and Research Purposes

João B. M. Santos[1], George Acioli Júnior[1], Henrique C. Barroso[1] and Péricles R. Barros[1]

[1]*LIEC – Department of Electrical Engineering, CEEI, UFCG*
Av. Aprígio Veloso, 882 Bairro Universitário CEP 58429-900, Campina Grande – PB, Brazil

Abstract

In this paper it is described a flexible laboratory-scale quadruple-tank coupled system designed for Control Education and Research purposes. The quadruple-tank process presented is a flexible structure for experimental implementation and evaluation of different control strategies in a multivariable non-linear process. The structure allows the implementation of different experiments, such as SISO (single-input single-output), TITO (two-input two-output) and MIMO (multiple-input multiple-output), with the option of coupling or cascade effect between tanks. A general description for the system is presented. A linear dynamic model of the tank system is presented for a SISO configuration and simple experiments are shown which illustrate the use of the laboratory-scale system in undergraduate control course.

Keywords: Control Education, Laboratory Plant, Coupled Tanks, PID Control, System Identification, Dynamical Modeling

1. Introduction

Laboratory-scale processes are important tools that help in the academic study of various phenomena observed in industrial plants, especially in the study of techniques for the multivariable control of non-linear processes.

The control of the liquid levels in storage tanks and reaction vessels is a common control problem in chemical and petrochemical processes industries. The tank process can be modeled with some constraints, allowing a quick and efficient implementation and test of different strategies of control.

Some multivariable laboratory-scale tanks systems are commercially available, for example from Quanser Consulting in Canada, Educational Control Products in the United States, and Feedback Instruments and TecQuipment in the United Kingdom (Johanson, 2000).

In this paper, a new flexible laboratory-scale quadruple-tank system for experimental level and flow control is described. This system has characteristics of construction and application rather individual, becoming a very flexible platform for implementation and evaluation of different control strategies in multivariable non-linear processes. The laboratory setup is used in Control and Automation undergraduate courses at the DEE/UFCG (Departamento de Engenharia Elétrica/Universidade Federal de Campina Grande) and for control research purposes. A laboratory-scale quadruple-tank system photo is shown in Figure 1(a).

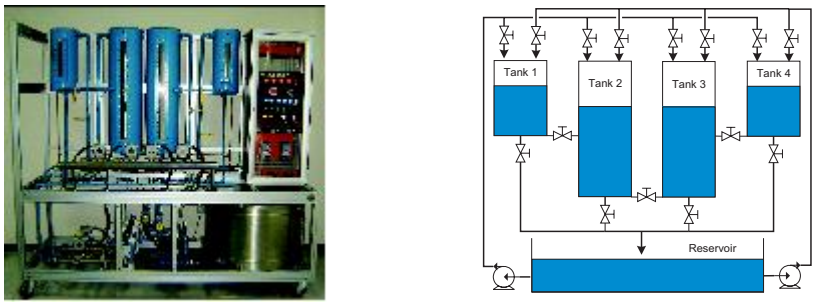


Figure 1: (a) Laboratory-scale Tank System (b) schematic diagram of the tank

The laboratory-scale quadruple-tank system gives the students the opportunity to solve control problems as well as to integrate theoretical knowledge obtained at lectures with practical experience. The experiments performed in the laboratory system for teaching undergraduate students in modeling, continuous-time identification and PI/PID controller design are presented here. For illustration, in this paper only experiments performed in SISO (Single-Input Single Output) configuration are presented.
This paper is organized as follows. In section 2, the laboratory-scale quadruple-tank system is described. In section 3, a linear dynamic model of the SISO tank system presented in literature is reviewed. In section 4, some of the experiments performed in the laboratory system are described. In section 5, the experiments results are presented and, finally, conclusions are presented in Section 6.

