MODELING, IDENTIFICATION AND CONTROL, 1988, VOL. 10, NO. 1, 35–51
doi:10.4173/mic.1989.1.3

State space model predictive control of a multistage electrometallurgical process

J. G. BALCHEN†, D. LJUNGQUIST† and S. STRAND†

Keywords: *Optimal control, predictive control, non-linear control system, state-space methods, iterative methods, on-line operation.*

The paper deals with the proposed application of a novel technique for Model Predictive Control to a multistage electrometallurgical process. The novel control technique is based upon high speed repetitive simulation of a non-linear state space model of the process including relevant constraints, and searching in a parameterized control space by an efficient optimization routine until an optimal set of control actions has been found. This MPC-technique is not limited to linear processes with quadratic objective functionals and is therefore believed to offer a major improvement to the control of many industrial processes where the standard, linear control solutions fail because of non-linearities and constraints in the process system.

1. Introduction

Since the time of Norbert Wiener, it has been known that optimal feedback control algorithms for dynamic processes will, in some way or other, reflect the models which describe the behaviour of the process and its environment and the measure of the performance to be minimized. Optimal control as derived by Pontryagin *et al.* (1962) is an example of this. Later, numerous contributions in the literature present control techniques which are model based. The algorithms for process control which were introduced from the mid 1970s, MAC, DMC and IDCOM are more recent examples of this (Richalet *et al.*, 1978; Cutler and Ramaker, 1979; Reid, 1981; Marchetti, Mellichamp and Seborg, 1983; Mehra and Mahmood, 1985; Garcia and Prett, 1986; Clarke, Mohtadi and Tuffs, 1987). The latter type of control algorithms, often referred to as Model Predictive Control (MPC), computes the process variables over a future time interval. The control variables which will minimize a performance criterion (objective functional) over this time interval, can be determined. If the process is regarded as linear and the objective functional is quadratic, a closed form solution is found which is easy to implement in a computer system. This is not the case if the process is non-linear, has a non-quadratic objective functional or is constrained. This has been generally known for a long time.

A simple way of deriving optimal control of a general dynamic process (non-linear and constrained) is to parameterize the control variables in a simple manner and optimize the control variables by searching in this parameter space by simulat-

---

Received November 1988.

This paper was presented at the IFAC Workshop on Model Based Process Control, Atlanta, Georgia, USA, June 13–14, 1988, and is reprinted in MIC with permission from IFAC.

† Division of Engineering Cybernetics, The Norwegian Institute of Technology, Trondheim, Norway.

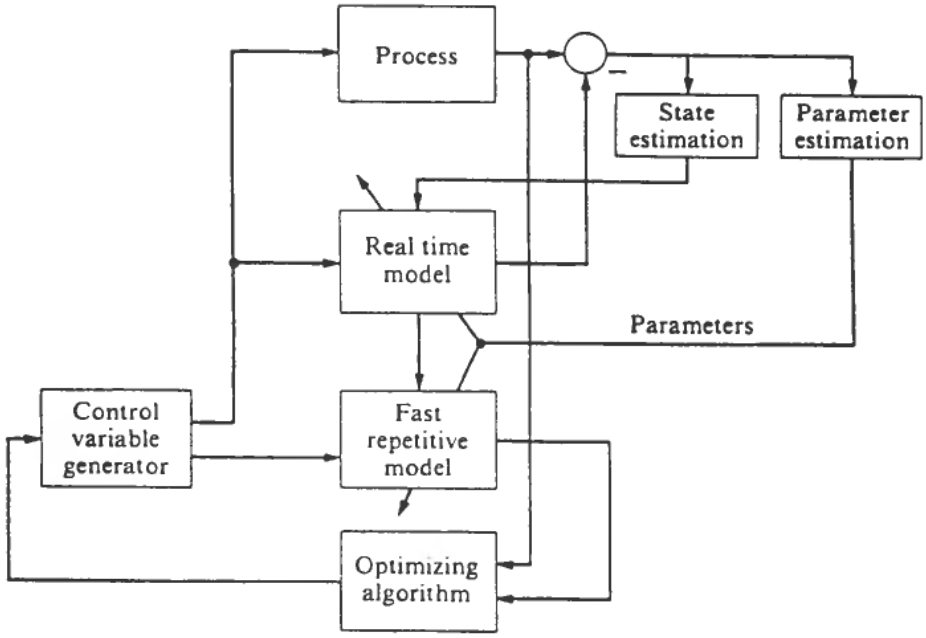

Figure 1. Components of the total system for predictive control.

ing the process behaviour over a future time interval using a fast repetitive computer (Strand 1987). This is a promising technique which has been employed in this paper. An alternative technique is to apply the Pontryagin Maximum Principle (Ljunguist 1986) which, due to its more efficient mathematical technique, yields better results in certain cases.

A block diagram illustrating the model-based repetitive dynamic optimization technique is shown in Fig. 1. It is based upon a state space model of the process which is used in two modes (any kind of dynamic model which can be reset can be used). One model is used in a combined state- and parameter estimator which updates the model states and parameters to agree with the behaviour of the real process. This makes it possible to adjust the other model which is running repetitively at a much higher speed. This second model can be identical to the first or may be reduced in complexity. The major prerequisite for the present method is the availability of a high speed on-line computing capacity. This should allow repetitive simulation and optimization calculations to be done so fast that the search in the control parameter space converges nearly instantaneously compared to the dynamics of the real process. Such computer capacity is available today at a reasonable price, which makes this technique applicable to most processes encountered in the process industry. In case the process dynamics is too fast for the simple searching technique suggested, a modified algorithm is available which employs a LQG-strategy in parallel with the searching strategy (see Balchen *et al.* 1988).

2. Properties of the present method

The properties of the present method compared to conventional process control techniques and MPC strategies based upon a linear model are briefly as follows:

The process (and its model) may be non-linear, highly interacting and with constraints on both control variables and state variables provided the optimization algorithm converges. This becomes a matter of the relative bandwidth (speed) of the optimization and the process. A similar statement is out of the question for other available process control strategies.

The technique is very straightforward and obvious and directly appeals to operators because of its simplicity.

The objective functional to be employed should be based on economic aspects (profit functions) including the cost of raw materials, energy, labour, maintenance etc. and value of the products.

The present technique, as well as other MPC techniques, assumes the availability of a reasonably good model of the process which is continuously updated against the real process. In many cases of industrial practice, this requirement is not controversial at all. In other cases it will limit the applicability of the method.

