PREDICTIVE CONTROLLER DESIGN FOR SISO SYSTEMS

Paul R. Maurath, Duncan A. Mellichamp, and Dale E. Seborg

Department of Chemical and Nuclear Engineering, University of California, Santa Barbara

ABSTRACT

This paper presents a fundamental analysis of the stability of single-input/single-output closed-loop systems with predictive controllers. The analysis can incorporate modeling errors and be used to calculate allowable modeling errors for a given system and controller. Design parameter selection guidelines for predictive controllers in SISO systems have been developed by considering performance, robustness, and ease of tuning. The performance and robustness of the resulting controllers are demonstrated on four numerical examples and compared to controllers designed using other parameter choices.

INTRODUCTION

In the late 1970's, two new control algorithms emerged in the open literature which approached the problem of process control somewhat differently than before. Model Algorithmic Control (MAC) (Richalet et al., 1978) and Dynamic Matrix Control (DMC) (Cutler and Ramaker, 1980) both: i) utilize a discrete convolution-type model to represent the process, and ii) control the process by optimizing the process output(s) over some finite future time interval. These techniques share several characteristics in addition to those mentioned above and have constituted a new category of process control referred to as "predictive control". A number of papers have appeared on these and related predictive control algorithms (Maurath, 1985).

The goal of this paper is to present a series of guidelines which will lead to reasonable starting values for the design parameters in a predictive controller. The present paper is restricted to unconstrained single-input, single-output systems. A related paper (Maurath et al, 1985) is concerned with multiple-input, multiple-output systems.

A brief overview of predictive control will be given followed by a stability analysis which can be applied in the presence of plant/model mismatch. The effects of controller design parameters on closed-loop system performance and robustness will be explored with their implications for controller design. Recommendations for controller design parameters for SISO predictive control systems will be presented and compared to some other potential design parameter choices.

DEVELOPMENT OF THE SISO PREDICTIVE CONTROLLER

The form of the predictive controller which will be used in these studies closely parallels that used in DMC. We have chosen DMC rather than MAC since more detailed information about DMC is available in the open literature. However, the available publications on DMC have presented a description of the algorithm but very few guidelines on design parameter selection. The effects of the design parameters on the performance of the system are not always obvious to the casual observer.

The predictive controller is based on the following discrete convolution model of the process:

$$y_k = \sum_{i=1}^{N} h_i u_{k-i} + d \tag{1}$$

where y_k is the system output at the k-th sampling time; u_{k-i} is the system input at sampling time k-i; d is the steady-state value of y when u=0. The model parameters $\{h_i\}$ are the values of the impulse response for the system in eq. 1. They are related to the step response coefficients $\{a_i\}$ by the expression:

$$h_k = a_k - a_{k-1} \tag{2}$$

In eq. (2), N is the number of terms in the sampled impulse response sequence which has been retained.

The development of the predictive controller formulation is based on the minimization of the error from the set point of the process output at a finite number (R) of future sampling instants. A more detailed development is available in Marchetti et al. (1983). The optimization problem can be formulated mathematically using the following performance index:

$$J = \underline{E}^T Q \underline{E} \tag{3}$$

where $\underline{E}$ is a vector of length R, and Q is an RxR positive-definite weighting matrix. The elements of $\underline{E}$ are differences between the set point and the predicted output of the process at the next R sampling instants in the future, i.e., $\underline{E}^T$ = col $[E_{k+1}, \ldots, E_{k+R}]$. These could be calculated using eq. 1 or by an analogous method. The variables used to minimize this index are the values of the system input at the next L sampling times, $u_k, \ldots, u_{k+L-1}$. Thus there are L degrees of freedom in the optimization. By utilizing eqs. 1 and 2, the effects of past and future inputs on the performance index can be separated as follows:

$$\underline{E} = \underline{E}' - A\,\underline{\Delta u} \qquad (4)$$

where

$$\Delta u_k = u_k - u_{k-1} \qquad (5)$$

$$\underline{\Delta u} = [\Delta u_k, \Delta u_{k+1}, \Delta u_{k+2}, \ldots, \Delta u_{k+L-1}]^T \qquad (6)$$

