Self-optimizing control: The missing link between steady-state optimization and control

Sigurd Skogestad

Department of Chemical Engineering, Norwegian University of Science and Technology, N-7491 Trondheim, Norway
Email: skoge@chembio.ntnu.no; Fax: +47-7359-4080

Abstract: Real-time steady-state optimization (RTO) has become increasing popular in recent years. But how should this optimal policy be implemented in the control system? It is argued that the the goal is to find a set of controlled variables which, when kept at constant setpoints, indirectly lead to near-optimal operation with acceptable loss. This is denoted "self-optimizing" control.

Keywords: Chemical process control, control structure design, plantwide control

1. INTRODUCTION

If we consider the control system in a chemical plant, then we find that it is structured hierarchically into several layers, each operating on a different time scale. Typically, layers include include scheduling (weeks), site-wide (real-time) optimization (day), local optimization (hour), supervisory/predictive control (minutes) and stabilizing and regulatory control (seconds); see Figure 1. The layers are interconnected through the **controlled variables** c, whereby the upper layer computes the setpoint value c_s to be implemented by the lower layer. We usually assume **time-scale separation** which for our purposes implies that the setpoints c_s can be assumed to be immediately implemented by the layers below. Which should these internal controlled variables c be? That is, what should we control? This paper attempts to answer this question.

More generally, the issue of selecting controlled variables is the first subtask in the **plantwide control** or **control structure design** problem (Foss, 1973); (Morari, 1982); (Skogestad and Postlethwaite, 1996):

(1) *Selection of controlled variables* c
(2) *Selection of manipulated variables* m
(3) *Selection of measurements* v (for control purposes including stabilization)
(4) *Selection of a control configuration* (structure of the controller that interconnects measurements/setpoints and manipulated variables)
(5) *Selection of controller type* (control law specification, e.g., PID, decoupler, LQG, etc.).

Even though control engineering is well developed in terms of providing optimal control algorithms, it is clear that most of the existing theories provide little help when it comes to making the above structural decisions.

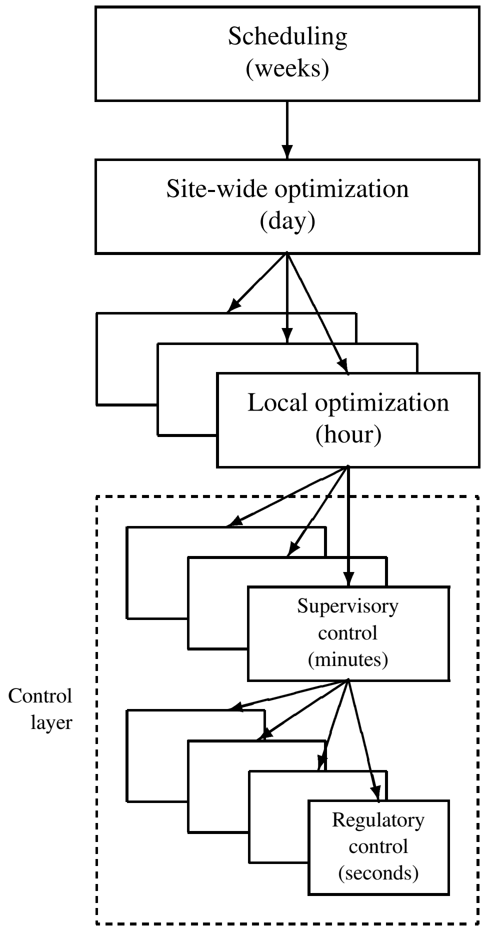

Fig. 1. Typical control hierarchy in a chemical plant.

