16th European Symposium on Computer Aided Process Engineering
and 9th International Symposium on Process Systems Engineering
W. Marquardt, C. Pantelides (Editors)

Risk analysis and robust design under technological uncertainty

R.F. Blanco Gutiérrez[a*], C.C. Pantelides[a†] and C.S. Adjiman[a]

[a]Centre for Process Systems Engineering, Department of Chemical Engineering, Imperial College London, London SW7 2AZ, U.K.

Technological innovation in process design often leads to increased technological risk arising from incomplete knowledge. We propose a systematic approach to manage this risk using mathematical models that are sufficiently detailed to quantify risk. Global sensitivity analysis is used to determine the complete probability distributions for the key performance indicators of the process, thereby allowing informed decisions to be taken regarding the acceptability of the risk inherent in a given design. It also produces global sensitivity indices which allow the identification of the critical uncertain parameters on which additional R&D needs to be focused if the risk is deemed to be unacceptably high. If the risk is acceptable, then scenario-based approximation is used to handle the residual uncertainty in the critical parameters. Issues regarding the robust and efficient solution of problems involving large numbers of scenarios based on nonlinear models with thousands of variables are considered. The methodology is demonstrated via a case study concerning the design of a catalytic tubular reactor.

1. INTRODUCTION

Technological innovation in processes and products almost inevitably implies increased risk with respect to performance, operability and safety. Although this risk can often be reduced by investing time, money and other resources in R&D activities, the increased cost and time spent can significantly reduce the competitive advantage arising from this innovation, e.g. by reducing the probability of achieving a leading market position. Therefore, the potential implications of any residual risk have to be weighed against the potential benefits that may be realised by the deployment of new technology.

The use of model-based methodologies for process design and operation can accelerate R&D activities by complementing experimental investigations at the laboratory, pilot plant and industrial plant scales. In principle, instead of searching the, often large, space of possible designs and operations, experimental R&D can be focused on deriving an accurate model (e.g. by identifying the fundamental chemistry associated with a new catalyst). The model can then be used for the relatively rapid and inexpensive consideration and screening of many alternatives. Once one or more promising alternatives are identified, their predicted performance may be verified again experimentally (e.g. using pilot plants).

Clearly, the effectiveness of this three-step approach depends crucially on the accuracy of the model derived at the first step. Recent years have witnessed significant advances in this context. It is now practically feasible to use detailed models of experimental apparatus to interpret experimental measurements correctly, estimating multiple model parameters from measurements taken from multiple steady-state and/or dynamic experiments. *A posteriori* statistical significance analysis can provide estimates of the errors in the parameter estimates. We also have at our disposal model-based techniques for experiment design techniques which can determine the optimal conditions for executing further experiments aiming at achieving maximum model accuracy. Nevertheless, it has to be recognised that, irrespective of the above advances, model uncertainty cannot be fully eliminated, and consequently, a number of important questions need to be addressed:

1. Given a certain level of model accuracy and external disturbances, what is the resulting uncertainty in the key performance indicators (KPIs) of a process or product designed using this model?

*Financial support: CONACyT and CPSE.
†Corresponding author. Tel.: +44 (0) 20 7594 6622; fax: +44 (0) 20 7594 6606. E-mail address: c.pantelides@imperial.ac.uk

2. If the risk associated with this uncertainty is unacceptable, and further R&D is required to resolve some of the inaccuracies in the model, which are the critical model aspects on which such R&D needs to be focused?

3. If the risk is, in principle, acceptable, then what is the best design that can take account of the residual model uncertainty?

There is already a large body of research aiming to provide answers to the questions posed above, with particular emphasis on the last one. Work to date has employed different metrics such as flexibility indices [1], trade-offs between flexibility indices and maximum regret [2], expected economic performance [3], and the cost of R&D [4]. Different tools have been proposed for the analysis of feasible regions in the presence of uncertainty, e.g. [5–8], and specific aspects such as technology evolution can be included [9]. The problem can be formulated as a two-stage stochastic optimisation problem (e.g. [10]).

This paper aims to complement the above work by providing a quantitative model-based methodology for addressing the first two of the questions posed above. Of course, the use of models for the quantification of technological risk will be successful only if the models can predict the situations that potentially give rise to such risk, e.g. the formation of undesirable by-products through side reactions, or the occurrence of hot spots in reactors through imperfect mixing. Almost always, such models will be more complex than those used for the prediction of nominal performance (e.g. the yield of the main reactor product or the average reactor temperature), and risk-management techniques need to be able to cope with such increased model complexity. The issue of model complexity also affects the practical feasibility of techniques for addressing the last of the three questions above. Most of the work to date reported in the open literature (e.g. scenario-based optimisation) has been applied only to rather small models involving small numbers of uncertain parameters which can be explored using a relatively small number of scenarios.