3. The application of model predictive control to a multistage electrometallurgical process

A simplified block diagram indicating how the multistage electrometallurgical process operates, is shown in Fig. 2. Each unit is identical and all units in the same row are electrically connected in series. Each unit is being fed with the raw material which appears in two aggregates, A and B. These aggregates are chemically similar, but differ in energy content. The product of this process is a valuable metal and a less valuable gas. Each of the units can be manipulated by the flow of the two different aggregates of feed, by the electric current through the unit and additional

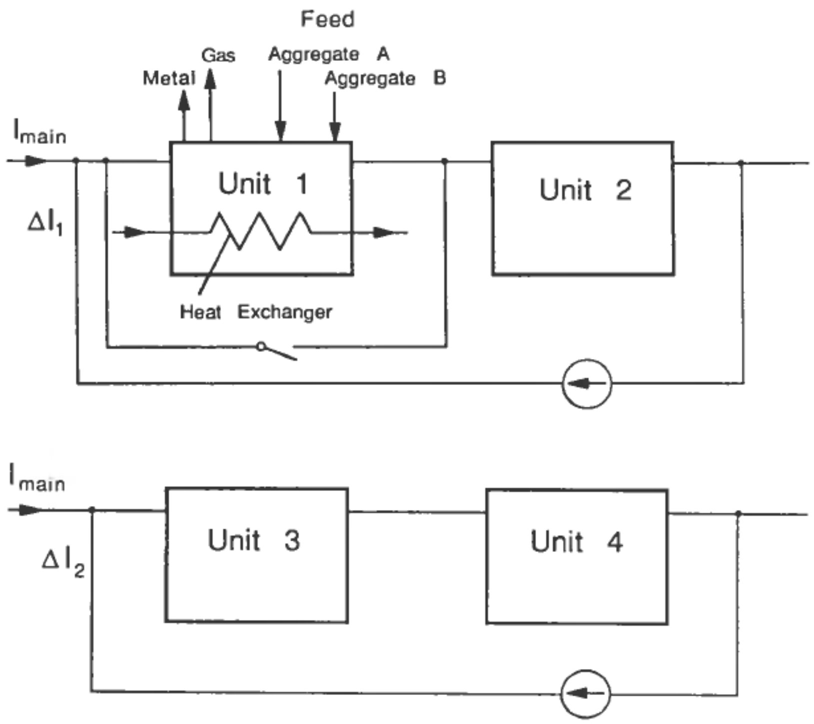

Figure 2. The multistage electrometallurgical process.

heating through a special circuit. The main objective of the control system is to maintain the mass and energy balances in each unit, that means maintaining the concentration of the active material and the temperature at optimal values. Due to operational reasons, the temperature must be kept between rather narrow constraints. The energy balance is influenced by the ratio of the two aggregates of feed, by the electric current and the additional heating. Additional heating is undesirable because it is unproductive. The yield of the process is not constant, but influenced by certain impurities in the feed, and must be regarded as a disturbance. Since all units in one row are connected in series, the electric current may be regarded as both a control variable and a disturbance. Thus this process becomes very interesting for the type of control presented in this paper, since

The control actions are discontinuous (quantised).

There are severe constraints on both control variables and states.

The units are highly interacting and interdependent.

All units have identical models, but with different parameters and states.

The objective functional of the system is easily established in terms of economic factors.

4. The model

First a detailed, precise description of the process dynamics was developed. The result was a model containing about ten state variables for each unit and several unknown non-linear functions and unknown parameters. Model reduction to five or six states per unit could be done by aggregation of related effects. Simulation results showed that even two states represented the dynamics of an *undisturbed* unit in a satisfactory way. These results seemed to fit perfectly with the control strategy illustrated in Fig. 1.

A simplified model with three state variables per unit is to be used in the optimization.

In the identification/estimation algorithm two or three extra state variables are used to make the model fit the real process.

The extra variables may be regarded as time-varying parameters in the simplified model. The almost unpredictable dynamics of the contents of inpurities and the other disturbances in the units however, make the assumption of constant parameters during the optimization horizon nearly as good as any other assumption. On-line estimation of these parameters then assures updated estimates before every new optimization. The slow dynamics of the actual process makes this strategy applicable even with a conventional personal computer.

The simplified model may be developed based upon elementary energy and mass balances.

(i) Energy balance in unit i

The following factors are taken into account in the energy balance:

energy due to the electric current that flows through the unit

additional heating

energy contents in the feed aggregate A
energy contents in the feed aggregate B
energy contents in the produced metal removed at certain points in time
energy in the escaping gas
energy loss to the surroundings
other losses

Aggregate A contains more energy than aggregate B.

(ii) Mass balance of the active component in unit i

The rate of change of the active component in the unit equals the difference between the sum of the two aggregates of the feed and the consumption of the component in the reaction of the unit. Moreover, some of the active component is inevitably removed each time the product is removed.

(iii) Mass balance, total mass in unit i.

Aggregates A and B of the feed come into the unit, whilst the metal product together with a small amount of feed is tapped off and the gas is continuously escaping. This leads to the following differential equation (nomenclature listed at the end of the paper):

$$\begin{aligned}
\dot{x}_{1i} &= \frac{1}{b_{1i}}\,[f_{1i}I_i^2 + u_{3i} - a_1(x_{1i} - a_2) + a_4(a_5 - x_{1i})u_{1i} \\
&\quad - (a_3 + a_4 x_{1i})u_{2i} + (a_6 - a_7 f_{2i})I_i] \\
\dot{x}_{2i} &= \frac{(a_{13} - x_{2i})^2}{b_{4i}a_{13}}\,[a_{14}(u_{1i} + u_{2i}) - a_{15} f_{2i} I_i - a_{16} x_{2i} f_{3i}] \\
\dot{x}_{3i} &= a_{14}(u_{1i} + u_{2i}) - a_{17} f_{2i} I_i - (a_{18} + a_{16} x_{2i}) f_{3i} \qquad (1)
\end{aligned}$$

where

$$\begin{aligned}
f_{1i} &= b_{2i} + a_8 x_{2i} - a_9 x_{1i} - a_{10} x_{1i} x_{2i} \\
f_{2i} &= b_{3i} - a_{11}(x_{1i} - a_{12}) \\
f_{3i} &= x_{3i} - \frac{b_{4i}a_{13}}{a_{13} - x_{2i}}
\end{aligned}$$

Each unit may be electrically short-circuited. If this is done to unit number i, the effect is that $I_i \equiv 0$. I_i therefore represents three control variables; the main current being fed through all the units, additional current fed through all units in a series and the possibility of short circuiting each unit individually. The aggregate A of the feed is fed in batches (u_{1i}), whilst the other control variables are continuous. Because of the nature of the process equipment, some of the control actions are implemented manually by the operators.

In the model (1), three different constants appear. The a_j's are constants common to all the units, while the b_{ji}'s may vary from unit to unit. Moreover f_{ji} indicates time-varying functions; f_{1i} equals the internal electrical resistance of the unit, f_{2i} equals the yield factor and f_{3i} is the predicted mass removal. f_{1i} and f_{2i} are very simplified expressions. Therefore b_{2i} and b_{3i} have to be updated in the estimation algorithm, and will typically change from one optimization to the next.

5. The objective functional

The simple model presented in the previous section can only be justified when the concentration of the active component is kept between 7 and 13 per cent, and it is desirable to keep the concentration in the middle of this range. For operational reasons the temperature must be kept between narrow limits. Equation (1) is scaled such that the temperature varies around 1·0, and this relative temperature is to stay within ± 3 per cent.

These two constraints are taken care of by means of penalty functions in the objective function. The other terms in the objective function reflect economic considerations:

$$J = \int_{t_0}^{t_0+T} \left[\sum_{i=1}^{n_u} L_i \right] dt \tag{2}$$

$$\begin{aligned} L_i = & -\beta_0 f_{2i} I_i + \beta_1((R_{yi} + f_{1i})I_i + e_i)I_i \\ & + \beta_2 u_{1i} + \beta_3 u_{3i} + \beta_4 u_{2i} + \beta_5 \dot{x}_{1i}^2 \\ & + c_{1i}(x_{1i}) + c_{2i}(x_{2i}) \end{aligned} \tag{3}$$

$c_{1i}(\cdot)$ and $c_{2i}(\cdot)$ are the penalty functions which keep the temperature and concentration respectively within the given boundaries. Different functions may be used, but their value should be nearly negligible in the permitted area and increase rapidly when the boundaries are reached (or just before). A penalty function suggested by Balchen (1984) has the desired properties. An additional linear penalty on relative temperatures above 1·0 reflects increased unit maintenance.

