Center for Quality and Productivity Improvement
UNIVERSITY OF WISCONSIN
610 Walnut Street
Madison, Wisconsin 53705
(608) 263-2520
(608) 263-1425 FAX

Report No. 111

**Discrete Proportional-Integral Control
with Constrained Adjustment**

George E.P. Box and Alberto Luceño

February 1994

The Center for Quality and Productivity Improvement cares about your reactions to our reports. Please direct comments (general or specific) to: Report Editor, Center for Quality and Productivity Improvement, 610 Walnut Street, Madison, WI 53705; (608) 263-2520. All comments will be forwarded to the author(s).

Discrete Proportional-Integral Control with Constrained Adjustments

George Box

Center for Quality
and Productivity Improvement

*University of Wisconsin
Madison, Wisconsin*

Alberto Luceño

Applied Statistics Department

*University of Cantabria
Santander, Spain*

ABSTRACT

It is well known that discrete feedback control schemes chosen to produce minimum mean square error at the output can require excessive manipulation of the compensating variable. Also very large reductions in the manipulation variance can be obtained at the expense of minor increases in the output variance by using constrained schemes. Unfortunately, however, both the form and the derivation of such schemes are somewhat complicated. The purpose of this article is to show that suitable "tuned" proportional-integral (PI) schemes in which the required adjustment is merely a linear combination of the two last observed errors can do almost as well as the more complicated optimal constrained schemes. If desired, these PI schemes can be applied manually using a feedback adjustment chart which is no more difficult to use than a Shewhart chart. Several examples are given and tables are provided that allow the calculation of the optimal constrained proportional-integral scheme and the resulting adjustment variance and output variance. Methods of tuning such controllers using Evolutionary Operation and experimental design are briefly discussed.

KEYWORDS: *Constrained adjustment; Exponentially weighted moving average (EWMA); Feedback adjustment chart; Proportional-integral (PI) control; Time series.*

This work was supported in part by the Alfred P. Sloan Foundation and the Spanish DGICYT under Grant No. PB92-0502.

Copyright © 1994 by George E.P. Box and Alberto Luceño.

Discrete Proportional-Integral Control with Constrained Adjustment

George E.P. Box and Alberto Luceño

It is well known that discrete feedback control schemes chosen to produce minimum mean square error at the output can require excessive manipulation of the compensating variable. Also very large reductions in the manipulation variance can be obtained at the expense of minor increases in the output variance by using constrained schemes. Unfortunately, however, both the form and the derivation of such schemes are somewhat complicated. The purpose of this article is to show that suitable "tuned" proportional-integral (PI) schemes in which the required adjustment is merely a linear combination of the two last observed errors can do almost as well as the more complicated optimal constrained schemes. If desired, these PI schemes can be applied manually using a feedback adjustment chart which is no more difficult to use than a Shewhart chart. Several examples are given and tables are provided that allow the calculation of the optimal constrained proportional-integral scheme and the resulting adjustment variance and output variance. Methods of tuning such controllers using Evolutionary Operation and experimental design are briefly discussed.

1. INTRODUCTION

Suppose for some operating process, observations of an output quality characteristic are made at equally spaced intervals of time. The length of such intervals would depend on the rate of response of the system; for a waste water treatment plant it might be a day, for a chemical process an hour, and for an atomic reactor a second. Suppose that in each such interval an adjustment to an input compensating variable X can be made in an attempt to bring the output as close as possible to some target value.

A simple scheme of feedback control* to achieve such compensation is shown in Figure 1 where, for illustration, we consider the problem of controlling a dyeing process by manipulating the input dye addition rate X so as to maintain the output color index of the dyed material as close as possible to the desired target value. Continual adjustment of this kind may be needed, for example, because of drifts in the ambient temperature or in the quality of raw materials or in other "noise" variables which can

neither be eliminated nor allowed for in advance. The net effect of such noise on the output of the process when no compensatory action is taken will be called the disturbance y_t . For simplicity we will suppose that y_t is measured as a deviation from a fixed output target value T . Also we suppose that at time zero the adjustment variable X was set so that the output was on target. Then, ideally, at time t , the net effect Y_t at the output of all the previous compensatory adjustments $X_{t-1} - X_{t-2}$, $X_{t-2} - X_{t-3}$, would just cancel the difference y_t (so that, $Y_t = -y_t$). In practice, this would not be possible to do exactly and an error (deviation from target) of $e_t = y_t + Y_t$ would occur. These errors $e_t, e_{t-1}, e_{t-2}, \dots$ may then be "fed back" to decide the next appropriate adjustment $X_t - X_{t-1}$. Thus in the dyeing example, the level X_t at time t of the adjustment (input) variable is the Dye Addition Rate and this is adjusted in the light of current and past deviations $e_t, e_{t-1}, \dots$ from target of the (output) color index. This paper is about the choice of the control equation

$$X_t = f(e_t, e_{t-1}, \dots).$$

It will be noted from the above that the values $y_t, y_{t-1}, \dots$ of the disturbance represent deviations from target that would occur if no adjustments were applied (i.e., if $X_t = X_{t-1} = X_{t-2} = \dots = X_0$). The errors $e_t, e_{t-1}, \dots$ represent such deviations when some kind of feedback control, involving manipulation of X_t , is

* More detailed discussions of discrete feedback control in relation to time series and dynamic models will be found in either one of the first two editions of Box & Jenkins (1970, 1976) or in the third edition Box, Jenkins, & Reinsel (1994) to which frequent references will be made. For brevity, we denote such references by (BJ,R). Other basic references are Astrom (1970), Astrom & Wittenmark (1984), and MacGregor (1972, 1987).

Figure 1. A feedback control loop illustrated for a dyeing operation. Y_t is the net effect at the output of all past adjustments in X_t and y_t is the net effect at the output of disturbances in the system.

operating.

Perhaps the most intuitive form of feedback control is such that will make the adjustment $x_i = X_i - X_{i-1}$ of the input variable at time i proportional to e_i , the last error at the output, so that

$$x_i = c_1 e_i \quad (1.1)$$

where c_1 is a constant whose choice is considered later. By summing equation (1.1) from 1 to t , we obtain a control equation of the form

$$X_t = k_0 + k_f \sum_{i=1}^t e_i \quad (1.2)$$

where $k_f = c_1$ and $k_0 = X_0$ is a constant defining the initial settings of X . Thus the control produced by equation (1.1) is such that the level of the adjustment variable X_t depends linearly on the cumulative sum of past errors. The control equation (1.2) is the discrete analog of the control engineer's continuous *integral control* in which the adjustment depends linearly on the integral of past errors.

A somewhat surprising aspect of the simple form of control defined by (1.1) and (1.2) is its relation to exponential smoothing. Suppose that g is the *steady state gain* of the system; that is to say that a unit adjustment in the input eventually will produce g units of change at the output. Suppose also that this change becomes fully effective in one time interval

so that $Y_t = gX_{t-1}$. Then (as shown below) the control equation (1.2) can alternatively be written

$$X_t = -\frac{1}{g} \hat{y}_{t+1}, \quad (1.3)$$

where, with $\tau = 1 + c_1 g$,

$$\hat{y}_{t+1} = (1 - \tau)(y_t + \tau y_{t-1} + \tau^2 y_{t-2} + \dots) \quad 0 < \tau < 1. \quad (1.4)$$

Thus the control equation (1.2) adjusts X_t so as to just cancel an *exponentially weighted moving average* (EWMA) of present and past values $y_t, y_{t-1}, \dots$ of the disturbance.

To see this, from Figure 1 and equation (1.1),

$$X_t - X_{t-1} = c_1 e_t = c_1 (y_t + gX_{t-1})$$

and so defining τ as above,

$$X_t - \tau X_{t-1} = -\frac{1 - \tau}{g} y_t, \quad (1.5)$$

which is a first order difference equation with the solution (1.3).