2. Laboratory-Scale System Description

The laboratory-scale system (shown in Figure 1(a)) consists of four tanks of different sizes, two hydraulic pumps, two frequencies inverters, six differential pressure transmitters, a PLC and a PC (Personal Computer) with SCADA (Supervisory Control and Data Acquisition).
Four differential pressure transmitters are used for measuring the level of the liquid in each tank. Two are used for measuring the liquid-flow of the hydraulic pumps. The primary element for flow measurement is the orifice plate. The two hydraulic pumps are driven by frequency inverters. This way, the liquid is pumped into the top of each tank. A simplified schematic diagram of the quadruple-tank system is shown in Figure 1(b).
The laboratory-scale system allows multidisciplinary use to support several disciplines at the electrical engineering undergraduate course at UFCG. The undergraduate students have the opportunity to test modeling and simulation techniques, linear control systems design, PLC (Programmable Logic Controller) programming, supervisory design and industrial instrumentation. It is also used in the graduate course for experimental design, validation of SISO and TITO control techniques, system identification and linear modeling of non-linear process.

Solenoid valves are installed to ensure the physical integrity of the plant when an abnormal condition is detected. They interrupt the liquid-flow from the pumps to the tanks. An alternative way had to be created because the pumps are not turned off.
The control is implemented in an Allen-Bradley's Controllogix PLC or using Matlab with its OPC (OLE for Process Control) toolbox. To improve the interface between the user and the platform for the PLC, a supervisory window turns available using Wonderware Intouch SCADA software. The entire configuration parameters,

monitoring and any change in the process control is done using the supervisory. The PLC turns available the process data for the supervisory through the RSlinx OPC Server.
The arrangement of the set of valves allows the configuration of the plant for carrying out different experiments, such as SISO, TITO and MIMO, with the option of coupling or cascade effect between tanks.
In this paper a SISO configuration schema is adopted. In this schema, the liquid is pumped into the top of the second tank from the reservoir, and outflows of the tank only due to gravity. The general objective of the control is to reach and stabilize the level in the tank 2 by controlling the entry flow.

3. Dynamic Model

In this section, the dynamic model for the SISO configuration system is presented (Ogata, 2003). Industrial processes generally have turbulent flux resulting in non-linear model equations. So, linearization is used to simplify the analysis and controller design. The resistance R to the liquid flux is defined as the level variation ΔH necessary to cause a unit variation on the outflow ΔV. The capacitance C of a tank is defined as the liquid quantity variation necessary to cause a unit variation in the liquid level.
Assuming the outflow turbulent, the steady state flow is $Q = K\sqrt{H}$ where, Q is the flow in m^3/s, K is the coefficient in $m^{2,5}/s$ and H is the level in m.
In the SISO configuration q_i is the entry flow in m^3/s, q_o is the outflow in m^3/s, V_1 is the entry valve, V_o is the out valve and C_1 is the capacitance. Small letters are used for the small deviations in the stead state values of the correspondent variables.
We have

$$C_1 \frac{dh}{dt} = q_i - q_o \qquad (2)$$

and

$$q_0 = 2\frac{h}{R_o} \qquad (3)$$

where, R_o is the out valve resistance. Substituting Eq. (3) in Eq. (2) results

$$\dot{h}(t) = -\frac{2h}{R_o C_1} + \frac{1}{C_1} q_i \qquad (4)$$

Eq. (4) is the system dynamic equation. This equation relates the liquid level variation with the incoming input.

4. Experiments Description

4.1. Standard Relay Experiment – Gain Margin Estimate
The standard relay test presented in (Åström,1995) is used to estimate the critical point gain and frequency. It can be shown (Schei, 1994) that if this relay test is applied to a closed loop system, with transfer function *G(s)*, the limit cycle occurs at the closed loop critical frequency and the gain margin can be computed from the loop gain

$$L(j\omega_{gm}) = G(j\omega_{gm})C(j\omega_{gm}) \qquad (8)$$

4.2. Loop-Gain Relay Experiment - Phase Margin Estimate

A general relay procedure to estimate the frequency point for which a given transfer function has a desired gain is presented in (Arruda, 2003a). The conditions of the limit cycle operation are defined by the following proposition, with proof found in (Arruda, 2003a):