The objective functional (2) has two important qualities:

Each of the objective functions (L_i) represent economic factors (expenses or incomes) or penalties with scaling factors to be chosen.

The economic scaling factors are obvious whereas the scaling factors of the penalty functions are easily tuned until the constraints are maintained in a satisfactory way.

This results in an optimal control problem which can be solved using an efficient searching algorithm.

6. Solution strategy

Given an arbitrary, non-quadratic objective functional and a non-linear state-space model, there are several ways to reach the optimal choice of the control vector. The following outlines the strategy used to obtain the results in this paper.

The original problem may be formulated as:

$$\min_{\boldsymbol{u} \in U} J(\boldsymbol{u}(t))$$

$$J(\boldsymbol{u}(t)) = S[\boldsymbol{x}(t_0 + T)] + \int_{t_0}^{t_0+T} L(\boldsymbol{x}(t), \boldsymbol{u}(t), t)\, dt \tag{4}$$

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{u}(t), t)$$

where U is the space of admissible controls, i.e. the control variables may be subject to hard constraints. In the present case, hard constraints are not applied to the state variables, since we are using penalty functions to limit them.

To make this original problem formulation numerically solvable, the control variables have to be parameterized, leading to somewhat different formulations and to suboptimal solutions of the given problem. A piecewise constant parameterization of each control variable is chosen since this is a practical way of implementing the control signal in the process. Thus the problem may be stated as follows:

$$\begin{aligned}
&\min_{\boldsymbol{\theta} \in \boldsymbol{\Theta}} J(\boldsymbol{\theta}) \\
&J(\boldsymbol{\theta}) = S[\boldsymbol{x}(t_0 + T)] + \int_{t_0}^{t_0+T} L(\boldsymbol{x}(t), \boldsymbol{\theta}, t)\, dt \qquad (5) \\
&\dot{\boldsymbol{x}}(t) = \boldsymbol{f}(\boldsymbol{x}(t), \boldsymbol{\theta}, t)
\end{aligned}$$

where $\boldsymbol{\theta}$ is the set of parameters describing the control variables. Now the original problem (4) has been turned into a finite-dimensional numerically solvable optimization problem.

The dimension of the parameter space has a great influence upon the solution speed, and should be kept to a minimum. However it is known that reduction of the number of parameters describing each control variable will make the corresponding suboptimal solution less satisfactory compared to the optimal solution of the originally stated problem in (4). Hence the choice of parameterization is a trade-off between solution speed and the quality of the suboptimal solution achieved.

In the present example as few as three parameters for each control variable are used. This turns out to be a suitable number in process control cases where trajectory tracking is not required. The discretization intervals do not have to be equal lengths (Balchen *et al.* 1988).

Moreover a quite simple optimization technique is used. The gradient of $J(\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$ is computed experimentally, perturbing each parameter in turn and measuring the corresponding change in $J(\boldsymbol{\theta})$. The search in the admissible parameter space is performed using a modified Newton algorithm, which aggregates the information retrieved by sequential computation of the gradients at points in the parameter space, in an approximation to the inverse Hessian. Special care is taken for parameters which have reached their constraints, according to the constraints on the control variables. In order to reduce the number of gradient computations, the line search in each direction is done quite accurately.

This optimizing technique is more time-consuming than those more mathematically advanced, such as the Pontryagin Maximum Principle. But it is easy to implement in a parallel processing system, since the gradient components may be computed on different processors. Furthermore it is an advantage that process operators are able to understand how optimal control is computed using this technique. Finally, if the solution speed is too slow compared with the process dynamics, which is not the case in the present example, a LQG-strategy with immediate feedback around the computed optimal trajectory is suggested (Balchen *et al.* 1988).

7. Simulation experiments

As pointed out in the presentation of the model, there are several uncertainties related to the parameters of the simplified model. Simulation experience shows that the inaccuracies related to the flow of feed and the removal of the produced metal

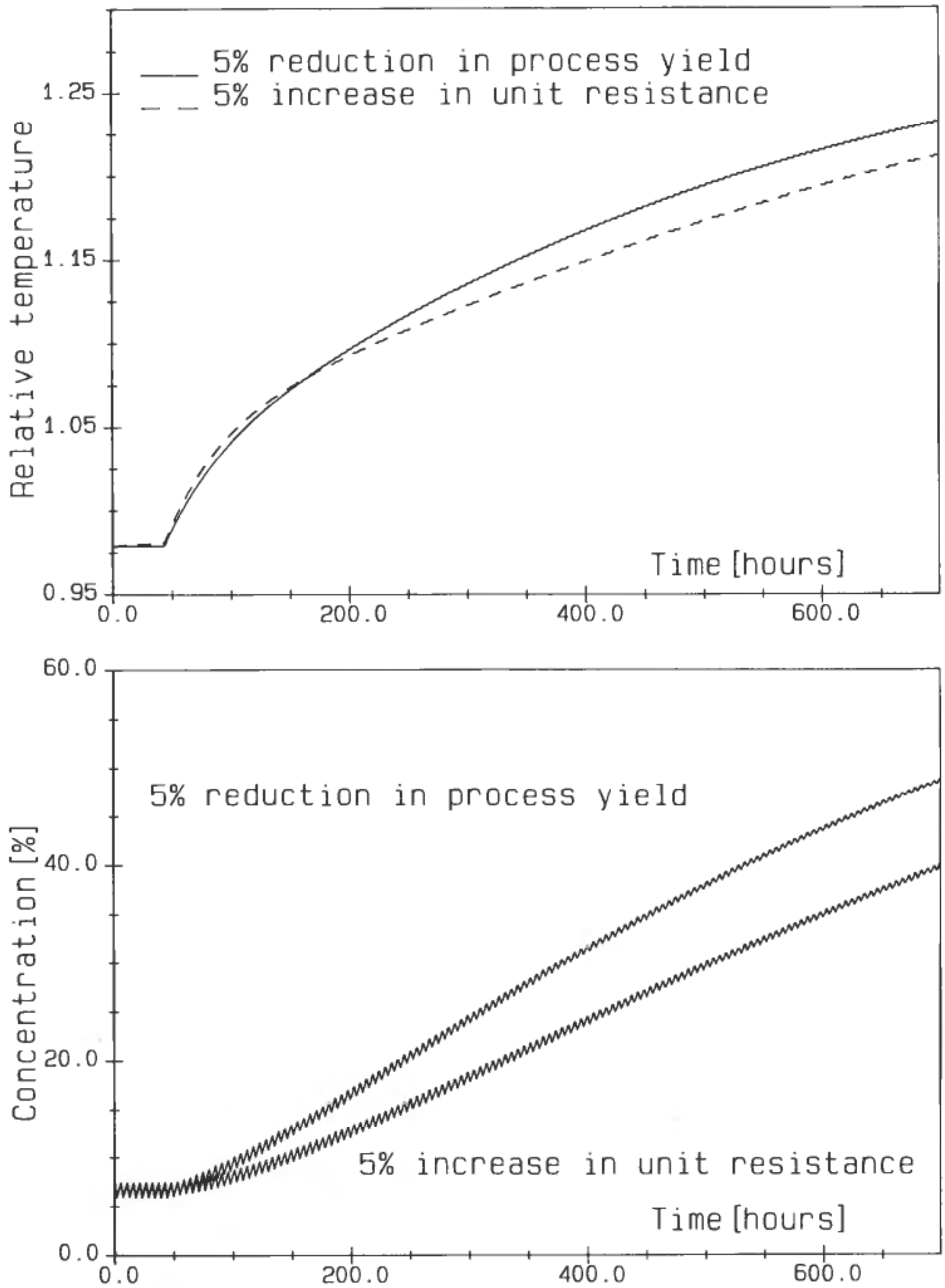

Figure 3. (*a*) and (*b*) Step responses, control variables unchanged.

only contribute marginally to the process dynamics. Changes in the electrical resistance and the yield factor of each unit, however, represent considerable disturbances as can be seen from Fig. 3(*a*) and 3(*b*).