Now exponentially weighted moving averages are often used empirically to estimate (forecast) a future value in a non-stationary time series. The control equation is thus equivalent to continually adjusting X_t so as to cancel out the forecast $\hat{y}_{t+1}$, made at time t , of the level the disturbance would

have at time $t+1$.

A slight elaboration of (1.1) would make the adjustment depend linearly on the last two errors:

$$x_t = c_1 e_t + c_2 e_{t-1}. \quad (1.6)$$

Summing as before we get

$$X_t = k_0 + k_p e_t + k_I \sum_{i=1}^t e_i \quad (1.7)$$

where $k_0 = X_0 + c_2 e_0$, $k_p = -c_2$, and $k_I = c_1 + c_2$. Thus the control produced by equation (1.7) is such that X_t is a linear combination of the last error e_t and the cumulative sum of previous errors $\sum_{i=1}^t e_i$.

The simple adjustment equation (1.6) thus produces control which is a discrete analog of continuous *proportional-integral* (PI)* control widely used in the process industries. These continuous controllers and their corresponding discrete counterparts have, when appropriately tuned, been remarkably successful. See for example Mayr (1970), Fearn & Maris (1991), and Box (1991). They are also easy to apply. Indeed, (Box, et. al., 1994), the discrete controllers can be put into operation manually using *feedback adjustment charts* that are no more difficult to use than Shewhart monitoring charts.

2. FEEDBACK ADJUSTMENT CHARTS

Feedback control is often thought of as exclusively a part of "automatic" process control or "computer" control. However, this form of control is also of value to the quality practitioner when manual adjustment of the process is required. Before resorting to feedback control it is important, of course, to first eliminate all assignable causes of variation with the help of process *monitoring* charts such as Shewhart, Cusum, and Roberts' EWMA charts (see, for example, Box & Kramer, 1992). Such charts are not designed for process adjustment, however, and when this is needed manual feedback control charts should be used.

INTEGRAL ACTION

For illustration consider again the dyeing process referred to earlier in which, by manipulating dye

* Automatic "PID" controllers include a third "differential" term which takes account of the rate of change of the error. If in the corresponding discrete analog the derivative is replaced by a finite difference, the adjustment $x_t = X_t - X_{t-1}$ becomes a linear function of the last *three* errors.

addition rate, the color index of the dyed material is kept close to a target value ($T = 9$). Suppose at first that the desired discrete control was of the "integral" type of equation (1.2) with $k_I = c_1 = -5$. The control equation is then

$$X_t = k_0 - 5 \sum_{i=1}^t e_i$$

and the corresponding adjustment equation is

$$x_t = -5e_t.$$

A simple chart that the operator can use to achieve this action is shown in Figure 2. The input adjustment scale shown on the left and the output scale shown on the right are such that a unit deviation from target in the color index calls for minus five units of adjustment in the dye addition rate. The operator records on the chart the current value of the color index and immediately reads off on the left hand scale the required adjustment to the dye addition rate. Thus at time 1:30 P.M. the color index was 9.14 calling for reduction of -0.7 units in the dye addition rate.

PROPORTIONAL-INTEGRAL ACTION

If instead, proportional-integral action was required (say with $k_p = -1.25$ and $k_I = -5$), then to achieve the required action

$$X_t = k_0 - 1.25e_t - 5 \sum_{i=1}^t e_i.$$

The chart of Figure 3 would be employed with dotted lines placed at a fraction $k_p/k_I = 0.25$ of each interval. The operator would then read off the desired control action at each stage by drawing a line through the last two points and extrapolating 0.25 intervals ahead (as indicated by the arrows). The adjustment would be obtained by reading off the extrapolated value on the adjustment scale. Thus in the figure the last two readings at 1:15 P.M. and 1:30 P.M. were 9.06 and 9.14. The projected value of 9.16 requires reduction of the dye addition rate of -0.8 units. Note that the use of these charts requires no calculation, i.e., they are entirely graphical. They are also very robust so that in practice "eyeballing" the extrapolation would be good enough.

If ∇ is the back difference operator so that $\nabla e_t = e_t - e_{t-1}$, then the adjustment equation (1.6) can be written

$$x_t = -G(1 + P\nabla)e_t$$

where $G = -k_I$ and $P = k_p/k_I$. In general then, in terms

Figure 2. A feedback control chart to achieve Integral action.

Figure 3. A feedback control chart to achieve Proportional-Integral (PI) action.

of the manual adjustment chart, the meaning of G and P is as follows:

The constant G determines the relative size of the units on the scales of x_t and e_t .

The constant P determines the required degree of interpolation or extrapolation of the last two points on the chart. If P is positive, the required action is obtained by *extrapolating* P units along the line joining e_t to e_{t-1} . If P is negative, a corresponding *interpolation* between e_t and e_{t-1} is called for.

3. SIMPLE TIME SERIES AND DYNAMIC MODELS AND MINIMUM MEAN SQUARE ERROR CONTROL

So far we have discussed discrete PI control from an empirical point of view. No models or theoretical justification has been so far introduced. This is deliberate for it is important to understand that this form of control can be (and historically has been) developed from a purely common sense standpoint. With this understanding it is hardly surprising that PI control is useful over a much wider field than that covered, for example, by the minimum mean square error (MMSE) justification briefly discussed below.

Processes sometimes contain inertia such that some time elapses before a change at the process input becomes fully effective at the output. Useful discrete models to represent such dynamic behavior are provided by linear difference equations. Correspondingly, models to describe the disturbance y_t are provided by stochastic linear difference equations (sometimes called ARIMA models) for which the input is a "white noise" sequence $\{a_t\}$ of independent and identically distributed random variables having mean zero and variance $\text{Var}(a_t) = \sigma_a^2$. Given such models, it is easy to derive the appropriate MMSE feedback control equation which minimizes the output mean square error about the target (BJR). This general class of models and the resulting control equations are extremely broad in scope. Practical situations can, however, often be adequately modeled by simple special cases.

A SIMPLE DISTURBANCE MODEL

It will be recalled that the disturbance model must represent the output y_t of the system when no effort of any kind is made to control it. It is argued in Box and Kramer (1992) that the entropy of such a disturbance

should increase monotonically and that a simple but realistic model is one for which the variogram of the output $V_m = \text{Var}(y_{t+m} - y_t)$ increases linearly with m . This property is enjoyed by the integrated moving average (IMA) model in which the first difference of the disturbance follows a first order moving average time series

$$y_t - y_{t-1} = a_t - \theta a_{t-1}. \quad (3.1)$$

From an empirical standpoint, identification of actually occurring disturbances has often led to this model or some close approximation to it. As is well known, the MMSE forecast at time t for all future time of a series generated by (3.1) is given by $\hat{y}_{t+1}^{(\theta)}$ the EWMA defined by equation (1.4) with $\tau = \theta$. If the system is such that corrective action has its full effect in one time interval so that $Y_t = gX_{t-1}$, it is easy to show that, for a disturbance represented by (3.1), "integral" control action defined by equations (1.1) and (1.2) yields minimum mean square error at the output if the constant k_I is chosen so that

$$k_I = c_1 = -\lambda/g, \quad 0 < \lambda < 1,$$

where $\lambda = 1 - \theta$.