$$\underline{E}' = \underline{y}_{sp} - \underline{y}_{o1}$$

$\underline{E}'$ contains the effects of past inputs on the process at the next R sampling instants, $\underline{y}_{sp}$ is a vector of set points for the next R sampling instants, and $\underline{y}_{o1}$ is the prediction of the process output using eq. 2, assuming that no future control action is taken. The $\underline{\Delta u}$ vector contains the next L input changes. The $\underline{A}\,\underline{\Delta u}$ term includes the effects of the current and future inputs. A is the RxL "Dynamic Matrix" (Cutler, 1983) and is constructed from the process step response in the following way:

$$A = \begin{bmatrix} a_1 & & & & \mathbf{0} \\ a_2 & a_1 & & & \\ a_3 & a_2 & a_1 & & \\ \vdots & & & \ddots & \\ \vdots & & & & a_1 \\ \vdots & & & & \vdots \\ a_R & \cdots & \cdots & \cdots & a_{R+L+1} \end{bmatrix} \qquad (8)$$

It is important to note that the A matrix contains information about the process dynamics over the next R sampling intervals. Note that complete dynamic information about the process (i.e., $a_1, a_2, \ldots, a_n$) is not required in the optimization, but is required in the calculation of the error vector $\underline{E}'$.

The changes in inputs which result from optimal control formulations can often be unacceptably large. In order to prevent excessive changes, weighting of input changes in the performance index has been proposed (Marchetti et al., 1983). Utilizing the above results, the performance index can be written as follows:

$$J = (\underline{E}' - A\,\underline{\Delta u})^T Q(\underline{E}' - A\,\underline{\Delta u}) + f\,\underline{\Delta u}^T\underline{\Delta u} \qquad (9)$$

where f is a non-negative scalar called the "move suppression factor". The resulting optimal control law is:

$$\underline{\Delta u} = (A^TQA + fI)^{-1}\,A^TQ\underline{E}' = K\underline{E}' \qquad (10)$$

A key design decision is to select appropriate values for the optimization horizon R, the control horizon L, the move suppression factor f, and the weighting matrix Q.

The control law in eq. 10 generates a series of L control movements which minimizes the predicted error in the process output at the next R sampling instants. Thus the dimensions of the matrices have physical significance. For the above solution to exist, the A matrix must be of full rank. This requires that R be greater than $L + \theta/T_s$ rounded up to the nearest integer, where θ is the process time delay and T_s is the sampling period. It should be noted that the choices, Q = I and f = 0, result in a one-step deadbeat controller, which is not stable for most practical systems.

If the next L control moves were implemented without feedback from the process, modeling errors and disturbances would go undetected for that period of time. This constitutes a potential problem in practical application of the predictive controller. The DMC implementation of predictive control as presented by Cutler and Ramaker (1980) solves this problem by executing only the first move in this calculated sequence. Then at the next sampling time, a new value of $\underline{E}'$ is calculated using feedback from the process to correct the predictions and a new L-step control sequence generated. Therefore only the first element of $\underline{\Delta u}$ is ever implemented and the actual control law can be expressed as:

$$\Delta u_k = \underline{K}^T\underline{E}' \qquad (11)$$

where $\underline{K}^T$ is the first row of the pseudo-inverse matrix in eq. 10. This method of implementation results in a forward shift in the optimization horizon (R sampling periods in the future) and the control horizon (L sampling periods in the future) at each sampling time.

The extension of this predictive controller to multivariable systems is straightforward and presented in Marchetti (1982) and Marchetti et al. (1985). It involves partitioning the error, prediction, and input vectors and adding additional blocks to the Dynamic Matrix.

A STABILITY ANALYSIS FOR PREDICTIVE CONTROL

Since predictive control was originally developed by industrial groups, their primary concern was to obtain good system performance. Therefore, complete theoretical analyses of these techniques have not yet appeared. Marchetti (1981) derived a controller transfer function for a predictive controller. For the special case where R=N-1, Marchetti (1982) developed a closed-loop characteristic equation for use in pole assignment. This equation assumed an exact process model.