The two curves in Fig. 3(*a*) illustrate the temperature responses due to 5 per cent reduction in the yield factor (at $t = 45$ h) and 5 per cent increase in the electrical

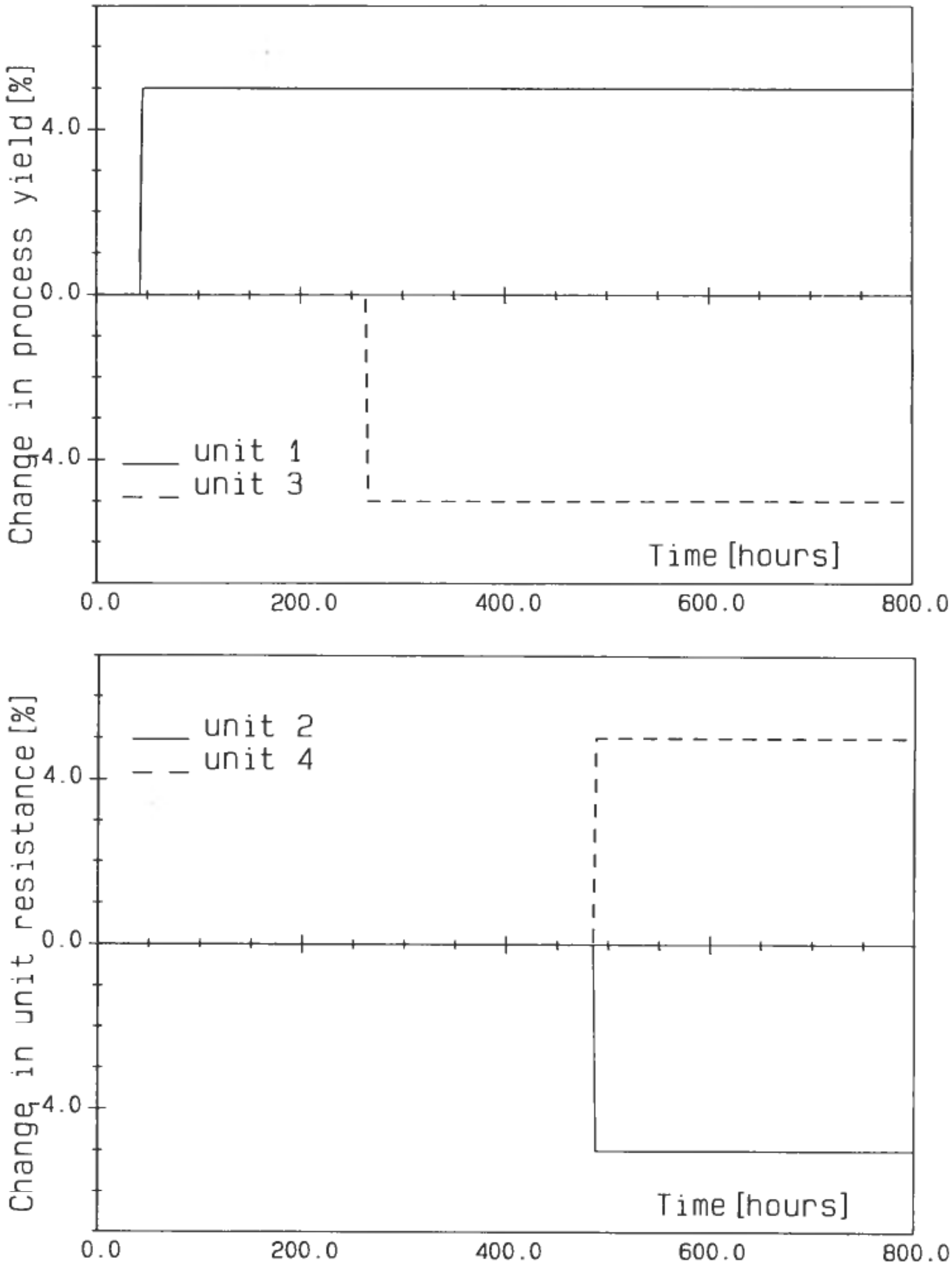

Figure 4. (*a*) Disturbances in process yields; (*b*) disturbances in unit resistances.

resistance in the unit when the control variables are kept unchanged. Figure 3(*b*) shows the corresponding changes in concentration of the active component in the unit. Both temperature and concentration must remain within narrow limits, however, therefore the control variables have to change.

In the simulations, 4 units are studied, units 1 and 2 in row number 1 and units 3 and 4 in row number 2. The electrical current can be somewhat different in

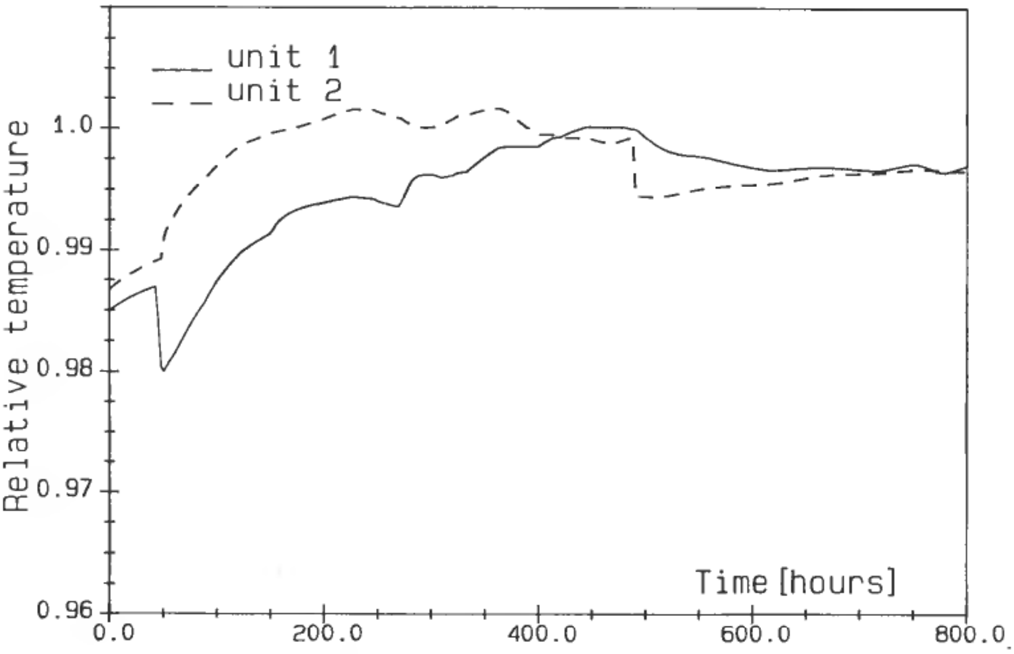

Figure 5. Unit temperatures.