More generally the process may have inertial characteristics such that a change at time t would not be realized at the output during one interval of time. It is frequently possible to represent such dynamic behavior by a "linear filter". Specifically, the output from the system is modeled by $g\tilde{X}_t$, where $\tilde{X}_t$ is a weighted average of previous levels of the input $X_{t-1}, X_{t-2}, \dots$. A simple dynamic model which provides such a filter is such that the change $\tilde{X}_t - \tilde{X}_{t-1}$, which occurs at the (scaled) output in the time interval between t and $t-1$, is proportional to the deviation $X_{t-1} - \tilde{X}_{t-1}$ between the input and the (scaled) output at time $t-1$ so that

$$\tilde{X}_t - \tilde{X}_{t-1} = (1 - \delta)(X_{t-1} - \tilde{X}_{t-1}), \quad 0 < \delta < 1,$$

where the constant $1 - \delta$ can be interpreted as the proportion of the eventual output change that will occur in one time interval after a "step" change has been made in the input. Thus the close δ is to unity the slower the system will react. Equivalently,

$$\tilde{X}_t - \delta \tilde{X}_{t-1} = (1 - \delta)X_{t-1}. \quad (3.2)$$

This first order difference equation provides a discrete analog of a first order dynamic system typified, for example, by Newton's law of cooling. If the system dynamic can be represented in this way,

then it is easy to show (see BJ,R) that proportional-integral action defined by equations (1.6) and (1.7) will yield a minimum mean square error if the constants k_I and k_P are chosen such that

$$k_I = -\lambda/g, k_P = -(\lambda\delta)/(g(1-\delta)). \quad (3.3)$$

The solution of (3.2) (cf. 1.5) is

$$\tilde{X}_t = \hat{X}_t^{(\delta)} = (1-\delta)(X_{t-1} + \delta X_{t-2} + \delta^2 X_{t-3} + \dots).$$

Thus the linear filter associated with the difference equation (3.2) is an exponentially weighted moving average with smoothing constant δ . It is then easy to show that the PI control equation with constants given by (3.3) is equivalent to continuously arranging that

$$\hat{X}_t^{(\delta)} = -\frac{1}{g} \hat{y}_t^{(\theta)}.$$

That is, an EWMA with smoothing constant δ of past values of the input variable is chosen to cancel an EWMA with smoothing constant θ of the disturbance.

Sometimes the dynamic model must allow, in addition, for "dead time" that is pure delay of, say, f units in response to adjustment. The equation for a first order system with f units of delay (see equation 3.2) is

$$\tilde{X}_t - \delta \tilde{X}_{t-1} = (1-\delta)X_{t-1-f}. \quad (3.4)$$

In that case MMSE is not given by PI control but by a scheme for which the adjustment equation is of the form

$$x_t = d_1 x_{t-1} + d_2 x_{t-2} + \dots + d_f x_{t-1-f} + d(e_t - \delta e_{t-1}),$$

where the d 's are suitable coefficients.

From the above it will be seen that one rationalization for the use of proportional-integral control might be that, for the case where $f = 0$, it produces minimum mean square error at the output for perhaps the simplest models for disturbance and for dynamics which can approximate reality. However, from equation (3.3) it will be seen that the coefficients G and P of the feedback chart that then yields MMSE must be chosen such that

$$G = \lambda/g, P = \delta/(1-\delta).$$

We see, therefore, that with $0 < \delta < 1$ the chart will

always require extrapolation rather than interpolation. Furthermore, a chart thus constructed for a slow acting system in which δ is close to unity can require very extensive extrapolation and, consequently, instability. Even with δ as small as 0.5, we find that $P = 1$ and one whole unit of extrapolation is called for.

4. CONSTRAINED ADJUSTMENT SCHEMES

Thus while, at first sight, the criterion that the control scheme should produce MMSE at the output seems eminently sensible, unfortunately, such MMSE control can require excessive manipulation of the compensating variable X_t . Various schemes have been devised, therefore, in which the range of compensatory manipulation is constrained (Box and Jenkins, 1970; Wilson, 1970; MacGregor, 1972). In particular, such constrained schemes can be obtained by finding an unconstrained minimum of the expression

$$\text{Var}(x_t) + \alpha \text{Var}(e_t) \quad (4.1)$$

where α is a constant which allocates the relative quadratic costs of the departures of e_t and x_t from zero. Options may be explored by making calculations for various choices of α .

MAJOR REDUCTIONS IN THE MANIPULATION VARIANCE $\text{VAR}(X_T)$ POSSIBLE WITH CONSTRAINED SCHEMES

A characteristic of constrained schemes obtained in this way is that, frequently, extremely large reductions in the manipulation variance $\text{Var}(x_t)$ can be obtained at the expense of very minor increases in the output variance $\text{Var}(e_t)$. To illustrate this point consider the following example. Suppose the disturbance model is of the form of equation (3.1) with $\lambda = 0.4$ and $\sigma_a^2 = 1.0$, also that the dynamics can be represented by a first order system with $g = 0.4$ and $\delta = 0.5$ and with dead time characterized by $f = 1$.

The minimum mean square error scheme is

$$x_t = -0.40x_{t-1} - 2.00e_t + 1.00e_{t-1},$$

giving variances $\text{Var}(e_t) = 1.16$ and $\text{Var}(x_t) = 5.00$.

Two examples of optimal constrained schemes obtained by minimizing (4.1) for different values of α are

$$(A) x_t = 0.29x_{t-1} - 0.11x_{t-2} - 0.98e_t + 0.49e_{t-1},$$

giving $\text{Var}(e_t) = 1.21$ and $\text{Var}(x_t) = 1.00$,

$$(B) x_t = 0.75x_{t-1} - 0.21x_{t-2} - 0.39e_t + 0.19e_{t-1},$$

giving $\text{Var}(e_t) = 1.36$ and $\text{Var}(x_t) = 0.25$.

Thus for scheme (A), at the expense of an increase of about 4% in the output variance, an 80% reduction in the input variance is possible. Scheme (B) produces an increase of 17% in the output variance and a reduction of 95% in the adjustment variance.

SUBSTITUTION OF PI SCHEMES FOR OPTIMAL CONSTRAINED SCHEMES

Although optimal constrained schemes such as those illustrated above can produce dramatic reductions in the adjustment variance $\text{Var}(x_t)$ at the expense of only minor increases in the output variance $\text{Var}(e_t)$, they can be somewhat complicated in form and in their derivation. The main purpose of this article is to show that for a simple and useful class of models simple PI schemes, when suitably "tuned", can do almost as well. Thus in what follows, we shall assume that we are to use a PI adjustment equation

$$-x_t = G(1 + PV)e_t,$$

or correspondingly a control equation

$$-X_t = -k_0 + GPe_t + G \sum_{i=1}^t e_i$$

and we will consider the appropriate choices of G and P for various constrained schemes.

That this approach is useful is illustrated by considering again the example in the last sub-section. Constrained PI schemes matching (A) and (B) in the sense that they produce the same reduction in $\text{Var}(x_t)$ are

$$(A') x_t = -0.61(1 + 0.62V)e_t,$$

giving $\text{Var}(e_t) = 1.21$ and $\text{Var}(x_t) = 1.00$,

$$(B') x_t = -0.45(1 - 0.11V)e_t,$$

giving $\text{Var}(e_t) = 1.37$ and $\text{Var}(x_t) = 0.25$.

Since (A) and (B) are *optimal* constrained solutions, (A') and (B') are necessarily inferior, but we see that to two decimal accuracy the PI schemes do almost as well as the much more complicated constrained schemes. In Appendix B we further demonstrate that PI schemes doing almost as well as optimal constrained schemes can be found for a wide range of situations.

Note that for the more "heavily damped" scheme (B') the constant $P = -0.11$ is negative. This calls for a linear *interpolation* at a point 89% of the way between e_{t-1} and e_t , rather than for an "extrapolation", which accounts intuitively for the greater stability of

the scheme and the smaller variance $\text{Var}(x_t)$. Study of the tables, given later, shows that this behavior is typical. Notice also that the values $P = +0.62$ and $P = -0.11$ for schemes (A') and (B') imply that an intermediate scheme exists for which $P = 0$. This is

$$(C) x_t = -0.49e_t, \text{ giving } \text{Var}(e_t) = 1.33 \text{ and } \text{Var}(x_t) = 0.32.$$

We see that this extremely simple scheme requiring only integral action produces a 94% reduction in the input variance for an increase of only 14% in the output variance.