Proposition 1: Consider a stable closed loop *M(s)*, with loop-gain *L(s)*, and a real positive number *r*. So the transfer function

$$F(s) = \frac{2}{r}\frac{M(s)}{M(s)\left(\frac{1-r}{r}\right)+1} - 1 \qquad (9)$$

is also stable. Then if a limit cycle is present, it oscillates at a frequency ω_o such that

$$|L(j\omega_0)| \approx r \qquad (10) \quad \blacksquare$$

4.3. Relay Based Gain and Phase Margins Redesign

The closed-loop performances are evaluated here on the phase and gain margins sense using the relay estimations previously described. This information is used to redesign the controllers as proposed in (Arruda, 2003b). This problem is solved using an iterative approach applied to the following equations:

$$\begin{aligned} &\angle G(j\omega_u)C(j\omega_u) = -\pi \\ &|G(j\omega_u)C(j\omega_u)| = \frac{1}{A_m} \\ &|G(j\omega_g)C(j\omega_g)| = 1 \\ &\angle G(j\omega_g)C(j\omega_g) = -\pi + \phi_m \end{aligned} \qquad (11)$$

where A_m is the desired gain margin and ϕ_m the desired phase margin. The iterative algorithm only requires the knowledge of the frequencies ω_u and ω_g. These frequencies are the solutions to Eq. (11a) and Eq. (11c). This algorithm uses the following lemmas to update the controller's parameters.

4.3.1. Controller Redesign for Gain Margin

The controller gain can be calculated for achieving the gain margin A_m using Eq. (11b). That is, with the current gain margin GM^k and the critical frequency $\omega_u{}^k$ one can compute the controller proportional gain $\overline{K}_c^{k+1}$ from

$$\overline{K}_c^{k+1} = \frac{K_c^k GM^k}{A_m} \qquad (12)$$

4.3.2. Controller Redesign for Phase Margin:

The controller gain can be calculated for achieving the phase margin ϕ_m using Eq. (11d). This step is separated into two parts:

1. Determine T_i^{k+1} such that Eq. (11d) is satisfied, i.e.,

$$T_i^{k+1} = \frac{\tan[-\pi + \phi_m - PM^k + \tan^{-1}(\omega_g^k T_i^k)]}{\omega_g} \qquad (13)$$

2. Now, update the controller proportional gain K_c^{k+1} such that the loop gain at the frequency ω_g^k is equal to one,

$$K_c^{k+1} = \overline{K}_c^{k+1} \frac{\sqrt{(1/T_i^k)^2 + \omega_g^2}}{\sqrt{(1/T_i^k)^2 + \omega_g^2}} \qquad (15)$$

The controller at the end of the iteration is finally given by

$$C^{k+1} = K_c^{k+1}\left(\frac{s + 1/T_i^{k+1}}{s}\right) \qquad (16)$$

5. Experimental Results

5.1. Controller Desing

There exist several design techniques based on the FOPDT model (Åström, 1995). Consider the Chien, Hrones and Reswick (CHR) SetPoint 0% and 20% methods. In these techniques the PI controller is designed based on the parameters τ, T_p, K_p and $\theta=\tau/T_p$. Two PI controllers are designed.

$$C_{chrset0\,pi}(s) = 182.45\left(1 + \frac{1}{1213.92s}\right)$$

$$C_{chrset20\,pi}(s) = 312.77\left(1 + \frac{1}{1011.6s}\right)$$

The closed-loop step responses are shown in Figures 6(a) and 6(b). The closfed-loop with $C_{chrset0\,pi}(s)$ is more aggressive than closed-loop with $C_{chrset20\,pi}(s)$.

5.2. Closed-Loop Performance Evaluation

The closed-loop performance is evaluated here on the phase (PM) and gain (GM) margins sense using the standard relay and loop gain experiments. The margins estimated for the two closed-loops are in Table 1. The relay experiments in closed-loop with $C_{chrset20\,pi}(s)$ are shown in Figures 8(a) and (b).