2. METHODOLOGY

The proposed methodology outlined in figure 1 starts by constructing a detailed process model and validating it using techniques of the type outlined in the introduction. This validation process results in optimal estimates of the values of model parameters and also in estimates of the accuracy of these values (e.g. in the form of confidence ellipsoids or joint probability density functions). In the second step, the model, with the nominal values of its parameters, is used to determine an optimal design and operating conditions using standard deterministic optimisation techniques.

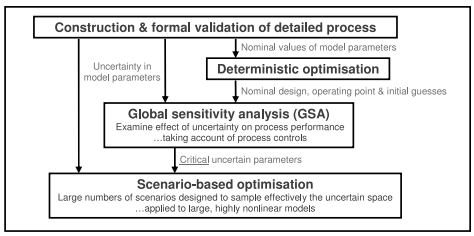

Figure 1. Proposed methodology.

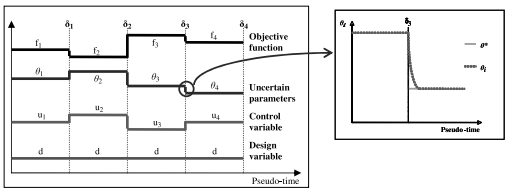

Figure 2. Pseudo-dynamic optimisation approach.

2.1. Global sensitivity analysis

The third step of the methodology aims to quantify the effects of parametric uncertainty on the process KPIs, including the objective function (e.g. economic performance) and constraints (e.g. relating to product quality and process safety and operability). This task is often performed using local sensitivity analysis based on the partial derivatives of the KPIs with respect to the uncertain parameters. Albeit conceptually simple and computationally inexpensive, this approach has certain important deficiencies. First, local values may fail to capture the KPI variability induced by the model parameters varying over ranges of values. Secondly, most processes have controls which can be used during operation to counteract

the effects of parameter uncertainty; hence, the effective sensitivity with respect to a certain parameter may be smaller than that implied by the local sensitivity value. Finally, any single measure of sensitivity is unlikely to contain sufficient information for assessing whether the risk inherent in a certain design is acceptable. Consequently, here we adopt a different approach based on global sensitivity analysis (GSA). This involves solving the following optimisation problem for a fixed design d and a sequence of values of the uncertain parameters θ:

$$\begin{array}{lll} & \Phi(d,\theta) = \max\limits_{u\in\mathcal{U}} \Phi(d,u,x,\theta) & \\ \text{s.t.} & f(d,u,x(z),x_z(z),x_{zz}(z),\theta) = 0 & \forall\, z\in\Omega \\ & h(d,u,x(z),x_z(z),x_{zz}(z),\theta) = 0 & \forall\, z\in\Gamma(\Omega) \\ & g(d,u,x(z),x_z(z),x_{zz}(z),\theta) \le 0 & \forall\, z\in\Omega \\ & y = \mathcal{Y}(d,u,x(z),x_z(z),x_{zz}(z),\theta) & \forall\, z\in\Omega \\ & y^L \le y \le y^U & \end{array} \tag{1}$$

Here Φ represents the objective function (e.g. an economic performance criterion), u a vector of control variables that may be varied over a space $\mathcal{U}$, and x is a vector of state variables which may be distributed over a domain Ω of independent variables z (e.g. spatial position). The model equations f are generally mixed systems of partial differential and algebraic equations involving d, u, x and the latter's partial derivatives, and subject to boundary conditions h and performance constraints g. The KPIs y are given functions $\mathcal{Y}$ of the other variables and are subject to lower and upper bounds.

The above optimisation determines the best set of operating conditions for the given design under a certain realisation of the uncertain parameters θ. The latter vary over a given domain Θ with a given probability distribution[3]. For the purposes of the GSA, the space Θ is sampled using a low-discrepancy sequence due to Sobol'[11] which has a number of desirable properties. First, for any positive integer k, a sequence of 2^k points covers the uncertainty space uniformly. Secondly, and unlike uniform grids, the projection of N sample points onto any parameter axis results in N distinct values of that parameter.