In this section, we develop the closed-loop transfer function and characteristic equation for any predictive controller of the form in eq. 11 and implemented as described above regardless of how the gain vector, $\underline{K}^T$, has been calculated. Different convolution models can be used for the process and the controller model, thus allowing for analysis to include the effects of modeling errors. The subsequent derivation assumes that the process is open-loop stable and can therefore be represented by a finite sampled impulse response. The details of the derivation are available elsewhere (Maurath, 1985).

Using a classical feedback structure, the controller transfer function can be written as:

$$D(z) = \frac{\Delta u(z)}{e(z)} = \frac{B}{1 + \sum_{s=1}^{N-1} z^{-s} \sum_{V=s+1}^{s+R} \sum_{t=v-s}^{R} K_t \hat{h}_v} \tag{12}$$

$$B = \sum_{r=1}^{R} K_r \tag{13}$$

where $e = y_{sp} - y$, z is the z-transform variable, K_i is the ith element of the gain vector $\underline{K}$ in eq. 11, and $\hat{h}_v$ is the v-th sampled impulse response coefficient of the process model used to design the controller. This model will be referred to as the controller model. The process transfer function is:

$$HG_p(z) = \frac{y(z)}{\Delta u(z)} = \frac{\sum_{j=1}^{N} h_j z^{-j}}{1 - z^{-1}} \tag{14}$$

where h_j is the j-th element of the process sampled impulse response. The closed-loop characteristic equation of the system is:

$$0 = 1 + D(z)HG_p(z) \tag{15}$$

or

$$0 = z^N + z^{N-1} \left(\sum_{v=2}^{R+1} \sum_{t=v-1}^{R} K_t \hat{h}_v + Bh_1 - 1 \right)$$

$$+ \sum_{s=2}^{N-1} z^{N-s} \left[\sum_{v=s+1}^{s+R} K_{v-s} \hat{h}_v + B(h_s - \hat{h}_s) \right] + B(h_N - \hat{h}_N) \tag{16}$$

The order of this equation is N. For the case of no modeling error, the order reduces to N-1. The number of terms in the model, N, in most cases will be fairly large, between 20 and 50. Since the elements of the controller gain vector are complicated functions of the controller design parameters and the process model, it is not possible to directly calculate stability limits as functions of these controller design parameters. Also, the roots of eq. 16 can be very sensitive to small changes in the polynomial coefficients. Hence, the determinations of controller stability reported in this paper have been verified via closed-loop simulations as well as the solution of eq. 16. The above development does allow direct determination of whether a particular closed-loop system is stable and what degree of plant/model mismatch can be tolerated by the controller. This evaluation can be made by using two different models for the controller model coefficients $\{\hat{h}_j\}$ and the true process model $\{h_j\}$.

Cutler (1983) has stated that a stability analysis is not required for DMC because it is so robust. However, the stability analysis presented above is a valuable design tool for determining the degree of robustness of the controller and the tradeoff between robustness and closed-loop system performance in the design of the controller. These considerations affect the choice of controller design parameters for predictive controllers. Examination of the robustness characteristics of predictive controllers has played a key role in the development of guidelines for controller design parameter selection which will be presented in the next section.

DESIGN PARAMETERS FOR SISO PREDICTIVE CONTROLLERS

In considering the general effects that the design parameters have on the closed-loop system performance and robustness we will consider five parameters; L (control horizon), R (optimization horizon), f (move suppresion parameter), N (dimension of the convolution model), and T_s (controller sampling period).

General Considerations

As L is increased, more degrees of freedom are available for controller optimization. This often (but not always) translates into tighter control of the process. The impact of changing L on the closed-loop system is not always significant because only the first element of the "optimal" input vector $\underline{\Delta u}$ is ever implemented. Increasing L may result in better control system performance but at the expense of larger changes in the manipulated variable and a reduction in the controller's robustness. This assertion and other results mentioned here will be illustrated by simulation examples in the next section.

Increasing the optimization horizon R generally has a stabilizing effect on the closed-loop system. Increasing R increases the number of equations which must be satisfied in a least-squares sense by the optimal control law of eq. 10. As R increases, the optimization horizon is extended to regions where the step response of the process normally changes more slowly. Consequently, the optimal inputs tend to be smoothed out. The final calculated input (which is never implemented unless L=1) asymptotically approaches the long-range deviation from set point divided by the process gain.