2. PREVIOUS WORK

Inspired by the work of Findeisen *et al.* (1980), the basic idea of self-optimizing control was formulated about twenty years ago by Morari *et al.* (1980). Morari *et al.* (1980) write that "in attempting to synthesize a feedback optimizing control structure, our main objective is to translate the economic objectives into process control objectives. In other words, *we want to find a function c of the process variables which when held constant, leads automatically to the optimal adjustments of the manipulated variables, and with it, the optimal operating conditions.* [...] This means that by keeping the function $c(u, d)$ at the setpoint c_s, through the use of the manipulated variables u, for various disturbances d, it follows uniquely that the process is operating at the optimal steady-state." If we replace ithe term "optimal adjustments" by "acceptable adjustments (in terms of the loss)" then the above is a precise description of what we in this paper denote a self-optimizing control structure. The only factor they fail to consider is the effect of implementation error $c - c_s$. Somewhat surprisingly, the idea of "feedback optimizing control" of Morari *et al.* (1980) has up to now received very little attention. One reason is probably that the paper also dealt with the issue of finding the optimal operation (and not only on how to implement it), and another reason is that the only example in the paper happened to result in an implementation with the controlled variables at their constraints. The constrained case is "easy" from an implementation point of view, because the simplest and optimal implementation is to simply maintain the constrained variables at their constraints ("active constraint control").

At about the same time, Shinnar (1981) published a more intuitive process-oriented approach for selecting controlled variables, and applied it to the control of a fluidized catalytic cracker (FCC). The work may at first seem unrelated, but if one translates the words and notation, then one realizes that Shinnar's ideas are close to the ideas presented in this paper and in Morari *et al.* (1980). The similar later paper by Arbel *et al.* (1996) extended the FCC case study, and introduced the concepts of "dominant variables" and "partial control". Tyreus (1999) provides some interesting ideas on how to select dominant variables.

Luyben (1988) introduced the term "eigenstructure" to describe the inherently best control structure (with the best self-regulating and self-optimizing property). However, he did not really define the term, and also the name is unfortunate since "eigenstructure" has a another unrelated mathematical meaning in terms of eigenvalues. Fisher *et al.* (1988) discuss selection of controlled variables, mainly focused towards active constraint control. However, somewhat hidden in their HDA example (p. 614) one finds statements about selecting controlled variables which optimal values are insensitive to disturbances. Narraway *et al.* (1991), Narraway and Perkins (1993) and Narraway and Perkins (1994)) strongly stress the need to base the selection of the control structure on economics, and they discuss the effect of disturbances on the economics. However, they do not formulate any rules or procedures for selecting controlled variables. In his book Rijnsdorp (1991) gives on page 99 a stepwise design procedure for designing optimizing control systems for process units. One step is to "transfer the result into on-line algorithms for adjusting the degrees of freedom for optimization". Mizoguchi *et al.* (1995) and Marlin and Hrymak (1997) stress the need to find a good way of implementing the optimal solution in terms how the control system should respond to disturbances, "i.e. the key constraints to remain active, variables to be maximized or minimized, priority for adjusting manipulated variables, and so forth." Finally, Zheng *et al.* (1999) present a procedure for selecting controlled variables based on economic penalties that is similar to the approach presented in this paper (apparently, the work has been performed independently), but they also do not consider the implementation error.

3. OPTIMAL OPERATION AND ITS IMPLEMENTATION

When controlling a chemical plant our first concern is to stabilize the plant and keep its operation within given constraints. These issues may consume some degrees of freedom (e.g. to stabilize levels with no steady-state effect and to satisfy "active" product specifications), but there will generally be many degrees of freedom u left. What should these be used for?