One valuable output of the GSA is an estimate of the complete probability distribution of each and every KPI. This provides a good assessment of the "upside" and "downside" inherent in design d and allows a more detailed assessment of the risk than what can be achieved based on aggregate measures such as expected value and variance. If the risk is deemed to be unacceptable, then one may have to go back to step 1 of the methodology to obtain more accurate estimates of the model parameters. Usually, this implies further experimentation, the cost of which may not be trivial. It is, therefore, important to focus this experimentation on those parameters which have the most impact on the process KPIs. Such critical parameters may be identified via global sensitivity indices also computed by GSA. Here we employ the indices proposed by Sobol'[12] which are derived from the "analysis of variances" (ANOVA) decomposition of the nonlinear functions $\Phi(d,\theta)$ and $y(d,\theta)$ defined by the solution of optimisation problem (1). For example, in the case of two parameters θ_1 and θ_2, the decomposition is expressed as:

$$\Phi(d,\theta_1,\theta_2) = \Phi_0(d) + \Phi_1(d,\theta_1) + \Phi_2(d,\theta_2) + \Phi_{12}(d,\theta_1,\theta_2) \tag{2}$$

For a given d, the variances of the functions Φ_1, Φ_2 and Φ_{12} can be calculated from the values of Φ determined during the sampling. A global sensitivity index is then defined as the ratio of the variance of each of these functions to the overall function variance. For example, the first-order global sensitivity index for parameter θ_1 in Eq. (2) is defined as:

$$S_1^{\Phi}(d) = \frac{Var[\Phi_1(d,\theta_1)]}{Var[\Phi(d,\theta_1,\theta_2) - \Phi_0(d)]} = \frac{Var_{\theta_1}[\mathcal{E}_{\theta_2}(\Phi(d,\theta_1,\theta_2))]}{Var_{\theta_1,\theta_2}(\Phi(d,\theta_1,\theta_2) - \Phi_0(d))} \tag{3}$$

The first-order global sensitivity indices S_i^{Φ} and S_i^{y} are quantities in the range $[0,1]$ which satisfy $\sum_i S_i^{\Phi} = \sum_i S_i^{y} = 1$. The parameters θ with the largest sensitivity indices in Φ and/or y are flagged as critical uncertain parameters on which any further experimental R&D effort needs to be focused. This method also allows the calculation of higher-order parameter interactions. For instance, given Eq. (2), S_{12}^{Φ} is the global sensitivity index of Φ for the interaction between θ_1 and θ_2.

[3] An estimate of this is produced by the *a posteriori* statistical significance analysis during the model validation step.

The variances $Var[.]$ and expected values $\mathcal{E}(.)$ in expressions of type (3) are multidimensional integrals calculated using a technique developed by Sobol' [12]. In our implementation, the process model is constructed in the gPROMS modelling tool[13] which is also used for the solution of the optimisation problem (1). The GSA is implemented in C++ as a gPROMS-Based Application (gBA) interacting directly with the gPROMS solution engine. The application has been parallelised for execution on distributed computer networks using MPI-based communication to allow the simultaneous evaluation of multiple sampling points.

2.2. Scenario-based optimisation

Even if the GSA indicates that the risk associated with a nominal design is acceptable, the design may not be optimal when one considers the variation of the objective function value caused by parameter variability. For some values of the uncertain parameters, it may not even be feasible, violating some of the inequality constraints and bounds in (1). Therefore, we need to determine a new design which takes explicit account of the parameter variability. This is a well-known problem which has been the focus of attention of much of the literature mentioned in the introduction to this section. A standard technique for solving the problem is by postulating a set of scenarios $s = 1, ..., NS$, each corresponding to a different realisation of the parameters $\theta^{[s]}$, and then determining a design d and controls $u^{[s]}, s = 1, ..., NS$ which maximise some probabilistic measure of performance (e.g. the expected value of Φ).

To obtain a good estimate of expected values, it is desirable to use a large number of scenarios. However, this significantly increases computational cost. In addition, with complex models of the type of interest here, numerical convergence (i.e. obtaining a set of variable values that satisfy the equality constraints in (1)) becomes a difficult task, and this can compromise the robustness of the overall algorithm. Here we use a pseudo-dynamic optimisation formulation (cf. figure 2) to solve the scenario-based problem, where smooth transitions between scenarios are achieved via a homotopy/continuation approach. This facilitates the initialisation process by requiring only one set of initial guesses.

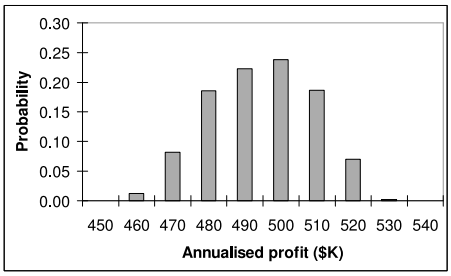

Figure 3. Probability distribution for profit from GSA.

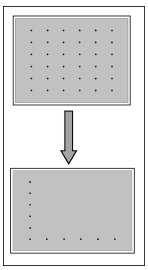

Figure 4. Reduction in number of scenarios when no parameter interactions exist.