The move suppression parameter, f, is used in the objective function to weight the changes in the input. An analogous procedure, "ridge regression", has been used in least-squares estimation problems over the last twenty years by many authors (Hoerl and Kennard, 1970). Ridge regression was originally used as a method of "shrinking" parameter values in estimated models back to the origin and reducing the sensitivity of those values to noise in the data. In the predictive controller optimization, the f parameter has been used in an analogous way to reduce the size of the calculated input changes. An improvement in closed-loop robustness results from the decreased sensitivity of the input changes to small changes in the process model (analogous to noise in the data). Significantly, no single method of determining the "proper" value for the ridge parameter has ever been widely accepted (Dempster et al., 1977), and the use of the ridge parameter has even come under periodic attack as a poor statistical practice (Smith and Campbell, 1980).

The choices of N and T_s are interrelated and also related to R. The settling time of the system model in eq. 1 is NT_s. Since we have assumed that the system is self-regulating, NT_s is finite. NT_s must be large enough to reduce truncation error from an available continuous model to an acceptable level or to adequately represent available experimental data. Small values of N are desirable to reduce necessary computation, but on the other hand, values of T_s which are too large result in reduced controller flexibility and impaired disturbance handling. Typically, values of $20 \leq N \leq 50$ have been employed by a number of investigators.

Design Parameter Selection

The first three controller design parameters (L,R, and f) are the principal means by which the designer can determine the closed-loop system's performance and robustness. While their individual effects on control system performance are relatively easy to understand, their net interactions and tradeoffs are not so obvious. When L is chosen to be large (greater than about 5), the effect of R on the closed-loop system is considerably reduced. When the controller is implemented, only the first control move calculated at any sampling time is used. When L is large, the controller optimization has enough degrees of freedom so that larger changes in the input variable can be allowed because more compensating moves are available. The first move of the controller is most affected by the predicted errors early in the predicted trajectory. Changes in R (unless it is very close to L) have very little effect on the actual control moves. If L and R are too close to each other, the controller approaches a deadbeat controller, which in most cases should be avoided. If a large value of L is used, a large value of R is required, and the only parameter left to "tune" the response is the move suppression parameter, f. The choice of f then involves a tradeoff between increasing the robustness of the controller and maintaining good controller performance.

Cutler (1983) presents a method for selecting L, R, and f for the DMC controller. He advocates setting R = L+N and finding L by increasing it until changes in L have no further effect on the first move of the controller in response to a step change in setpoint. The main tuning parameter is then the input suppression parameter f as noted above.

A Special Case: L=1

There is some motivation for avoiding the iterative procedure proposed by Cutler (1983) or at least reducing its complexity. One major motivation is that large values of L and R increase the off-line and on-line computational requirements for controller design and implementation. What about small values of L? When L is reduced to 1 or 2, R has a strong effect on the system response. For example, when L=1 and f=0, variation of R can generate a controller which ranges from a deadbeat controller (R = minimum possible value) to a steady-state controller (R very large). By varying R, the designer can easily trade speed of response for robustness. For L=1, the controller gain matrix K in eq. 10 reduces to a gain vector of this form:

$$\underline{K}^T = \frac{1}{\sum_{i=1}^{R} a_i^2} [a_1, a_2, a_3, \ldots, a_R] \qquad (17)$$

This controller is simple and easily tuned by changing R. Its design is much more straightforward than using larger values of L and move suppression to avoid robustness problems. The controller does not require move suppression for tuning. It can be easily adjusted to give responses comparable to those obtained by using larger values of L and move suppression; it also has excellent robustness characteristics. These properties will be illustrated in the next section.

Recommendations for Design Parameters

Since no significant benefits can be obtained by using more complicated controllers, we recommend that L=1 be chosen for predictive control of unconstrained single-input/single-output systems. Selection of controller design parameters for MIMO systems is a considerably more difficult task than for SISO systems. A new method for designing MIMO predictive controllers has also been developed by Maurath et al. (1985).