5. CONSTRAINED PROPORTIONAL- INTEGRAL SCHEMES

Now although minimum mean square error schemes based on the simple time series and dynamic models (3.1) and (3.4) are often unsatisfactory, this is principally the fault of the inappropriate choice of the *criterion* (output MMSE). As we have already mentioned, there is both theoretical and empirical evidence that supports the use of the IMA time series model (3.1) to approximate the behavior of an uncontrolled disturbance. Furthermore, there is much empirical evidence that the first order system, possibly with delay as represented by (3.4), frequently provides a useful approximation to the dynamic behavior of industrial processes. In practice, the sampling interval for a process will be chosen taking into account the rapidity with which the system reacts to adjustment so that schemes for $f = 0$ and $f = 1$ are of the most practical importance. In what follows we shall, therefore, retain these *models* to derive optimally constrained PI schemes.

The problem of finding the values of G and P for a PI scheme that, for a given value of $\text{Var}(e_t)$, minimizes $\text{Var}(x_t)$, is equivalent to the problem of finding the unconstrained minimum of the function

$$L_\alpha = \{g(1 - \delta)\}^2 \text{Var}(x_t) + \alpha \text{Var}(e_t). \quad (5.1)$$

The procedure we have used for doing this is given in Appendix A where it is shown that the resulting values of gG , P , $\text{Var}(e_t)/\sigma_a^2$, and $\text{Var}(gx_t)/\sigma_a^2$ are functions of λ , δ , f , and α . The results are given in Tables 1 and 2 for frequently encountered ranges of the parameters. Methods for estimating λ , δ , f , and g from actual operating data are given in time series texts and, in particular, in BJ,R (but, see also Section 7).

Table 1.

The table shows the values of gG , P , $\text{Var}(e_t)/\sigma_e^2$, and $\text{Var}(g_t)/\sigma_g^2$ for the optimal PI scheme when $f=0$ and $\delta=0.0$ (0.1) 0.5, for several values of α and λ .

$\lambda =$	$\delta = 0.0$						$\delta = 0.1$						$\delta = 0.2$					
	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000
$gG =$	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000
$P =$	0.000	0.000	0.000	0.000	0.000	0.000	0.111	0.111	0.111	0.111	0.111	0.111	0.250	0.250	0.250	0.250	0.250	0.250
$Var(e_t)/\sigma_e^2 =$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$Var(gx_t)/\sigma_g^2 =$	0.040	0.090	0.160	0.250	0.360	1.000	0.050	0.112	0.200	0.312	0.449	1.247	0.065	0.146	0.260	0.406	0.585	1.625
	0.181	0.261	0.336	0.406	0.473	0.698	0.181	0.267	0.344	0.417	0.485	0.715	0.187	0.272	0.352	0.427	0.498	0.733
	-0.221	-0.194	-0.167	-0.140	-0.113	0.000	-0.161	-0.133	-0.104	-0.076	-0.047	0.078	-0.080	-0.050	-0.020	-0.010	-0.041	0.177
	1.004	1.009	1.016	1.025	1.036	1.100	1.004	1.009	1.016	1.025	1.036	1.100	1.004	1.009	1.016	1.025	1.036	1.100
	0.022	0.049	0.087	0.135	0.194	0.537	0.025	0.057	0.101	0.157	0.226	0.628	0.030	0.068	0.122	0.190	0.275	0.763
	0.170	0.240	0.305	0.364	0.418	0.592	0.174	0.247	0.314	0.375	0.431	0.608	0.178	0.254	0.324	0.387	0.446	0.628
	-0.284	-0.252	-0.219	-0.185	-0.151	0.000	-0.244	-0.209	-0.173	-0.136	-0.099	0.066	-0.187	-0.149	-0.110	-0.070	-0.029	0.150
	1.007	1.017	1.031	1.050	1.072	1.200	1.007	1.017	1.032	1.050	1.072	1.200	1.008	1.018	1.032	1.050	1.072	1.200
	0.018	0.039	0.069	0.106	0.152	0.420	0.020	0.044	0.077	0.120	0.173	0.480	0.023	0.051	0.091	0.142	0.204	0.568
	0.151	0.209	0.261	0.306	0.347	0.465	0.156	0.216	0.269	0.317	0.358	0.478	0.161	0.224	0.279	0.329	0.372	0.494
	-0.347	-0.312	-0.275	-0.235	-0.193	0.000	-0.326	-0.287	-0.246	-0.203	-0.157	0.052	-0.295	-0.252	-0.207	-0.159	-0.108	0.119
	1.015	1.035	1.063	1.099	1.144	1.400	1.015	1.035	1.063	1.099	1.144	1.400	1.015	1.035	1.063	1.100	1.144	1.400
	0.014	0.029	0.050	0.077	0.110	0.303	0.015	0.032	0.055	0.085	0.122	0.337	0.017	0.036	0.062	0.095	0.138	0.385
	0.137	0.187	0.230	0.267	0.299	0.388	0.141	0.193	0.237	0.276	0.309	0.398	0.146	0.200	0.246	0.286	0.320	0.411
	-0.380	-0.345	-0.307	-0.265	-0.219	0.000	-0.366	-0.329	-0.287	-0.241	-0.191	0.043	-0.347	-0.306	-0.260	-0.209	-0.154	0.099
	1.023	1.053	1.095	1.150	1.216	1.600	1.022	1.052	1.095	1.149	1.215	1.600	1.022	1.052	1.094	1.149	1.216	1.600
	0.011	0.023	0.040	0.061	0.087	0.240	0.012	0.025	0.043	0.066	0.095	0.262	0.013	0.028	0.048	0.074	0.106	0.293
	0.126	0.169	0.206	0.238	0.264	0.333	0.129	0.174	0.212	0.245	0.272	0.342	0.134	0.181	0.220	0.254	0.282	0.352
	-0.399	-0.366	-0.328	-0.285	-0.237	0.000	-0.391	-0.354	-0.313	-0.266	-0.215	0.037	-0.378	-0.338	-0.293	-0.242	-0.185	0.087
	1.031	1.072	1.128	1.200	1.288	1.800	1.031	1.071	1.127	1.199	1.288	1.800	1.030	1.070	1.126	1.199	1.287	1.800
	0.010	0.020	0.033	0.051	0.072	0.200	0.010	0.021	0.036	0.055	0.078	0.216	0.011	0.023	0.039	0.060	0.086	0.237
	0.116	0.155	0.187	0.214	0.237	0.293	0.119	0.159	0.193	0.220	0.243	0.300	0.123	0.165	0.199	0.228	0.252	0.308
	-0.413	-0.380	-0.343	-0.299	-0.250	0.000	-0.406	-0.372	-0.331	-0.285	-0.232	0.033	-0.398	-0.360	-0.316	-0.265	-0.207	0.075
	1.040	1.091	1.161	1.251	1.361	2.000	1.039	1.090	1.160	1.250	1.360	2.000	1.038	1.088	1.158	1.249	1.359	2.000
	0.008	0.017	0.029	0.044	0.062	0.172	0.009	0.018	0.030	0.047	0.066	0.183	0.009	0.019	0.033	0.050	0.072	0.199
	0.098	0.129	0.153	0.173	0.189	0.225	0.100	0.132	0.157	0.177	0.194	0.230	0.103	0.136	0.162	0.183	0.199	0.236
	-0.433	-0.403	-0.366	-0.323	-0.272	0.000	-0.429	-0.398	-0.360	-0.313	-0.259	0.025	-0.425	-0.391	-0.350	-0.301	-0.242	0.057
	1.065	1.141	1.245	1.379	1.542	2.500	1.063	1.139	1.243	1.377	1.541	2.500	1.061	1.136	1.241	1.375	1.540	2.500
	0.006	0.012	0.021	0.032	0.046	0.127	0.006	0.013	0.022	0.034	0.048	0.134	0.007	0.014	0.024	0.036	0.052	0.142