Loosely speaking, they should be used to "optimize the operation". There may be many issues involved, and to trade them off against each other in a systematic manner we usually quantify a scalar performance (cost) index J which should be minimized. In many cases this index is an economic measure, e.g. the operation cost. For example, J could be of the form

$$J = J_u(u, d) = \int_0^T \phi(u, d)dt \tag{1}$$

where u are the degrees of freedom for optimization, d are time-varying disturbances, and T is the total operation time. The optimizing control problem may then be formulated as

$$\min_u J_u(u, d) \tag{2}$$

subject to the inequality constraints

$$g(u, d) \leq 0 \tag{3}$$

where u are the independent variables we can affect (degrees of freedom), and d are independent variables we can not affect (disturbances). In this paper we will for simplicity use steady-state models and the integration in (1) may be replaced by time-averaging over the various steady-states. The main justification for using a steady-state analysis is that the economic performance is primarily determined by steady-state considerations. The effect of the dynamic control performance can be partly included in the economic analysis by introducing a control error term as an additional disturbance.

There are two main issues when it comes to optimizing control. The first is the mathematical and numerical problem of solving the optimization problem in (2) to obtain the optimal operating point. The optimization problem may be very large, with hundreds of thousands of equations and hundreds of degrees of freedom (e.g. for a complete ethylene plant), but with todays computers and optimization methods this problem is solvable, and it is indeed solved routinely today in some plants. The second issue, the focus of this paper, is how the optimal solution should be implemented in practice, that is, which variables c should be held constant between each reoptimization. Surprisingly, this issue has received much less attention.

4. SELF-OPTIMIZING CONTROL

The method adcocated in this paper for selecting controlled variables (task 1) follows the ideas of Morari *et al.* (1980) and Skogestad and Postlethwaite (1996) and is very simple. The basis is to define mathematically the quality of operation in terms of a scalar cost function J to be minimized. To achieve truly optimal operation we would need a perfect model, we would need to measure all disturbances, and we would need to solve the resulting dynamic optimization problem on-line. This is unrealistic in most cases, and the question is if it is possible to find a simpler implementation which still operates satisfactorily (with an acceptable loss). More preciely, the **loss** L is defined as the difference between the actual value of the cost function obtained with a specific control strategy, and the truly optimal value of the cost function, i.e.

$$L = J - J_{\mathrm{opt}}$$

The simplest operation would result if we could select controlled variables such that we obtained acceptable operation with constant setpoints, thus effectively turning the complex optimization problem into a simple feedback problem and achieve what we here call "self-optimizing control":

> **Self-optimizing control** is when we can achieve an acceptable loss L with constant setpoint values c_s for the controlled variables

(The reader is probably familiar with the term self-regulation, which is when acceptable dynamic control performance can be obtained with constant manipulated inputs. Self-optimizing control is a direct generalization to the case where we can achieve acceptable (economic) performance with constant controlled variables.)

The idea is that by locally controlling the right variables c, we can take care of most of the disturbances, and thus reduce the need for continouos reoptimization. This also reduces the need for model information and tends to make the implementation more robust. On the other hand, it usually implies a performance loss compared to the "true" optimal (centralized) solution, and the challenge is to find a "self-optimizing" control structure (i.e. to find the right controlled variables c) for which the loss L is acceptable.

A well-known example of self-optimizing control is the "cake baking process", where the operation is indirectly kept close to its optimum by controlling the oven temperature at the setpoint given in the cook book (which is this case is the "optimizer").

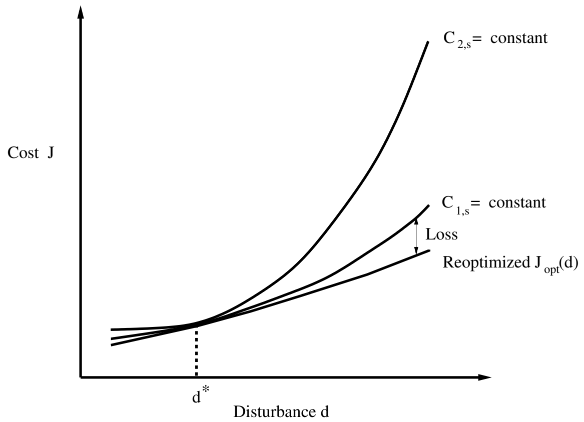

Fig. 2. Loss imposed by keeping constant setpoint for the controlled variable

The idea is further illustrated in Figure 2, where we see that there is a loss if we keep a constant setpoint rather than reoptimizing when a disturbance moves the process away from its nominally optimal operating point (denoted $*$). For the case illustrated in the figure it is better (with a smaller loss) to keep the setpoint c_{1s} constant than to keep c_{2s} constant.