each row (see Fig. 2). Units 1 and 3 are exposed to disturbances in the yield factors, increase and reduction of 5 per cent respectively, while units 2 and 4 are disturbed by decrease and increase of 5 per cent respectively in the electrical resistance of the unit. The step disturbances occur at the points in time illustrated in Fig. 4(*a*) and 4(*b*). It is assumed that the real time estimator described in Fig. 1 is able to give

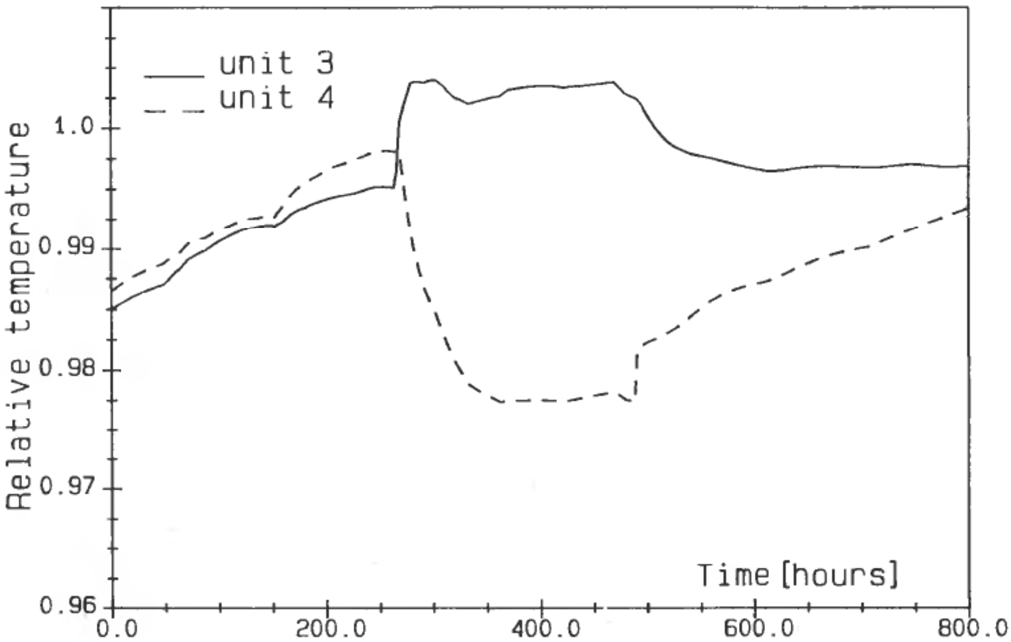

Figure 6. Unit temperatures.

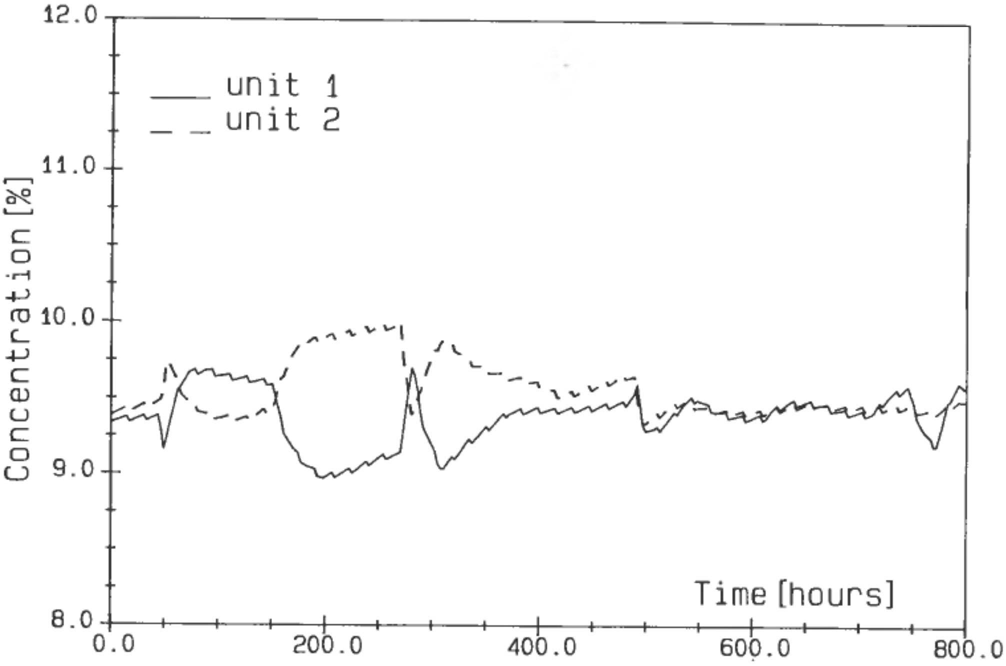

Figure 7. Unit concentrations.

estimates of these disturbances within a period of 2·7 h. The optimization horizon is 48 h. Moreover the control parameters are updated every 2·7 hour.

The disturbances indicated above are not supposed to be reasonable from a process engineer's point of view, but it is known that the step sizes are reasonable

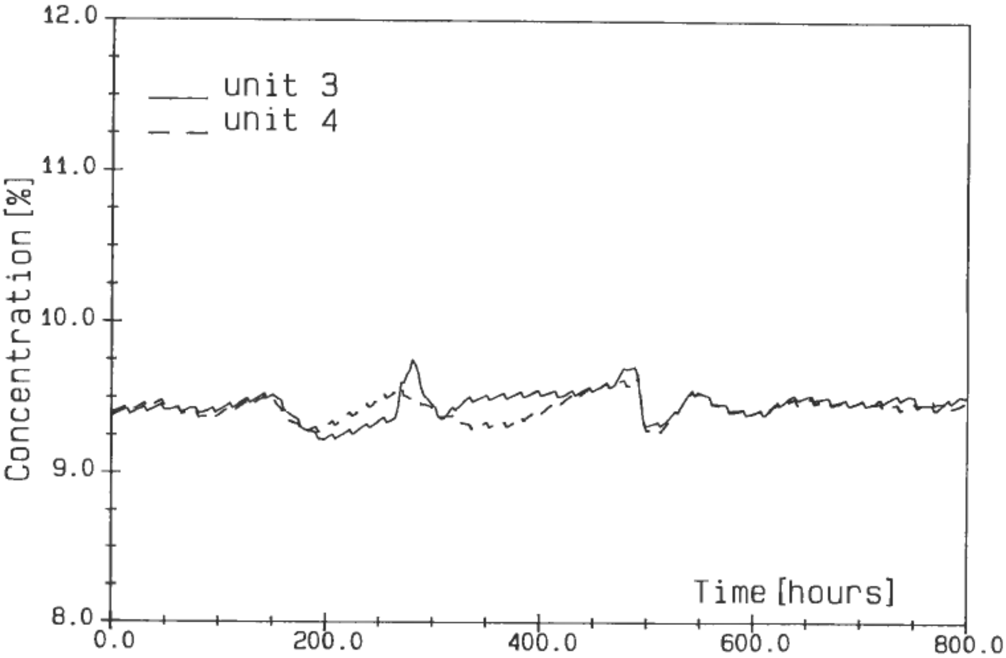

Figure 8. Unit concentrations.