What is a good starting value for R in the L=1 controller? After extensive simulation experience with several different types of systems, including non-minimum phase systems, we recommend that R be set equal to the number of sampling periods required for the process open-loop step response to reach 50% of its final value. For non-minimum phase systems, there is an additional requirement. Assuming that the process gain is positive, the following condition is necessary for stability:

$$\sum_{i=1}^{R} a_i > 0 \tag{18}$$

This condition implies that the first input change in response to a step change in setpoint must be in the direction of the eventual steady-state value of the input. This restriction normally imposes a lower limit on acceptable values of R.

The L=1 controller also has some philosophical appeal. When it is implemented, each calculated input change is actually implemented. By contrast, when L > 1, the results of the optimization are not fully utilized and the relationship between the previously calculated optimal inputs and those that are eventually implemented is not obvious. The results of the optimization for L = 1 are fully utilized and updated at each control interval.

For constrained SISO problems, the L = 1 controller may be inappropriate because the controller cannot take into account future input constraints when calculating the current input change. When larger values of L are required, the design method in Maurath et al. (1985) can provide guidance on the choice of the move suppression parameter f. Maurath (1985) addresses predictive controller design for input constrained problems. More general constrained problems may require use of quadratic programming techniques (Garcia et al., 1984).

SIMULATION RESULTS

To illustrate how these recommendations for controller design parameters might be applied, several simulation examples will be considered. These examples illustrate that an L=1 controller which has performance and robustness characteristics comparable to much more complicated predictive controllers can be easily designed and implemented.

Three numerical examples are presented in Table 1. Note that System 3 has a right-half plane zero which produces an inverse response to a step change in input.

Table 1. Simulation Examples

System 1	$\dfrac{e^{-6s}}{10s + 1}$	$T_s = 2$
System 2	$\dfrac{e^{-6s}}{(5s + 1)^5}$	$T_s = 2$
System 3	$\dfrac{(1 - 9s)}{(3s + 1)(15s + 1)}$	$T_s = 1.5$

Figure 1 shows closed-loop step responses for System 1 for three different predictive controllers. Controllers 1 and 2 are L=1 controllers with R=7 and R=5, respectively. R=7 is our recommended starting value for the L=1 predictive controller for this system. Controller 3 was designed with L=4, R=30, and f=0.1. The L = 4 controller response lies between the responses for the two L = 1 controllers. The different responses for the two L=1 controllers demonstrate the effect of R on performance. The only difference between the comparable controllers is their behavior as the set point is approached. The L=1 controllers make a smooth rise without overshoot, while the L=4 controller produces a slight overshoot. All three controllers would be acceptable in many situations. One disadvantage of the faster controllers is that the input variable changes are much larger than for the slower controller. The largest value of the input is a factor of 2 larger. This example shows how the L=1 controller can be easily tuned to provide the desired response by adjusting R. The design of Controller 3 required selection of the L and R parameters; many different combinations of L and R result in controllers with similar performance characteristics.

Figures 2 and 3 illustrate the robustness characteristics of the three controllers above. These and similar figures later in this section show the allowable variations in the model used to design the controller (i.e., the "controller model") relative to the fixed process model. The location of the process model in parameter space is designated with a "+". Figure 2 shows the stability limits for the three controllers in the gain/time delay parameter space for the controller model when the process time constant is known exactly. The robustness of the slower controller is somewhat greater than the faster ones in this parameter space. Figure 3 shows the stability regions in the time constant/time delay parameter space for the controller model.

Figure 4 illustrates closed-loop step responses for System 2 using three predictive controllers. Controller 1 is an L=1 controller using R=12. Our design criterion for this system would indicate R=15 as a starting value of R. Controllers 2 and 3 were designed with L=10, R=50, and f=0.01 or 0.1, respectively. The responses of these controllers and the size of the input variable changes are quite similar, with Controller 3 having smaller input changes than Controller 2, as would be expected. As in Figure 3, the L=1 controller exhibits less oscillation in its approach to set point. The oscillation would be increased if R were reduced.

Figure 5 shows the stability regions for these three controllers in the gain/time delay plane. The stability regions are comparable in size and shape for the L=1 controller and the more complicated controllers. The shapes of the regions are somewhat similar to those for the first-order system in Figure 3.