An additional concern with the constant setpoint policy is that there will always be an implementation error $d_c = c - c_s$, e.g. caused by measurement error. The implementation error may cause a large additional loss if the optimum surface is "sharp". To be more specific, we may, as illustrated in Figure 3 distinguish between three classes of problems when it comes to the actual implementation:

(a) *Constrained optimum: Implementation easy.* In the figure is shown the case where the minimum value of the cost J is obtained for $c = c_{min}$. In this case there is no loss imposed by keeping a constant $c_s = c_{min}$. In addition, implementation of an "active" constraint is usually easy, e.g., it is easy to keep a valve closed.

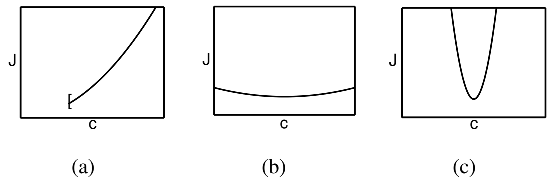

Fig. 3. Implementing the controlled variable

(b) *Unconstrained flat optimum: Implementation easy*. In this case the cost is insensitive to value of the controlled variable c, and implementation is again easy.

(c) *Unconstrained sharp optimum: Implementation difficult*. The more difficult problems for implementation is when the cost (operation) is sensitive to value of the controlled variable c. In this case, we want to find another controlled variable c in which the optimum is flatter.

5. REQUIREMENTS FOR CONTROLLED VARIABLES

We can derive the following **four requirements for of a good candidate controlled c variable** (also see Skogestad and Postlethwaite (1996), page 404):

Requirement 1. Its optimal value is insensitive to disturbances (so that the setpoint error e_{cs} is small)

Requirement 2. It is easy to measure and control accurately (so that the implementation error d_c is small)

Requirement 3. Its value is sensitive to changes in the manipulated variables u, that is, the gain from u to y is large (so that even a large error in the controlled variable c results in only a small error in u). Equivalently, the optimum should be "flat" with respect to the variable c.

Requirement 4. For cases with two or more controlled variables, the selected variables should not be closely correlated.

All four requirements should be satisfied. For example, assume we have a mixture of three components, and we have a measurement of the sum of the three mole fractions, $c = x_A + x_B + x_C$. This measurement is always 1 and thus independent of disturbances (so requirement 1 is satisfied), but it is of course not a suitable controlled variable because it is also insensitive to the manipulated variables u (so requirement 3 is not satisfied). Requirement 3 also eliminates variables that have an extremum (maximum or minimum) when the cost has its minimum, because the gain is zero for such variables.

6. DISTILLATION EXAMPLE

To give the reader som appreciation of the issues, we consider a distillation plant where the plant economics are mainly determinded by its steady-state behavior. With a given feed stream and a specified column pressure[1], a conventional two-product distillation column, as shown in Figure 4, has two degrees of freedom at steady state ($N_{\text{opt}} = N_u = 2$). (From a control point of view the column has $N_m = 5$ degrees of freedom, but two degrees of freedom are needed to stabilize the reboiler and condenser holdups, which have no steady-state effect, and one degree of freedom is used to control the pressure at its given value). The two steady-state degrees of freedom, e.g. selected to be the reflux flow L and the distillate flow D,

$$u = \begin{bmatrix} L \\ D \end{bmatrix}$$

(this is not a unique choice) may be used to optimize the operation of the plant. However the question is: How should the optimal solution be *implemented*, that is, which two variables c should be specified and controlled during operation?