Table 1. continued

$\lambda =$	$\delta = 0.3$						$\delta = 0.4$						$\delta = 0.5$					
	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000
$gG =$	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000
$P =$	0.429	0.429	0.429	0.429	0.429	0.429	0.667	0.667	0.667	0.667	0.667	0.667	1.000	1.000	1.000	1.000	1.000	1.000
$(e_1)/\sigma_{e_1}^2 =$	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$(gx_1)/\sigma_{gx_1}^2 =$	0.089	0.200	0.356	0.556	0.801	2.224	0.129	0.290	0.516	0.806	1.160	3.222	0.200	0.450	0.800	0.250	1.800	5.000
	0.190	0.278	0.360	0.438	0.511	0.752	0.193	0.283	0.368	0.449	0.524	0.773	0.195	0.287	0.375	0.458	0.536	0.793
	0.035	0.065	0.096	0.127	0.160	0.308	0.202	0.230	0.260	0.292	0.325	0.486	0.450	0.474	0.502	0.532	0.566	0.742
	1.004	1.009	1.017	1.026	1.037	1.100	1.004	1.010	1.017	1.026	1.037	1.100	1.004	1.010	1.017	1.026	1.037	1.100
	0.039	0.088	0.156	0.245	0.352	0.974	0.054	0.122	0.217	0.338	0.485	1.325	0.085	0.189	0.332	0.513	0.730	1.952
	0.182	0.262	0.334	0.401	0.462	0.648	0.186	0.269	0.345	0.415	0.478	0.671	0.190	0.276	0.356	0.429	0.495	0.696
	-0.105	-0.064	-0.022	0.022	0.066	0.263	0.019	0.063	0.108	0.154	0.202	0.418	0.212	0.257	0.303	0.352	0.404	0.644
	1.008	1.019	1.033	1.051	1.073	1.200	1.009	1.020	1.034	1.053	1.075	1.201	1.009	1.020	1.035	1.054	1.076	1.201
	0.028	0.062	0.111	0.175	0.253	0.703	0.036	0.082	0.147	0.232	0.334	0.924	0.053	0.121	0.216	0.338	0.486	1.319
	0.166	0.233	0.291	0.343	0.388	0.514	0.172	0.243	0.305	0.359	0.407	0.537	0.178	0.253	0.319	0.377	0.427	0.563
	-0.249	-0.201	-0.150	-0.096	-0.040	0.209	-0.177	-0.122	-0.064	-0.004	0.059	0.335	-0.058	0.005	0.071	0.140	0.211	0.521
	1.015	1.036	1.064	1.101	1.146	1.401	1.017	1.038	1.067	1.104	1.148	1.402	1.018	1.040	1.070	1.108	1.153	1.404
	0.019	0.041	0.073	0.114	0.164	0.457	0.022	0.038	0.090	0.142	0.205	0.573	0.029	0.067	0.122	0.193	0.280	0.778
	0.152	0.209	0.257	0.299	0.335	0.427	0.159	0.219	0.271	0.315	0.352	0.447	0.166	0.231	0.287	0.333	0.373	0.472
	-0.319	-0.272	-0.220	-0.163	-0.102	0.174	-0.274	-0.220	-0.160	-0.095	-0.026	0.281	-0.199	-0.133	-0.062	0.014	0.094	0.440
	1.022	1.052	1.095	1.150	1.217	1.601	1.023	1.054	1.098	1.153	1.221	1.602	1.025	1.058	1.103	1.159	1.227	1.606
	0.015	0.031	0.055	0.085	0.122	0.339	0.017	0.037	0.065	0.102	0.147	0.411	0.021	0.046	0.083	0.131	0.191	0.536
	0.139	0.189	0.230	0.265	0.294	0.366	0.146	0.199	0.243	0.279	0.309	0.383	0.154	0.211	0.258	0.297	0.329	0.405
	-0.359	-0.315	-0.264	-0.207	-0.144	0.150	-0.330	-0.278	-0.219	-0.154	-0.082	0.242	-0.279	-0.217	-0.146	-0.069	0.014	0.381
	1.030	1.070	1.126	1.199	1.288	1.801	1.030	1.071	1.128	1.202	1.292	1.803	1.032	1.074	1.133	1.208	1.299	1.807
	0.012	0.025	0.044	0.067	0.097	0.269	0.014	0.029	0.051	0.079	0.114	0.318	0.016	0.036	0.063	0.098	0.143	0.401
	0.128	0.172	0.208	0.238	0.262	0.320	0.135	0.181	0.219	0.250	0.276	0.334	0.143	0.193	0.234	0.267	0.293	0.354
	-0.384	-0.343	-0.293	-0.237	-0.173	0.132	-0.364	-0.316	-0.259	-0.195	-0.123	0.213	-0.329	-0.271	-0.203	-0.127	-0.043	0.336
	1.038	1.087	1.158	1.249	1.360	2.001	1.038	1.088	1.159	1.251	1.363	2.003	1.039	1.090	1.163	1.257	1.370	2.008
	0.010	0.021	0.037	0.056	0.080	0.222	0.012	0.024	0.042	0.064	0.092	0.258	0.013	0.029	0.050	0.078	0.113	0.318
	0.107	0.141	0.168	0.189	0.206	0.243	0.113	0.148	0.176	0.198	0.216	0.253	0.113	0.148	0.176	0.198	0.216	0.253
	-0.419	-0.382	-0.337	-0.283	-0.219	0.101	-0.409	-0.368	-0.317	-0.256	-0.186	0.163	-0.409	-0.368	-0.317	-0.256	-0.186	0.163
	1.060	1.134	1.239	1.374	1.539	2.501	1.058	1.132	1.238	1.374	1.540	2.503	1.058	1.132	1.238	1.374	1.540	2.503
	0.007	0.015	0.024	0.030	0.042	0.155	0.006	0.017	0.030	0.044	0.062	0.174	0.006	0.017	0.030	0.044	0.062	0.174

Table 2.

The table shows the values of gG , P , $\text{Var}(g_x)/\sigma_g^2$, and $\text{Var}(e_i)/\sigma_e^2$ for the optimal PI scheme when $f=1$ and $\delta=0.0$ (0.1) 0.5, for several values of α and λ .
 The table also shows the input variance for the optimal unconstrained scheme $\text{Var}_U(g_x)$. The output variance for the optimal unconstrained scheme is given by $\text{Var}_U(e_i)/\sigma_e^2 = 1 + \lambda^2 f$.