A further reduction in problem complexity can be achieved in cases where the higher-order sensitivity indices (cf. section 2.1) indicate no significant interactions between parameters. In this case, a reduced set of scenarios can be found by sampling the uncertain parameter space along each parameter domain independently, keeping other parameters fixed at a single set of values (see figure 4).

Even with the above reductions in the numbers of uncertain parameters and scenarios, the solution of the scenario-based optimal design problem may remain prohibitively expensive for complex systems. To address this issue, we use a Sample Average Approximation (SAA) approach[14,15] which approximates the optimal design through the solution of a sequence of problems, each involving a much smaller number N of scenarios. Such N-scenario problems are formulated and solved until the average values and standard deviations of the objective function and design variables obtained up to a certain point converge to constant values. If the total number of N-scenario problems solved is M, the computational cost is usually much smaller than what would be required for solving a single problem with $M \times N$ scenarios. Moreover, this approach is more amenable to paralellisation as several N-scenario problems can be solved in parallel.

Table 1
Uncertain parameters and first-order sensitivity indices for the objective function.

Parameter	Mean	Std Dev %	S
CO heat of formation, $\Delta H_{f(CO)}$ (kJ kmol^{-1})	-110440	3.3	0.1822
$COCl_2$ heat of formation, $\Delta H_{f(COCl_2)}$ (kJ kmol^{-1})	-222850	3.3	0.7332
Kinetic coeff., Ek_r (kPa m^3 kmol^{-1})	57686.3	3.3	0.0003
Radial heat transfer coeff., kr (kW m^{-1}K^{-1})	0.05	3.3	0.0050
Axial heat transfer coeff., kz (kW m^{-1}K^{-1})	0.05	3.3	0.0077
Overall heat transfer coeff., U (kW m^{-2}K^{-1})	0.096	3.3	0.0374
Cooling water inlet temp., T_{cwin} (K)	293	0.33	0.0127
Feed stream inlet temp., T_{in} (K)	293	0.33	0.0000

3. CASE STUDY

As an illustration of the proposed methodology, we consider the design of an externally cooled catalytic tubular reactor producing phosgene. It is desired to determine the reactor length and diameter which maximise the annualised profit. The available controls include the cooling water flowrate, the feed stream partial pressures for the two reactants and feed stream velocity. The model is steady-state and spatially distributed along the axial and radial dimensions, the spatial discretisation of which results in approximately 5,900 variables. It involves the 8 uncertain parameters listed in table 1. Here these are assumed to be described by independent normal distributions; however, any type of joint probability density function can be used.

As described in the methodology (figure 1), an optimal nominal design is first identified based on the mean values shown in table 1. The global sensitivity indices for the eight parameters are then calculated; this requires the evaluation of 17 (=2×8+1) multidimensional integrals, each computed using a Sobol' sequence of 2^{12} points. The computation requires the solution of 69,632 problems of type (1), with a total CPU time of 55 hours, spread over 32 parallel processes on a Linux cluster of Pentium 4 processors with speeds ranging from 1.79 GHz to 3.39 GHz. The global sensitivity indices for the objective function are shown in the last column of table 1. The heats of formation of CO and $COCl_2$ are clearly the critical parameters in this case. The second-order sensitivity index for these two parameters is $S_{\Delta H_{f(CO)},\Delta H_{f(COCl_2)}} = 0.020$, which indicates that there is little interaction between them.

The GSA also determines the complete probability distribution for the objective function, as shown in figure 3. Assuming that the risk is acceptable, the final step involves optimisation based on multiple realisations ("scenarios") of the two critical parameters. All other parameters are fixed at their nominal values, which greatly reduces the number of scenarios that need to be considered. We consider three different approaches. The first samples the space of $\Delta H_{f(CO)}$ and $\Delta H_{f(COCl_2)}$ using a 6 × 6 uniform grid. The second approach exploits the lack of interaction between the two parameters (as indicated by the low value of the corresponding second-order sensitivity index) to reduce the number of scenarios. Thus, each parameter is sampled independently at six points while keeping the other parameter constant, which results in a total of 11 scenarios (see figure 4).

The third approach employs the SAA method solving a sequence of 5-scenario problems. Figure 5 shows the evolution of the cumulative means and standard deviations of the objective function and the optimal reactor length with the number of problems solved. Both the optimal reactor length and the optimal reactor radius (not shown in the figure) converge rapidly to their final values. The convergence of the profit is somewhat slower. The average CPU time per 5-scenario problem is 409 s. Convergence to the optimal design is achieved after about 10 5-scenario problems requiring 3889 CPU s.