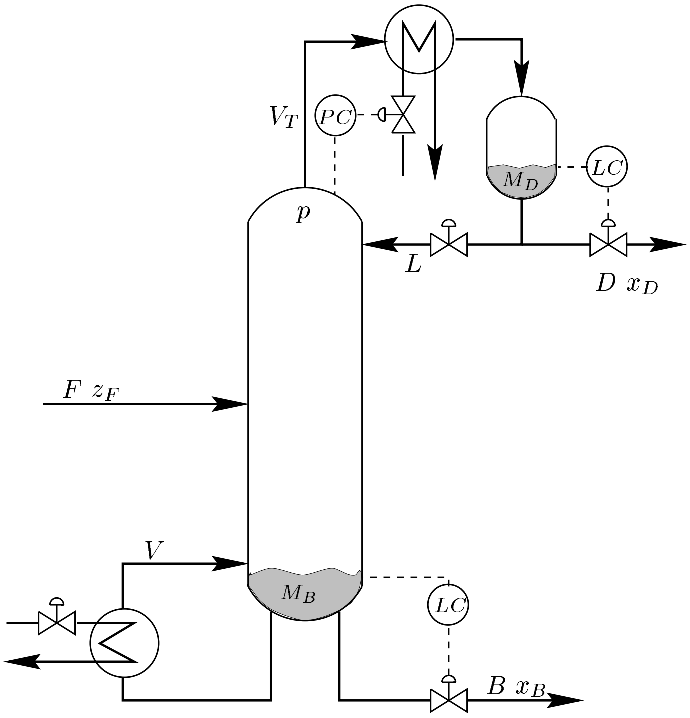

Fig. 4. Typical distillation column controlled with the LV-configuration

To answer this question in a quantitative manner, we need to define the constraints for the operation and the cost function J to be minimized.

Constraints. We assume that the mole fraction of light component in the distillate product x_D must be above

[1] If column pressure is free we often find that the optimal choice is to have maximum cooling corresponding to minimum pressure ("floating pressure control" as suggested by Shinskey (1984)). The reason is that in most cases the relative volatility is improved when pressure is lowered.

0.95, and that to avoid flooding the capacity of the column is limited by a maximum allowed vapor load.

$$x_D \geq x_{D,min} = 0.95; \quad V \leq V_{max}$$

Cost function. Rather than minimzing the cost J, it is more natural in this case to maximize the profit $P = -J$, which is the product value minus the feed costs and the operational (energy) costs which are proportional to the vapor flow V,

$$P = p_D D + p_B B - p_F F - p_V V \tag{4}$$

Case 1. Constrained operation

Let us first consider a case where

- distillate is the more valueable product ($p_D \gg p_B$)
- energy costs are low (p_V is small)

In this case, it is optimal to operate the column at maximum load (to reduce the loss of light component in the bottom) and with the distillate composition at its specification (to maximize distillate flow by including as much heavy component as possible) (Gordon, 1986), i.e.

$$V_{opt} = V_{max}; \quad x_{D,opt} = x_{D,min} = 0.95$$

Thus, the optimum lies at constraints and implementation is obvious: We should select the vapor rate V and the distillate composition x_D as the controlled variables,

$$c = \begin{bmatrix} V \\ x_D \end{bmatrix}; \quad c_s = \begin{bmatrix} V_{max} \\ x_{D,min} \end{bmatrix}$$

In practice, we may implement this using a lower-level feedback control system where we adjust the boilup V to keep the pressure drop over the column (an indicator of flooding, i.e. V_{max}) below a certain limit, and adjust the reflux L (or some other flow, depending on how the level and pressure control system is configured) so that x_D is kept constant.