$\lambda =$	$\delta = 0.00$						$\delta = 0.1$						$\delta = 0.2$					
	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000
$gG =$	0.166	0.228	0.278	0.319	0.351	0.420	0.165	0.227	0.276	0.316	0.347	0.412	0.165	0.225	0.274	0.312	0.342	0.404
$P =$	0.193	0.279	0.355	0.423	0.483	0.651	0.317	0.407	0.487	0.558	0.620	0.797	0.470	0.564	0.649	0.724	0.791	0.980
$Var(e_i)/\sigma_e^2 =$	1.040	1.090	1.162	1.256	1.374	2.137	1.040	1.091	1.163	1.258	1.378	2.158	1.040	1.091	1.164	1.260	1.382	2.180
$Var(g_x)/\sigma_g^2 =$	0.039	0.086	0.146	0.217	0.298	0.698	0.048	0.102	0.171	0.252	0.340	0.773	0.060	0.126	0.208	0.300	0.401	0.880
	0.164	0.224	0.273	0.313	0.344	0.411	0.164	0.224	0.272	0.311	0.341	0.405	0.164	0.223	0.271	0.309	0.338	0.398
	0.115	0.208	0.293	0.369	0.436	0.625	0.228	0.327	0.417	0.497	0.568	0.766	0.369	0.475	0.571	0.657	0.732	0.944
	1.040	1.091	1.163	1.257	1.375	2.139	1.040	1.091	1.163	1.259	1.379	2.160	1.040	1.092	1.164	1.261	1.383	2.182
	0.034	0.075	0.129	0.194	0.269	0.643	0.040	0.089	0.151	0.225	0.308	0.715	0.051	0.109	0.183	0.269	0.363	0.816
	0.150	0.202	0.243	0.275	0.301	0.356	0.152	0.204	0.246	0.279	0.304	0.358	0.154	0.207	0.249	0.282	0.307	0.359
	-0.176	-0.081	0.018	0.117	0.210	0.493	-0.112	-0.006	0.104	0.211	0.311	0.608	-0.029	0.093	0.215	0.332	0.440	0.754
	1.047	1.105	1.187	1.293	1.423	2.243	1.047	1.106	1.188	1.293	1.424	2.253	1.047	1.106	1.188	1.294	1.426	2.265
	0.018	0.040	0.072	0.112	0.160	0.416	0.020	0.046	0.082	0.128	0.182	0.466	0.024	0.054	0.097	0.150	0.213	0.534
	0.103	0.133	0.156	0.173	0.186	0.210	0.108	0.140	0.164	0.182	0.195	0.221	0.114	0.147	0.172	0.191	0.205	0.232
	-0.411	-0.362	-0.298	-0.219	-0.128	0.254	-0.398	-0.342	-0.271	-0.184	-0.086	0.314	-0.379	-0.315	-0.234	-0.137	-0.030	0.395
	1.084	1.190	1.340	1.534	1.770	3.147	1.080	1.183	1.329	1.517	1.745	3.076	1.078	1.178	1.321	1.504	1.726	3.016
	0.007	0.015	0.026	0.039	0.056	0.156	0.008	0.017	0.028	0.044	0.063	0.174	0.009	0.019	0.032	0.049	0.071	0.197
	0.052	0.064	0.072	0.077	0.081	0.087	0.056	0.069	0.078	0.084	0.088	0.094	0.060	0.075	0.084	0.091	0.095	0.103
	-0.467	-0.440	-0.400	-0.347	-0.279	0.094	-0.464	-0.435	-0.394	-0.338	-0.267	0.117	-0.460	-0.430	-0.385	-0.326	-0.251	0.147
	1.248	1.521	1.898	2.386	2.988	6.519	1.227	1.479	1.830	2.284	2.844	6.122	1.206	1.439	1.765	2.187	2.707	5.746
	0.002	0.005	0.008	0.013	0.018	0.050	0.003	0.005	0.009	0.014	0.020	0.056	0.003	0.006	0.010	0.016	0.023	0.063
$Var_{ext}(g_x)/\sigma_g^2 =$	0.040	0.090	0.160	0.250	0.360	1.000	0.050	0.112	0.200	0.312	0.449	1.247	0.065	0.146	0.260	0.406	0.585	1.625

Table 2. continued

$\lambda =$	$\delta = 0.03$						$\delta = 0.4$						$\delta = 0.5$					
	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000	0.200	0.300	0.400	0.500	0.600	1.000
$gG =$	0.165	0.224	0.272	0.309	0.338	0.396	0.164	0.223	0.270	0.306	0.334	0.388	0.164	0.222	0.268	0.303	0.329	0.380
$P =$	0.664	0.764	0.855	0.937	1.010	1.217	0.920	1.029	1.129	1.221	1.303	1.534	1.275	1.398	1.512	1.617	1.712	1.980
$Var(e_t)/\sigma_a^2 =$	1.040	1.091	1.165	1.262	1.387	2.203	1.040	1.092	1.166	1.265	1.392	2.227	1.040	1.092	1.167	1.268	1.397	2.252
$Var(gx_t)/\sigma_a^2 =$	0.078	0.161	0.261	0.372	0.490	1.039	0.108	0.217	0.346	0.484	0.628	1.285	0.158	0.312	0.487	0.672	0.861	1.696
	0.164	0.223	0.270	0.306	0.335	0.392	0.163	0.222	0.268	0.304	0.331	0.385	0.163	0.221	0.266	0.301	0.327	0.378
	0.550	0.665	0.768	0.861	0.944	1.174	0.791	0.917	1.031	1.135	1.227	1.483	1.128	1.270	1.399	1.517	1.623	1.918
	1.041	1.092	1.166	1.263	1.388	2.205	1.041	1.092	1.167	1.226	1.393	2.229	1.041	1.093	1.168	1.269	1.398	2.254
	0.066	0.140	0.231	0.334	0.445	0.965	0.091	0.189	0.306	0.435	0.572	1.195	0.134	0.273	0.434	0.607	0.785	1.580
	0.155	0.209	0.251	0.284	0.309	0.359	0.157	0.211	0.253	0.286	0.310	0.359	0.158	0.212	0.255	0.287	0.311	0.358
	0.085	-0.225	0.362	0.490	0.608	0.947	0.246	0.407	0.562	0.706	0.837	1.208	0.483	0.672	0.850	1.014	1.162	1.581
	1.048	1.107	1.189	1.296	1.429	2.281	1.049	1.108	1.190	1.298	1.432	2.298	1.049	1.109	1.192	1.300	1.435	2.317
	0.029	0.067	-0.119	0.185	0.260	0.633	0.038	0.088	0.157	0.240	0.333	0.785	0.056	0.128	0.223	0.335	0.459	1.038
	0.120	0.155	0.181	0.201	0.215	0.243	0.126	0.163	0.191	0.211	0.226	0.255	0.132	0.172	0.200	0.221	0.237	0.266
	-0.352	-0.277	-0.182	-0.071	0.049	0.505	-0.311	-0.218	-0.104	0.025	0.163	0.661	-0.244	-0.124	0.019	0.175	0.336	0.893
	1.076	1.175	1.316	1.495	1.712	2.968	1.076	1.176	1.315	1.493	1.706	2.933	1.078	1.180	1.320	1.496	1.705	2.912
	0.010	0.021	0.036	0.057	0.082	0.229	0.011	0.024	0.043	0.068	0.099	0.275	0.014	0.030	0.054	0.086	0.126	0.350
	0.065	0.081	0.092	0.099	0.104	0.113	0.072	0.089	0.101	0.109	0.115	0.124	0.080	0.099	0.113	0.122	0.128	0.138
	-0.455	-0.422	-0.373	-0.309	-0.228	0.191	-0.448	-0.410	-0.355	-0.283	-0.193	0.257	-0.437	-0.391	-0.326	-0.242	-0.138	0.361
	1.186	1.402	1.704	2.097	2.582	5.398	1.167	1.368	1.651	2.019	2.471	5.085	1.152	1.340	1.608	1.956	2.383	4.822
$Var(gx_t)/\sigma_a^2 =$	0.003	0.007	0.012	0.018	0.026	0.072	0.004	0.008	0.014	0.021	0.030	0.084	0.005	0.010	0.017	0.025	0.037	0.102
	0.089	0.200	0.356	0.556	0.801	2.224	0.129	0.290	0.516	0.806	1.160	3.222	0.200	0.450	0.800	1.250	1.800	5.000

6. DESCRIPTION AND USE OF THE TABLES

In Tables 1 and 2 the values of gG , P , $\text{Var}(e_t)/\sigma_a^2$, and $\text{Var}(gx_t)/\sigma_a^2$ for a variety of alternative schemes have been tabulated by using different values of α . The tables may be used to choose a suitable constrained PI scheme by consideration of the resulting input and output variances. Three cases have been considered. Table 1 with $\delta = 0.0$ is used for a system where the response to an adjustment occurs in one time interval. Table 1 for $\delta > 0.0$ is appropriate where the dynamics may be approximated by a first order dynamic system. Table 2 is for the case where the system inertia may be represented by a first order dynamic system with one period of delay. For intermediate values of the parameters linear interpolation is sufficient to find approximate values. The rationale for the choice of the α values is discussed in Appendix B.