Figure 6 presents closed-loop step responses for System 3 using three predictive controllers. Controller 1 is the recommended L=1 controller with R=15. Because System 3 is a non-minimum phase system, the speed of response of the L=1 controller goes through a maximum as a function of R. Controller 2 shows this maximum at R=20. The largest inverse response generated with the L=1 controller also occurs at this maximum speed of response. The amount of inverse response that the open-loop system exhibits limits the speed of response of the closed-loop system with the L=1 controller. The size of the inverse response in the closed-loop step response is also limited with the L=1 controller. Controller 3 was designed with L=4, R=30, and f=0.5. For smaller values of f, Controller 3 would exhibit a larger inverse response and a marginally faster rise to the new set point. Such a large inverse response (more than 50% of the set-point change) may not be acceptable in many situations.

In a previous paper, Marchetti et al. (1983) presented an application of a predictive controller to a stirred-tank heating system. They chose a controller designed with L=3, R=5, and f=0.1, with the goal being to keep L/R small. Figure 7 shows simulated step responses for an L=1 predictive controller with R=10 and Marchetti's controller. This example illustrates the advancement in our understanding of predictive control and the effectiveness of the predictive controller.

CONCLUSIONS

The development of a predictive controller has been reviewed for a single-input, single-output process. A stability analysis for this controller has been developed which allows the computation of closed-loop system stability in the presence of modeling errors. The effects of controller design parameters on the closed-loop system response and robustness have been examined for several representative systems. These studies have led to recommendations for the controller design parameter which are based on the process dynamics. The use of a control horizon of one sampling period (L=1) produces an effective controller which can be tuned using the optimization horizon R as the principal tuning parameter. Control systems designed using these recommendations were evaluated for four systems and were found to be comparable to more complicated predictive controllers with regard to both performance and robustness.

ACKNOWLEDGEMENT

We wish to express our appreciation to the National Science Foundation for its support of the first author through the National Science Foundation Graduate Fellowship program.

LITERATURE CITED

Cutler, C.R., Ph.D. Thesis, University of Houston, Houston, 1983.

Cutler, C.R., Ramaker, B.L., Joint Automatic Control Conference Preprints, Paper WP5-B, San Francisco, 1980.

Dempster, A.P., Schatzoff, M., Wermuth, N., J. Amer. Statist. Assoc., 1977, 72, 77.

Garcia, C.E., Morshedi, A.M., Fitzpatrick, T.J., Proc. Amer. Control Conf., Paper TA4-10:00, San Diego (1984).

Hoerl, A.E., Kennard, R.W., Technometrics 1970, 12, 55.

Marchetti, J.L., M.Sc. Thesis, University of California, Santa Barbara, 1981.

Marchetti, J.L., Ph.D. Thesis, University of California, Santa Barbara, 1982.

Marchetti, J.L., Mellichamp, D.A., Seborg, D.E., Ind. Eng. Chem. Proc. Des. Dev., 1983, 22, 488.

Maurath, P.R., Seborg, D.E. and Mellichamp, D.A., Proc. Amer. Control Conf., Boston, June, 1985.

Maurath, P.R., Ph.D. Thesis, University of California, Santa Barbara, 1985.

Richalet, J., Rault, A., Testud, J.L., Papon, J. Automatica, 1978, 14, 413.

Smith, G., Campbell, F., J. Amer. Statist. Assoc. 1980, 75, 74.