Case 2. Partly constrained operation

Next, consider a case where

- distillate is the more valueable product ($p_D > p_B$)
- energy costs are relatively high (such that the term $p_V V$ contributes significantly to J)

Also in this case it is optimal to have the distillate composition at its specification, $x_D = x_{D,min}$, but to save energy we will have $V < V_{max}$. Again, we should select x_D as a controlled variable, but it is not clear what the other variable should be. Some options may be $x_B, L, V, D, L/F$ or L/D. To find the best choice a more detailed choice is needed, for example, based on evaluating the economic loss for various disturbances. We often find that the purity x_B is a good choice, but "two-point" composition control is generally to be avoided because of poor dynamic performance.

	$x_B =$ 0.04	$D/F =$ 0.639	$L/F =$ 15.065	$L/D =$ 23.57
$F: 1 \to 1.3$	0	0	0	0
$z_F: 0.65 \to 0.75$	0.019	2.530	0.006	0.129
$q_F: 1 \to 0.5$	0.000	0.000	0.001	0.000
20%impl.error	0.012	inf.	0.119	0.130

inf. denotes infeasible operation

Table 1. Losses in $/min for various disturbances.

The losses for a propane-propylene column case study (Skogestad, 2000) are given in Table 1. We see that keeping L/F constant is a good alternative for "self-optimizing" control with a loss similar to that of keeping x_B constant. However, it is not a good choice to keep D/F or L/D constant.

Case 3. Unconstrained operation

For a case with

- bottoms product is the more valuable product ($p_B > p_D$), but its price p_B is reduced as it gets impure
- energy costs are relatively high

the optimum may be unconstrained in both variables, that is,

$$V_{opt} < V_{max}; \quad x_{D,opt} > x_{D,min}$$

Implementation in this case is again not obvious. Some candidate sets of controlled variables c are

$$c_1 = \begin{bmatrix} x_D \\ x_B \end{bmatrix}; c_2 = \begin{bmatrix} T_{top} \\ T_{btm} \end{bmatrix}; c_3 = \begin{bmatrix} x_D \\ V \end{bmatrix}; c_4 = \begin{bmatrix} L \\ V \end{bmatrix}; c_5 = \begin{bmatrix} L/D \\ V/B \end{bmatrix}$$

and there are many others. Controlled variables c_1 and c_2 will yield a "two-point" control system where we close two loops for quality control; c_3 yields a "one-point" control system where only one quality loop is closed; whereas c_4 and c_5 are "open-loop" policies which require no additional feedback loops (except for the level and pressure loops already mentioned). All of these choices of controlled variables will have different self-optimizing control properties.

Case 4. Feed rate as degree of freedom

Let us go back to case 1, but assume now that the feed rate is a degree of freedom. Again we will find that the capacity and purity constraints are active,

$$V_{opt} = V_{max}; \quad x_{D,opt} = x_{D,min} = 0.95$$

but since we will attempt to push more feed through the column the operation will nevertheless be different from

Case 1. We will get a a less pure bottom product, and thus a larger loss of the valueable light component.

Let us consider this in some more detail. In case 1 the objective of optimizing the profit P is equivalent to maximizing P/F (since F is given). On the other hand, in case 4 the feed rate is a degree of freedom and at the optimum we have $\partial P/\partial F = 0$, and the resulting value of P/F will obviously be smaller than in case 1.

7. A FUNDAMENTAL DIFFERENCE BETWEEN DESIGN AND OPERATION

The last finding has some interesting implications regarding the difference between design and operation. In both cases we want to maximize the profit (value increase) P. During design we always assume a fixed feed rate and thus maximize P/F. On the other hand, during operation the feed rate is often a free variable, and as shown above we will we will get that

$$(P/F)_{\text{operation}} < (P/F)_{\text{design}}$$

Thus, if we optimize the operation we will use the raw materials and energy less effectively than we designed for. An obvious question is: Is this an ethetical and environmental dilemma?