EXAMPLE

As an example of the use of the tables we consider a recent application presented in this Journal (Fearn and Maris 1991) of feedback adjustment for the control of gluten additions in a flour mill.

In this example extensive preliminary runs had provided data yielding estimates of $\lambda = 0.25$ and $\sigma_a = 0.12$. Also the dynamic response of the system was modeled by a simple delay of one time unit and the units had been chosen so that $g = 1$. Thus $f = 1$ and the appropriate minimum mean square error scheme which these authors employed was

$$-x_t = \lambda(x_{t-1} + e_t/g) \text{ or } -x_t = 0.25(x_{t-1} + e_t).$$

Fearn and Maris found this unconstrained scheme to be entirely satisfactory. However, the results of the present paper show that constrained PI schemes might have been used which for rather small increases in $\text{Var}(e_t)$ could give dramatic reductions in $\text{Var}(x_t)$. Linear interpolation in the first two columns of Table 2 yielded the following specimen schemes:

G	P	% increase in $\text{Var}(e_t)$	% decrease in $\text{Var}(x_t)$
0.19	0.16	0.1	16
0.18	-0.13	1.0	55
0.12	-0.39	6.7	83

We notice further that an even simpler scheme

$$-x_t = 0.19e_t$$

obtained by interpolation to make P zero would give an increase in $\text{Var}(e_t)$ of only 0.6% with a decrease in $\text{Var}(x_t)$ of 38%.

7. TUNING PI CONTROLLERS WITH EVOLUTIONARY OPERATION AND EXPERIMENTAL DESIGN

The ease with which estimates of the coefficients λ , g , δ , and f can be obtained varies with different applications. Thus, for some systems, it will be known that an adjustment at the input is fully realized at the output in one time interval and, therefore, that δ and f are both zero. Also the steady state gain constant g is often known from physical understanding of the process. Finally, when past records are available, λ and other needed constants can be estimated by direct calculation (see, for example, BJ,R).

Sometimes, however, especially at the initiation of an entirely new scheme, little information is available. In such a case, the PI scheme based on best guesses can be run initially and improved upon using evolutionary operation (EVOP) (Box, 1957; Box & Draper, 1968). The "variables" studied in the EVOP scheme will be the two constants G and P of the PI controller. Suppose G_0 and P_0 are the initial values obtained from best guesses of the parameters. A 2^2 factorial design can then be run in the evolutionary operation mode with G_0 and P_0 as the center point and levels $G_0 \pm d_G$ and $P_0 \pm d_P$. If the control scheme is run manually, then five different manual charts might be used in random sequence on successive days. The "responses" will be the estimates s_e^2 and s_x^2 of $\text{Var}(e_t)$ and $\text{Var}(x_t)$ obtained at the five points of the design (the 2^2 factorial plus a center point), and in practice, it is preferable to work with $\log(s_e^2)$ and $\log(s_x^2)$. (See for example, Bartlett and Kendall, 1946.) Experience suggests that initial guesses for λ might be 0.2 or 0.3. Also, for a system exhibiting dynamic characteristics, a guess for δ might be made by remembering that for a first order system a proportion of $1-\delta$ of the eventual change will occur in the initial time interval.

When, as sometimes happens, there is a number m of PI control schemes operating simultaneously on the same process, multi-factor designs can be used in which $G_1, P_1; G_2, P_2; \dots; G_m, P_m$ are the design

variables. Initially fractional factorial designs or other orthogonal arrays can be used to determine a direction of steepest ascent. If necessary, this can be followed later by a response surface study in some or all of the control variables. This approach has the advantage that it can take account of interaction between the different controllers.

8. CONCLUSION

Minimum mean square error feedback control schemes can require excessive manipulation of the adjustment variable. Constrained PI schemes may be used to significantly reduce the variance of the adjustments required at the expense of a small increase in the variance of the output. The performance of simple optimal constrained PI schemes appears to be almost as good as that of general linear constrained schemes for a wide range of practical situations. The former, however, are much easier to implement, since the required adjustment is simply a linear combination of the two last observed errors.

Tables have been provided that allow the calculation of these adjustments corresponding to optimal constrained PI schemes and the resulting variances for the system's input and output. Methods of tuning controllers using Evolutionary Operation and experimental design are briefly discussed.

ACKNOWLEDGMENTS

This research was supported in part by the Alfred P. Sloan Foundation and the Spanish DGICYT under Grant No. PB92-0502.

APPENDIX A: DERIVATION OF CONSTRAINED PI SCHEMES

As we saw previously, the PI adjustment equation $x_t = -G(1 + PV)e_t$ implies that x_t depends only on the last two errors e_t and e_{t-1} according to the linear relation

$$x_t = c_1 e_t + c_2 e_{t-1} \quad (\text{A.1})$$

with $G = -(c_1 + c_2)$ and $P = -c_2/(c_1 + c_2)$. Now let B be the backward shift operator so that, for example, $Be_t = e_{t-1}$. Then the IMA time series model (3.1) may be written

$$(1 - B)y_t = (1 - \theta B)a_t \quad (\text{A.2})$$

with parameters $\lambda = 1 - \theta$ and σ_a^2 . Also the delayed first order dynamic model with f periods of delay may be written

$$(1 - \delta B)\tilde{X}_t = (1 - \delta)X_{t-1-f} \quad (\text{A.3})$$

Now, introducing (A.1), (A.2), and (A.3) into the equation resulting from multiplying $e_t = y_t + g\tilde{X}_t$ by $(1 - B)$, one gets that e_t follows an ARMA($p, 2$) model of the form $(1 - \phi_1 B - \dots - \phi_p B^p)e_t = [1 - (\delta + \theta)B + \delta\theta B^2]a_t$, where the order p of the autoregressive part of the model is $f+2$.

For $f=0$, the autoregressive parameters are

$$\phi_1 = 1 + \delta + c_1 g(1 - \delta), \text{ and}$$

$$\phi_2 = -\delta + c_2 g(1 - \delta).$$

Similarly, for $f=1$, one has

$$\phi_1 = 1 + \delta,$$

$$\phi_2 = -\delta + c_1 g(1 - \delta), \text{ and}$$

$$\phi_3 = c_2 g(1 - \delta).$$

Using this ARMA($p, 2$) model for e_t , it is possible to calculate $\text{Var}(e_t)$ and $\text{Cov}(e_t, e_{t-1})$ as functions of the parameters $c_1, c_2, \lambda, \sigma_a^2, \delta, f$, and g . Using (A.1), we can now calculate $\text{Var}(x_t) = (c_1^2 + c_2^2)\text{Var}(e_t) + 2c_1 c_2 \text{Cov}(e_t, e_{t-1})$ as a function of the same set of parameters

The minimization of L_α (equation 5.1) with respect to c_1 and c_2 can be performed numerically for the set of values of c_1 and c_2 that makes the ARMA($p, 2$) model for e_t stationary. Because g enters L_α only as a factor affecting c_1 and c_2 , and σ_a^2 does not affect the point where the minimum of L_α is attained, the values of gc_1 and gc_2 that minimize L_α and provide the optimal constrained PI scheme are functions of λ, δ, f , and α . Then $gG, P, \text{Var}(e_t)/\sigma_a^2$, and $\text{Var}(gx_t)/\sigma_a^2$ are also functions of λ, δ, f , and α .