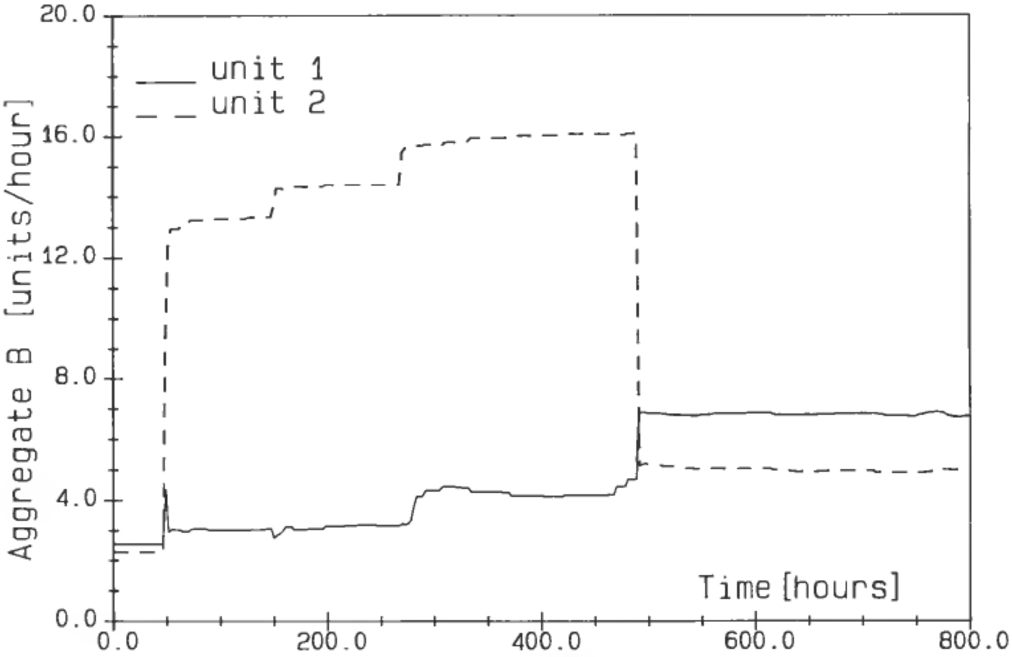

Figure 9. Use of aggregate B.

and that the disturbances will be a challenge to the optimization algorithm. The reader should be aware that the dynamic behaviour of the yield factor and the electrical resistance of the unit, still apply in the simplified model, i.e. just the parameters b_{2i} and b_{3i} in (1) are changed in steps. The exact numerical values of the

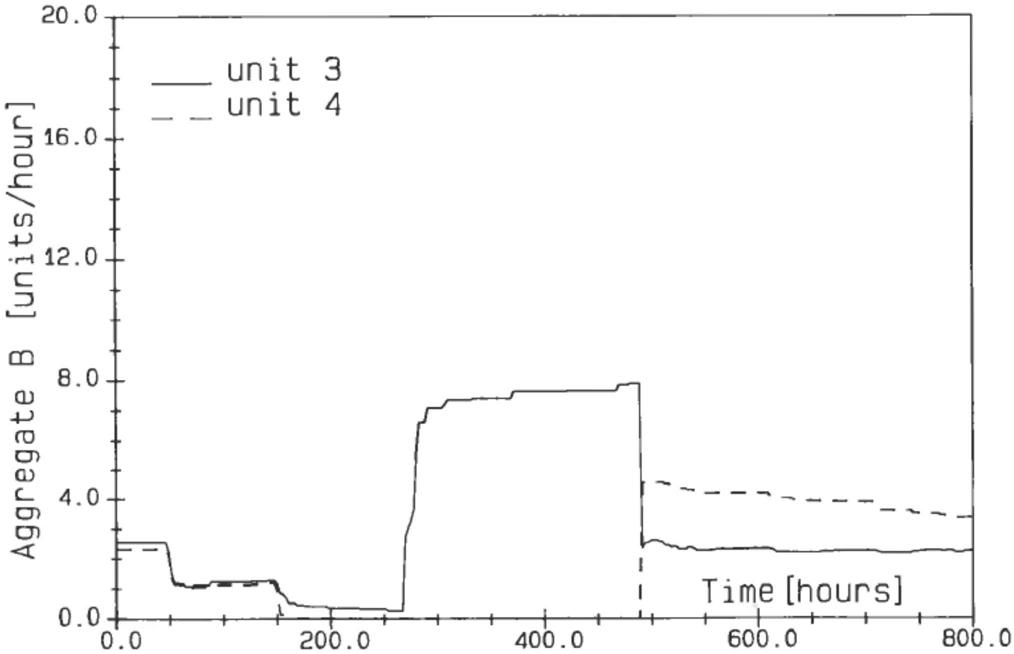

Figure 10. Use of aggregate B.

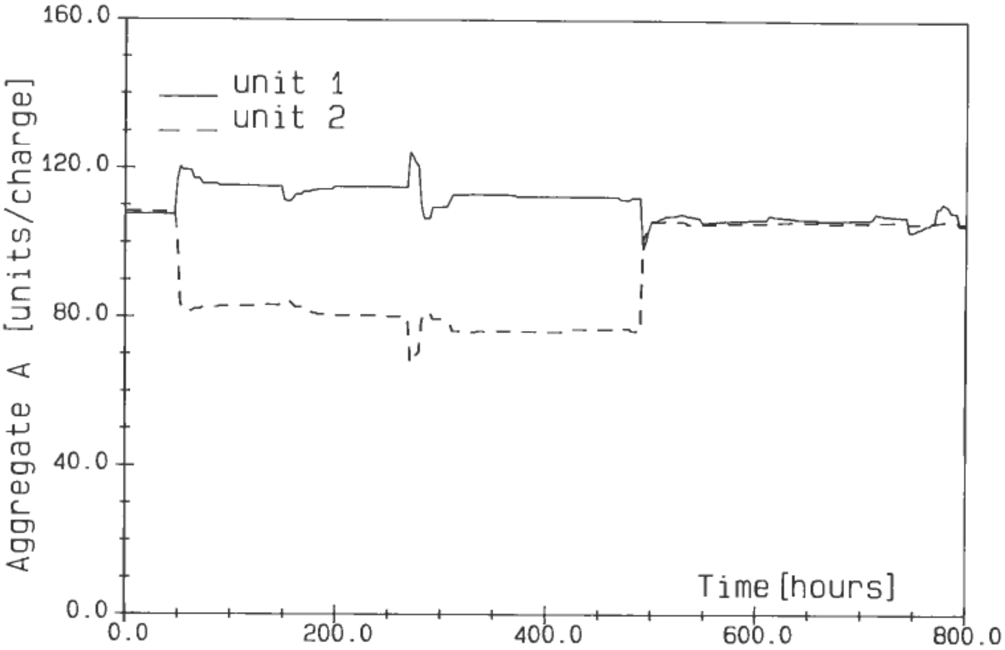

Figure 11. Use of aggregate A.

different prices in the objective functional are not important to a fundamental interpretation of the optimization results.

Many conclusions can be drawn from the simulation experiments. However, we wish to limit the amount of comments by just pinpointing the most important results obtained (see Figs. 5–16).