8. CONCLUSION

The results from the steady-state optimization are implemented by computing the setpoints of the controlled variables c. A set of controlled variables is denoted self-optimizing if we can achieve an acceptable (economic) loss with constant setpoints c_s. Important steps in evaluating self-optimizing control are degrees of freedom analysis, definition of optimal operation (cost and constraints), and evaluation of the loss for the set of disturbances.

Acknowledgment. *The ideas in this paper have developed gradually over several years. Valuable discussions with and contributions from J.C. Morud, B. Glemmestad I.J. Halvorsen, K. Havre, T. Larsson, and M.S. Govatsmark are gratefully acknowledged.*

9. REFERENCES

Arbel, A., I.H. Rinard and R. Shinnar (1996). Dynamics and control of fluidized catalytic crackers. 3. designing the control system: Choice of manipulated and measured variables for partial control. *Ind. Eng. Chem. Res.* **35**(7), 2215–2233.

Findeisen, W, F.N. Bailey, M. Brdys, K. Malinowski, P. Tatjewski and A. Wozniak (1980). *Control and coordination in Hierarchical Systems*. John Wiley & sons.

Fisher, W.R., M.F. Doherty and J.M. Douglas (1988). The interface between design and control. 3: Selecting a set of controlled variables.. *Ind. Eng. Chem. Res.* **27**(4), 611–615.

Foss, C.S. (1973). Critique of chemical process control theory. *AIChE Journal* **19**(2), 209–214.

Gordon, M.L. (1986). Simple optimization for dual composition control. *Hydrocarbon Processing* pp. 59–62.

Luyben, W.L. (1988). The concept of eigenstructure in process control. *Ind. Eng. Chem. Res.* **27**(1), 206–208.

Marlin, T.E. and A.N. Hrymak (1997). Real-time operations optimization of continuous processes. In: *Fifth international conference on chemical process control (CPC-5), Lake Tahoe, Jan. 1996*. Vol. 93 of *AIChE Symposium Series*. pp. 156–164.

Mizoguchi, A., T.E. Marlin and A.N. Hrymak (1995). Operations opimization and control design for a petroleum distillation process. *Canadian J. of Chem. Engng.* **73**, 896–907.

Morari, M. (1982). Integrated plant control: A solution at hand or a research topic for the next decade?. In: *Proc. of Second international conference on chemical process control (CPC-2), Sea Island, Georgia, Jan. 1981*. pp. 467–495.

Morari, M., G. Stephanopoulos and Y. Arkun (1980). Studies in the synthesis of control structures for chemical processes. Part I: Formulation of the problem. Process decomposition and the classification of the control task. Analysis of the optimizing control structures.. *AIChE Journal* **26**(2), 220–232.

Narraway, L. and J. Perkins (1994). Selection of process control structures based on economics. *Computers chem. Engng. (Suppl.)* **18**, S511–S515.

Narraway, L.T. and J.D. Perkins (1993). Selection of process control structure based on linear dynamic economics. *Ind. Eng. Chem. Res.* **32**(11), 2681–2692.

Narraway, L.T., J.D. Perkins and G.W. Barton (1991). Interaction between process design and process control: economic analysis of process dynamics. *J. Proc. Cont.* **1**, 243–250.

Rijnsdorp, J.E. (1991). *Integrated Process Control and Automation*. Elsevier.

Shinnar, R. (1981). Chemical reactor modelling for purposes of controller design. *Cheng. Eng. Commun.* **9**, 73–99.

Shinskey, F.G. (1984). *Distillation Control*. 2 ed.. McGraw-Hill Book Company.

Skogestad, S. and I. Postlethwaite (1996). *Multivariable Feedback Control*. John Wiley & Sons.

Tyreus, B.D. (1999). Dominant variables for partial control; part 1 and 2. *Ind. Eng. Chem. Res.* **38**, 1432–1455.

Zheng, A., R.V. Mahajannam and J.M. Douglas (1999). Hierchical procedure for plantwide control system synthesis. *AIChE Journal* **45**, 1255–1265.