APPENDIX B: FURTHER COMPARISON OF PI SCHEMES WITH OPTIMAL LINEAR SCHEMES

Tables B1 and B2 further illustrate the similarity and the results obtained from suitable tuned PI schemes and optimal linear constrained schemes. In what follows we use the subscripts C and U to indicate the

optimal linear constrained scheme and the unconstrained scheme respectively.

When $f = 0$ it is possible to show that

$$\text{Var}_C(e_t) = \sigma_a^2(1 + \lambda^2 Q), \text{Var}_C(x_t) = \text{Var}_U(x_t)W,$$

and

$$\frac{\text{Var}_U(x_t)}{\sigma_a^2} = \frac{\lambda^2(1 + \delta^2)}{g^2(1 - \delta)^2}, \quad (\text{B.1})$$

and Q and W only depend on δ and α (see Box and Jenkins, 1970, chapter 13). Because of this, we have used the same representation for the optimal

Table B1.

The table shows the values of $100Q$ and $100W$ that provide the output and input variances for some optimal constrained PI schemes and general linear constrained schemes according to the formulas $\text{Var}(e_t) = \sigma_a^2(1 + \lambda^2 Q)$ and $\text{Var}(gx_t) = \text{Var}_U(gx_t)W$ for $f = 0$, $\delta = 0.5$, and several values of λ and α . The percent increase in $\text{Var}(e_t)$ and the percent decrease in $\text{Var}(gx_t)$ with respect to the optimal unconstrained scheme are $100\lambda^2 Q$ and $100(1 - W)$, respectively.

$f = 0, \delta = 0.5$						
		OPTIMAL CONSTRAINED PI SCHEME				OPTIMAL CONSTRAINED SCHEME FOR ANY λ
	$\lambda =$	0.20	0.40	0.60	0.80	
$\alpha = 1.0742900$	100Q =	23.504	22.084	21.111	20.480	20.000
	100W =	26.660	27.040	26.989	26.706	26.179
$\alpha = 0.4153830$	100Q =	45.557	44.035	42.468	41.210	40.000
	100W =	14.663	15.220	15.542	15.616	15.489
$\alpha = 0.2152021$	100Q =	63.522	64.103	63.043	61.676	60.000
	100W =	10.319	10.389	10.600	10.706	10.694
$\alpha = 0.1288560$	100Q =	79.904	83.133	83.057	81.902	80.000
	100W =	8.130	7.846	7.926	8.006	8.025
$\alpha = 0.0844903$	100Q =	96.511	101.857	102.770	101.963	100.000
	100W =	6.744	6.281	6.278	6.329	6.354

Table B2.

The table shows the percent increase in $\text{Var}(e_t)$ and the percent decrease in $\text{Var}(gx_t)$ for some optimal constrained PI schemes, with respect to the optimal unconstrained scheme, for $f = 1$, $\delta = 0.5$, $\lambda = 0.4$, and several values of α . It also shows the percent increase in $\text{Var}(e_t)$ for the general linear constrained schemes that have the same percent decrease in $\text{Var}(gx_t)$ with respect to the optimal unconstrained scheme as the corresponding constrained PI schemes.

$f = 1, \delta = 0.5, \lambda = 0.4$			
	OPTIMAL CONSTRAINED PI SCHEME		OPTIMAL CONSTRAINED SCHEME
α	% increase in $\text{Var}(e_t) / \sigma_a^2$	% decrease in $\text{Var}(gx_t) / \sigma_a^2$	% increase in $\text{Var}(e_t) / \sigma_a^2$ for the same decrease in $\text{Var}(gx_t) / \sigma_a^2$
∞	0.64	39.08	0.38
100	0.64	39.82	0.40
1	2.72	72.13	2.50
0.01	38.6	97.93	38.32
0.0001	356.1	99.80	350.41

constrained PI schemes, i.e. $\text{Var}(e_t) = \sigma_a^2(1 + \lambda^2 Q)$ and $\text{Var}(x_t) = \text{Var}_U(x_t)W$, although here Q and W depend on δ , α , and also on λ . The increase in $\text{Var}(e_t)$ with respect to the corresponding optimal unconstrained schemes is then $100\lambda^2 Q\%$ and the decrease in $\text{Var}(x_t)$ is $100(1 - W)\%$.

Table B1 is for $f = 0$ with $\delta = 0.5$ and various values of λ . It compares values for $100Q$ and $100W$ for schemes of the two types and shows that optimal PI schemes give results that are always close to those of optimal linear schemes in terms of the decrease in $\text{Var}(g x_t)$ and increase in $\text{Var}(e_t)$ and can be used without any significant loss of efficiency.

A similar conclusion is reached from Table B2. For $f = 1$, $\delta = 0.5$, $\lambda = 0.4$, and several values of α , a comparison is made of the decrease in $\text{Var}(g x_t)$ and corresponding increase in $\text{Var}(e_t)$ that may be attained using optimal linear constrained schemes and optimal constrained PI schemes. In this case, $\text{Var}_U(e_t)/\sigma_a^2 = 1 + \lambda^2 f = 1.16$, and $\text{Var}_U(g x_t)/\sigma_a^2 = 0.8$ because of (B.1). We note that minimizing L_α/α when $\alpha = \infty$ is equivalent to minimizing $\text{Var}(e_t)$ without paying any attention to $\text{Var}(x_t)$ and that the PI scheme for $\alpha = \infty$ produces a small increase in $\text{Var}(e_t)$ in compensation for a much bigger decrease in $\text{Var}(g x_t)$.

REFERENCES

- Astrom, K.J. (1970). "Introduction to Stochastic Control," *Mathematics in Science and Engineering Series*, 70. Academic Press.
- Astrom, K.J. and B. Wittenmark (1984). *Computer Controlled Systems: Theory and Design*. Prentice Hall; New Jersey.
- Bartlett, M.S. and D.G. Kendall (1946). "The Statistical Analysis of Variance-Heterogeneity and the Logarithmic Transformation," *Journal of the Royal Statistical Society*, Ser. B, 8. pp. 128-150.
- Box, G.E.P. (1957). "Evolutionary Operation, a Method for Increasing Industrial Productivity," *Applied Statistics*, 6. pp. 81-101.
- Box, G.E.P. (1991). "Feedback Control by Manual Adjustment," *Quality Engineering*, Vol. 4. pp. 143-151.
- Box, G.E.P. and N.R. Draper (1968). *Evolutionary Operation*. John Wiley; New York.
- Box, G.E.P. and G.M. Jenkins (1970, 1976). *Time Series Analysis: Forecasting and Control*, 1st and 2nd eds. Holden-Day; San Francisco.
- Box, G.E.P., G.M. Jenkins and G. Reinsel (1994). *Time Series Analysis: Forecasting and Control*, 3rd. ed. Prentice Hall; Englewood Cliffs, NJ.
- Box, G.E.P. and T. Kramer (1992). "Statistical Process Monitoring and Feedback Adjustment. A Discussion," *Technometrics*, Vol. 34. pp. 251-285.
- Fearn, T. and P.I. Maris (1991). "An Application of Box-Jenkins Methodology to the Control of Gluten Addition in a Flour Mill," *Applied Statistics*, Vol. 40. pp. 477-484.
- MacGregor, J.F. (1972). "Topics in the Control of Linear Processes Subject to Stochastic Disturbances," Ph.D. Thesis, University of Wisconsin-Madison.
- MacGregor, J.F. (1987). "Interfaces Between Process Control and On-line Statistical Process Control," *Computing and Systems Technology Divisions Communications*, Vol. 10, No. 2. pp. 9-20.
- Mayr, O. (1970). *The Origins of Feedback Control*. M.I.T. Press; Cambridge, MA.
- Wilson G.T. (1970). "Modeling Linear Systems for Multivariate Control," Ph.D. Thesis, Department of System Engineering, University of Lancaster, England.