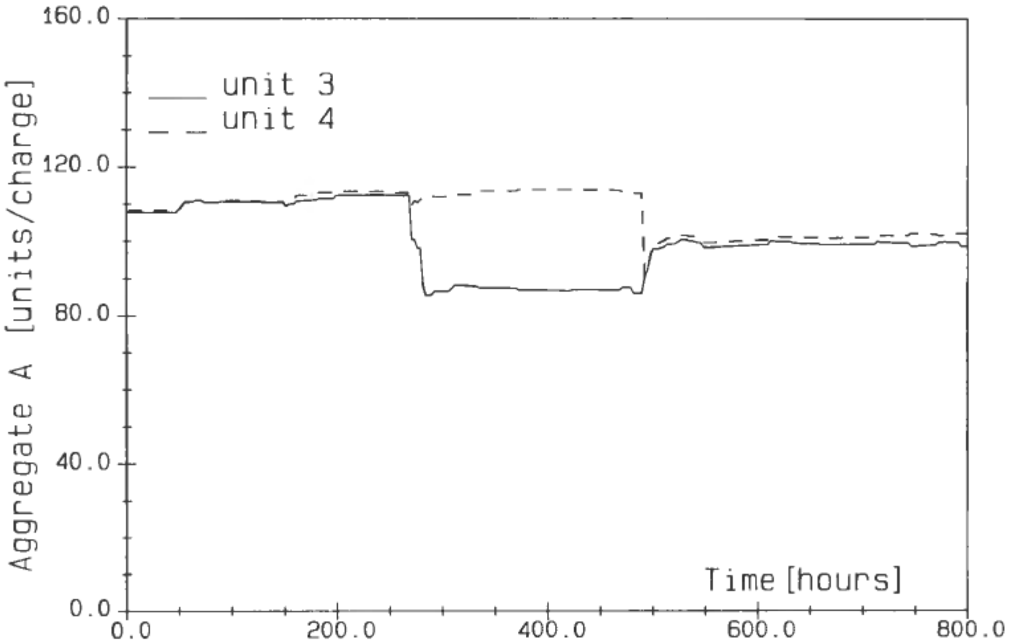

Figure 12. Use of aggregate A.

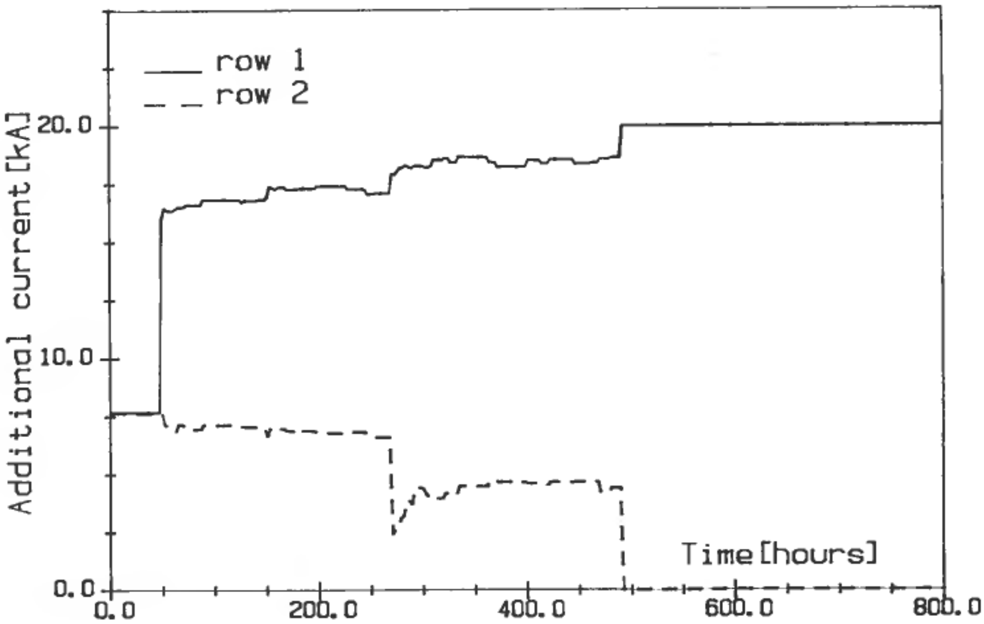

Figure 13. Use of row current.

Increase in the yield factor in unit 1 calls for an increase of electrical current in this row. The temperature in unit 2, however, increases as the row current increases. Thus the current cannot increase very much if the resistance in unit 2 does not decrease for some reason (Fig. 13).

The increase of current in units 1 and 2 demands increase in the feed. The responses in the feed of the two different aggregates A and B due to the step

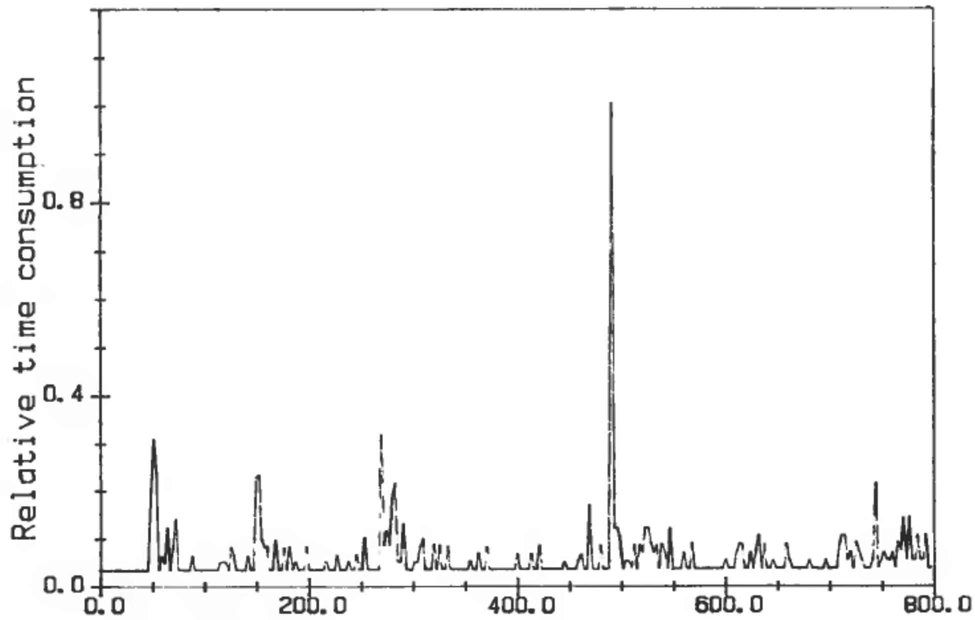

Figure 14. Computational effort in optimization.

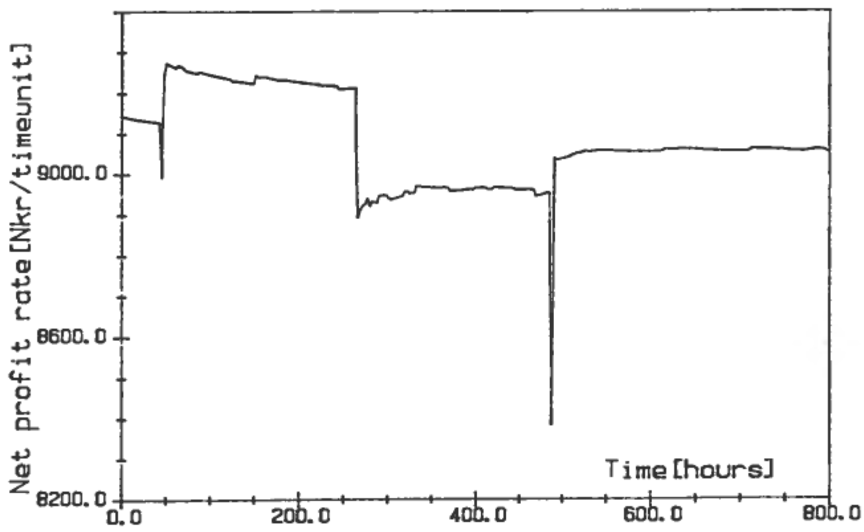

Figure 15. Net profit rate from 4 units.

changes in yield factor in unit 1 and resistance in unit 2 reflect the importance of the energy balance, and the fact that aggregate A contains more energy than aggregate B (Figs. 9, 11 and 13).