	GM	PM
$C_{chrset0\,pi}(s)$	12.16	12.68
$C_{chrset20\,pi}(s)$	8.96	25.05

Table 1: Margins for the two controllers

6. Conclusion

In control engineering courses, a laboratory experience that complements theoretical lecture material is essential for students. In this way, a flexible laboratory-scale quadruple-tank system and a set of experiments have been presented. The laboratory system is used in Control and Automation courses and in control researches at DEE/UFCG. In this paper only simple experiments performed in undergraduate courses are presented.

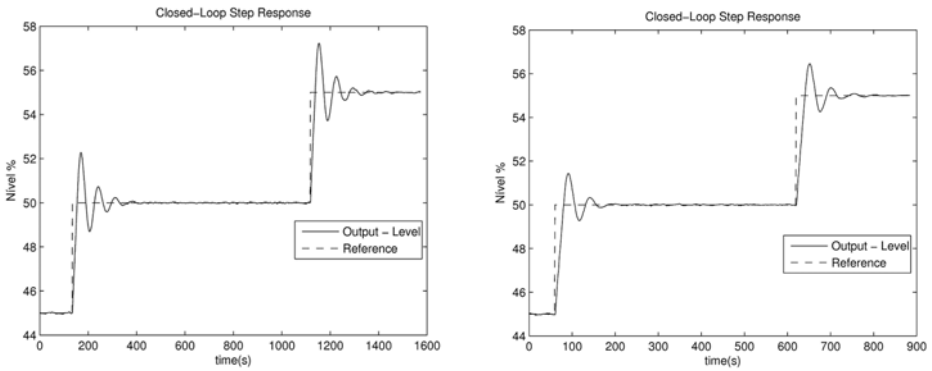


Figure 2: (a) CHR controller closed loop step response (set point 0%) and (b) CHR controller closed loop step response (set point 20%)

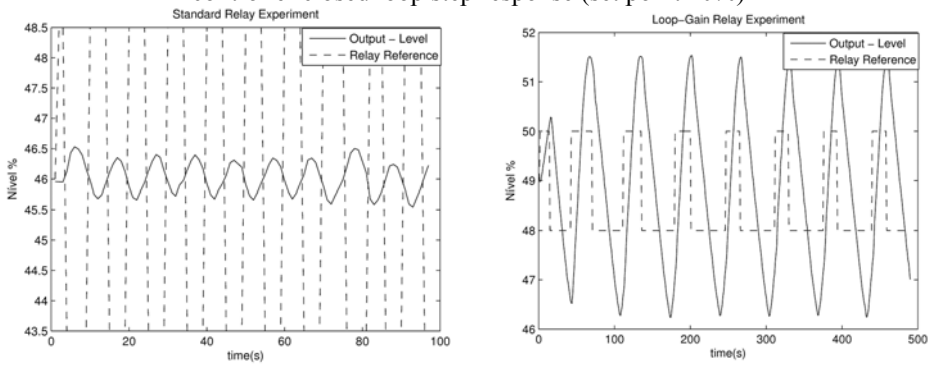


Figure 8: (a) standard relay experiment and (b) loop-gain relay experiment

References

Åström, K. J. and HÄagglund, T. (1995). PID Controllers: Theory, Design and Tuning, 2nd edn, Instrument Society of America, Re-search Triangle Park, North Carolina.

Coelho, F. S. and Barros, P. R. (2003). Continuous-time identification of first-order plus dead-time models from step response in closed loop, 13th IFAC Symposium on System Identification, Rotterdam (The Netherlands).

de Arruda, G. H. M. and Barros, P. R. (2003a). Relay-based closed loop transfer function frequency points estimation, Automatica 39(2): 309-315.

de Arruda, G. H. M. and Barros, P. R. (2003b). Relay based gain and phase margins PI controller design, IEEE Transctions on Inst. And Meas. Tech. 52(5): 1548-1553.

Johansson, K. H. (2000). The quadruple-tank process: A multivariable laboratory process with an adjustable zero, IEEE Transactions on Control Systems Technology 8: 456-465.

Ogata, K. (2003). Engenharia de Controle Moderno, Prentice-Hall, Inc.

Schei, T. S. (1994). Automatic tuning of PID controllers based on transfer function estimation, Automatica 10: 1983-1989.

Author Index