All three approaches give identical results in terms of the design variables, but show large variations in the expected value of the objective function value and its variation over the set of scenarios studied. These differences arise from the difficulty of obtaining good estimates of the corresponding two-dimensional integrals using relatively small numbers of scenarios. To illustrate this point, the last row of table 2 shows benchmark results evaluated *a posteriori* by applying 200 scenarios to the optimal reactor design. It can be seen that the values obtained by the SAA approach are nearest the benchmark values, as might be expected in view of the fact that the SAA makes use of a much larger number of scenarios than the other two methods.

Table 2
Results of the case study for different methods.

Method	Total number of scenarios	Profit (USD/yr) Mean	Profit (USD/yr) Std deviation	Reactor length (m)	Reactor radius (m)	CPU hours
Single optimisation	36	466,000	31,000	1.403	0.714	73
Single optimisation, independent params	11	474,000	26,000	1.403	0.714	0.9
SAA	20×5	479,000	10,000	1.403	0.714	2.3
Benchmark	200	477,000	7,000	1.403 (fixed)	0.714 (fixed)	–

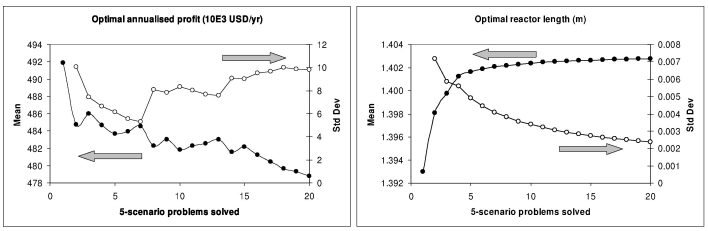

Figure 5. Convergence of SAA approach with number of 5-scenario problems solved.

4. CONCLUSIONS

A systematic methodology has been proposed to manage technological risk arising from incomplete knowledge at the process design stage. It uses mathematical models which are sufficiently detailed to establish a quantitative relationship between the uncertain parameters and the process KPIs. Although the construction of such models is not trivial, it is increasingly being undertaken in industrial practice. An optimisation-based global sensitivity analysis, based on sampling via low-discrepancy sequences, is performed to identify critical parameters affecting the KPIs while exploiting the flexibility afforded by the process control variables. The complete probability distribution of the KPIs is also obtained, thus allowing informed decisions to be made regarding acceptability of the inherent risk. If the risk is deemed to be acceptable, then scenario-based optimisation is employed to determine a design that performs optimally given the variability of the critical parameters. As the number of scenarios may increase exponentially with the number of parameters being considered, the GSA plays a crucial role in eliminating non-critical parameters, and in assessing the extent to which interactions among the critical ones need to be considered. Our results indicate that sample average approximation methods may provide an effective means of handling large numbers of scenarios using nonlinear models involving thousands of variables.

REFERENCES

1. R.E. Swaney and I.E. Grossmann, AIChE J. 31 (1985) 621.
2. M.G. Ierapetritou and E.N. Pistikopoulos, Comput. Chem. Eng. 18 (1994) 163.
3. F.P. Bernardo, E.N. Pistikopoulos and P.M. Saraiva, Ind. Eng. Chem. Res. 38 (1999) 3056.
4. F.P. Bernardo, P.M. Saraiva and E.N. Pistikopoulos, Comput. Chem. Eng. 24 (2000) 1695.
5. W.C. Rooney and L.T. Biegler, Comput. Chem. Eng. 23 (1999) 1563.
6. W.C. Rooney and L.T. Biegler, AIChE J. 47 (2001) 1794.
7. N.J. Samsatli, M. Sharif, N. Shah, L.G. Papageorgiou, AIChE J. 47 (2001) 2277.
8. V. Goyal and M.G. Ierapetritou, AIChE J. 49 (2003) 1233.
9. L. Cheng and E. Subrahmanian and A.W. Westerberg, Comput. Chem. Eng. 27 (2003) 781.
10. J. Acevedo and E.N. Pistikopoulos, Comput. Chem. Eng. 22 (1998) 647.
11. I.M. Sobol', USSR Computational Math. and Mathematical Phys. 16 (1976) 236.
12. I.M. Sobol', Mathematics and Computers in Simulation 55 (2001) 271.
13. gPROMS ModelBuilder v. 2.3.6, Process Systems Enterprise Ltd, www.psenterprise.com (2005).
14. A.J. Kleywegt, A. Shapiro and T. Homem-de-Mello, SIAM J. Optim. 12 (2002) 479.
15. J. Wei and M.J. Realff, Comput. Chem. Eng. 28 (2004) 333.