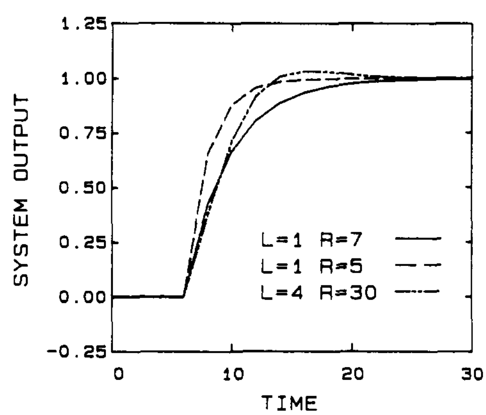

Fig 1 System 1
SP Changes

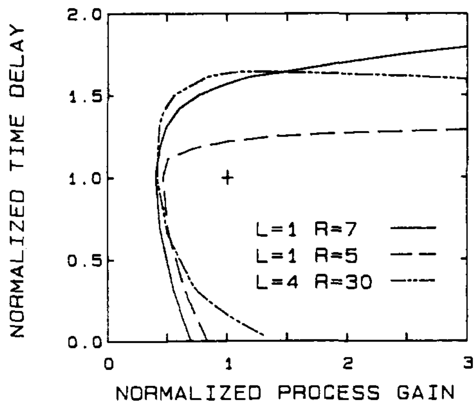

Fig 2 System 1
Stability Regions

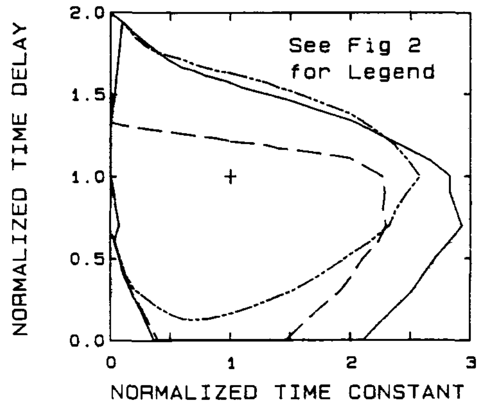

Fig 3 System 1
Stability Regions

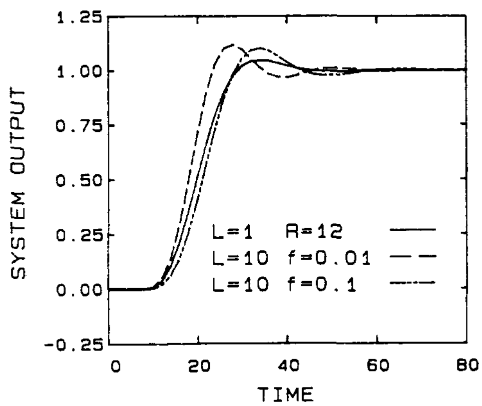

Fig 4 System 2
SP Changes

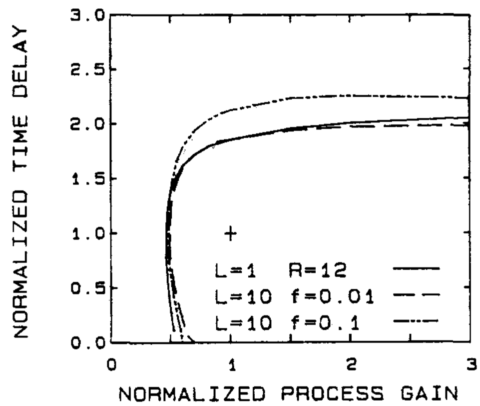

Fig 5 System 2
Stability Regions

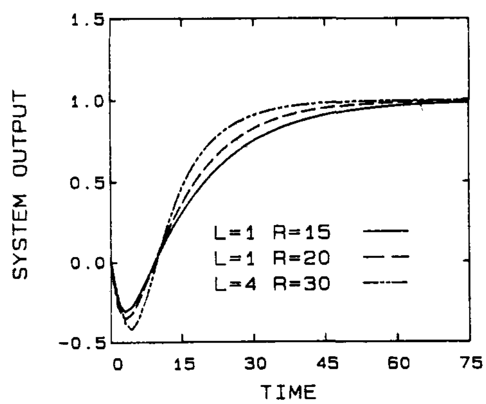

Fig 6 System 3
SP Changes

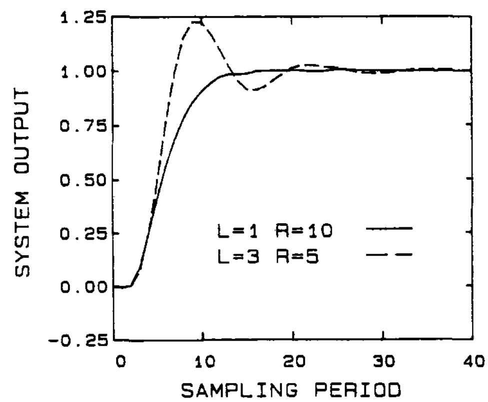

Fig 7 Stirred-Tank Heater
SP Changes