Similar statements are valid for the control variables in units 3 and 4 (Figs. 10, 12 and 13).

In the present experiments there is no need for adding auxiliary heat or short circuiting the units, but such conditions may occur under severe disturbances.

The main peaks in the computational effort (Fig. 14) occur immediately after the disturbances are detected, except the one at $t = 150$ h. The latter is caused by

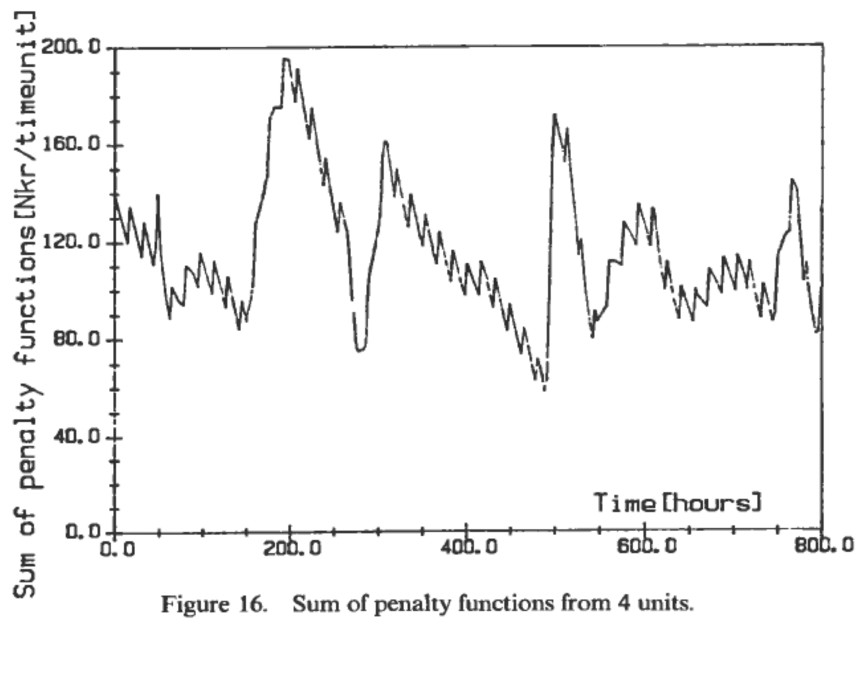

Figure 16. Sum of penalty functions from 4 units.

the objective functional being flat near optimum and an imperfect searching algorithm. Figure 14 is scaled from 0·0 to 1·0. With the present algorithm installed on an IBM AT, 1·0 corresponds to 10 min. Updating the control parameters with a 2·7 h interval, this means that more than 50 units can be controlled by an IBM AT. This updating interval is small compared to the process dynamics, and a parallel LQG-strategy is not required.

Figure 15 shows the net profit rate of the process. The step disturbances have direct consequences for this rate. The peaks are due to the time-delay in the estimation of the disturbances. As shown in Fig. 16 the sum of penalty functions is nearly negligible.

The following statements are justified by the simulation experiments.

A reasonable objective functional based upon economic considerations and physical interpretation is easily derived.

When the state variables are highly coupled through the dynamics and the control variables it is not straightforward to work out how the control variable should change due to physical disturbances or changes in the operational conditions. It is not difficult, however, to understand the optimality of the new controls derived by the present strategy.

8. Conclusion

From an economic point of view optimal control should be based on non-quadratic objective functionals. The corresponding control variables will drive the process over a wide range of operational conditions. Therefore in many cases non-linear models and constraints have to be considered. Existing control strategies cannot handle such problems in a satisfactory way. The proposed method, however, permits the flexibility required as illustrated by a simple simulation experiment.

ACKNOWLEDGMENT

This work has been sponsored in part by The Royal Norwegian Council for Scientific and Industrial Research through the Predictive Control Research Program.

Nomenclature

x_{1i}	temperature in unit number i
x_{2i}	concentration of active component in unit number i
x_{3i}	total mass of unit number i
I_i	the current through unit i
u_{1i}	feedrate of aggregate A to unit i
u_{2i}	feedrate of aggregate B to unit i
u_{3i}	additional heating to unit i
t_0	the initial time for the optimization
J	the objective functional
L_i	the objective function (cost function) for unit number i
T	the optimization horizon

$S(x(t_0 + T))$	a scalar function weighting the final value of the state vector
n_u	number of units
β_i	prices corresponding to the different terms in the objective functional (in Nkr)
a_j, b_{ji}	model parameters, i indicates unit No.
f_{1i}	internal resistance, unit i
f_{2i}	yield factor, unit i
f_{3i}	predicted mass removal, unit i
R_{yi}, e_i	parameters used to compute the voltage across unit i.

References

Balchen, J. G. (1984). A quasi-dynamic optimal control strategy for non-linear multi-variable processes based upon non-quadratic objective functionals. *Modeling, Identification and Control*, **5**, 195–209.

Balchen, J. G., Ljungquist, D. and Strand, S. (1988). Predictive Control Based upon State Space Model, 1988 American Control Conference.

Clarke, D. W., Mohtadi, C. and Tuffs, P. S. (1987). Generalized predictive control. Part I and II *Automatica*, **23**, 137–160.

Cutler, C. R., and Ramaker, B. L. (1979). Dynamic Matrix Control—a Computer Control Algorithm. *AIChE 86th National Mtg*, Houston, TX, Apr. 1979.

Garcia, C. E., and Prett, D. E. (1986). Advances in Industrial Model Predictive Control. *CPC III*, Asilomar, Calif., Jan. 1986.

Ljungquist, D. (1986). Predictive Control Based on Pontryagin's Maximum Principle (in Norwegian), Thesis (Sivilingeniør). Division of Engineering Cybernetics, The Norwegian Institute of Technology, Trondheim. Dec. 1986.

Marchetti, L., Mellichamp, D. A., Seborg, D. E. (1983). Predictive control based on discrete convolution models. *Ind. Eng. Chem. Proc. Des. Develop.*, **22**, 488–495.

Mehra, R. K., and S. Mahmood. (1985). Model Algorithmic control, chapter 15 in Distillation Dynamics and Control, ed. P. B. Desphande (Instruments Society of America).

Pontryagin, L. S., Boltianskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F. (1962). *The Mathematical Theory of Optimal Processes.* (J. Wiley & Sons).

Reid, J. G. (1981). Output predictive algorithmic control: precision tracking with application to terrain following. *AIAA, Journal of Guidance and Control*, Vol. 4, No. 5, Sept–Oct, 1981.

Richalet, J. A., Rault, A., Testud, J. D. and Papon, J. (1978). Model Predictive Heuristic Control: Applications to Industrial Processes, *Automatica*, **14**, 413–428.

Strand, S. (1987). System for Predictive Control Based on Repetitive Simulation of a Dynamic Model with a Parameterized Control Vector (in Norwegian), Thesis (Sivilingeniør), Division of Engineering Cybernetics, The Norwegian Institute of Technology, Trondheim. March, 1987.