Centerfor Quality and Productivity Improvement
University of Wisconsin
610 Walnut Street
Madison, Wisconsin 53705

(608) 263-2520
(608) 263-1425 FAX
quality@engr.wise.edu

Report 180

Feedforward as a Supplement to Feedback Adjustment in Allowing for Feedstock Changes

George E. P. Box
Center for Quality and Productivity Improvement, University of Wisconsin Madison

Alberto Luceño
University of Cantabria, 39005 Santander, Spain

Feedforward as a Supplement to Feedback Adjustment in Allowing for Feedstock Changes

George E. P. Box

Center for Quality and Productivity Improvement, University of Wisconsin—Madison

Alberto Luceño

University of Cantabria, 39005 Santander, Spain

Many industrial processes must be adjusted from time to time to continuously maintain their mean close to target. Compensations for deviations of the process mean from target may be accomplished by feedback and/or by feedforward adjustment. Feedback adjustments are made in reaction to errors at the output; feedforward adjustments are made to compensate anticipated changes. This article considers the complementary use of feedback and feedforward adjustments to compensate for anticipated step changes in the process mean as may be necessary in a manufacturing process each time a new batch of feedstock material is introduced. We consider and compare five alternative control schemes: (1) feedforward adjustment alone, (2) feedback adjustment alone, (3) feedback-feedforward adjustment, (4) feedback and indirect feedforward to increase the sensitivity of the feedback scheme, and (5) feedback with both direct and indirect feedforward.

KEY WORDS: Feedback control; Feedforward control; First-order dynamics; Nonstationary disturbance; Repeated adjustment scheme; Statistical process control.

1. INTRODUCTION

Traditionally methods for statistical process control (SPC) have had as one important objective the detection, and where possible, the assignment and removal of disturbances due to "special causes" (e.g., see Deming 1986). Methods of this kind have been highly successful because such continued debugging can produce steady improvement in the system by reduction of variation and simplification of operation. However, it is impossible for all process disturbances, even when detectable, to be removed economically in this way. Indeed, as has been argued elsewhere (Box and Kramer 1992), the hypothesis of an *unadjusted* process remaining in a perfect state of statistical control contradicts the second law of thermodynamics and must be regarded therefore as a purely theoretical concept. This is borne out by careful study of the performance of a number of operating quality control schemes. For example, Alwan and Roberts (1988) found that process instability was common *after* standard techniques of quality control had been applied. Also it is noteworthy that in the highly successful Six Sigma system for process improvement (e.g., see Smith 1992, Harry 1994, Hoerl 1998, Box and Luceño 2000) additional allowance is made for a one and a half sigma *drift* of the local mean about the target value. It would be expected then that complementary methods based on process adjustment would be of value as an adjunct to the debugging process mentioned above, and recently there has been a resurgence of interest in the use of adjustment techniques suitable for application in the SPC context such as was discussed for example by Box and Jenkins (1970); MacGregor (1972); Fearn and Maris (1991); Vander Wiel, Tucker, Faltin, and Doganaksoy (1992); Jensen and Vardeman (1993); Tucker, Faltin, and Vander Wiel (1993); Box, Jenkins, and Reinsel (1994); Luceño, González, and Puig-Pey (1996); Box and Luceño (1997); Montgomery and Woodall (1997); and Luceño (2000).

Feedback adjustment occurs when compensatory changes are made in some suitable adjustment variable X in *reaction* to output deviations from target. Feedforward adjustment occurs when such changes are made to compensate *anticipated* deviations. We consider here the situation where feedforward control might be used to compensate for expected level shifts. A common example is when adjustments are made to compensate for batch to batch differences based on pretest of the incoming raw material. However, such feedforward adjustment if used alone could prove inadequate because pretesting is subject to measurement errors. Thus even if the background disturbance was stationary, feedforward control used alone could result in systematic deviations from the process target. In this article therefore we consider to what extent is it helpful in the SPC context to use feedforward adjustment of one kind or another in *addition* to feedback adjustment.

1.1 A Simple System of Feedback Adjustment

We suppose that the output quality characteristic we desire to control is measured at equally spaced intervals of time. Thus e_t is the *output error* observed at time t for the system under control and $x_t = X_t - X_{t-1}$ is the consequent input *adjustment* made at time t. Then in what follows we consider a system of feedback control, sufficiently simple to be readily usable in the SPC context, that consists of repeatedly making an adjustment just sufficient to cancel a fixed *proportion* G of the current output deviation from target.

The resulting feedback adjustment equation is then

$$g\, x_t = -G\, e_t, \tag{1.1}$$

where $0 < G \leq 1$ will be called the *damping factor*. The constant g is the system gain, that is the eventual change in the output that is induced by a unit adjustment at the input. Appropriate values for G are often in the range .1 to .4.

Although this type of feedback adjustment is of great simplicity, it has interesting characteristics, labeled below as A, B, and C, that are discussed more fully and justified in Box and Luceño (1997):

(A) The adjustment equation (1.1) is the discrete analog of continuous integral control because, by summing Equation (1.1), the overall correction in a period t intervals long is

$$g\, X_t = g\, X_0 + g \textstyle\sum_{i=1}^{t} x_i = g\, X_0 - G \sum_{i=1}^{t} e_i = k_0 + k_1 \sum_{i=1}^{t} e_i\, , \tag{1.2}$$

where $k_0 = g\, X_0$ and $k_1 = -G$ are constants.

(B) It further follows from Equation (1.1) that such control is equivalent to arranging that the overall correction is

$$g\, X_t = -\tilde{y}_t, \tag{1.3}$$

where

$$\tilde{y}_t = G\left[y_t + H \;\; y_{t-1} + H^2 \;\; y_{t-2} + \ldots\right]$$

is an exponentially weighted moving average (EWMA) of the output deviations from target y_t, y_{t-1}, ... *that would have occurred if no control had been applied.* The constant $H = 1 - G$, in such an EWMA, is often called the *smoothing constant* or *discount factor*. Notice that Equations (1.2) and

(1.3) make no specific model assumptions about y_t, y_{t-1}, ... but follow from (1.1) using $e_t = y_t + g X_{t-1}$. Equation (1.3) does show however that the adjustment equation (1.1) is intuitively reasonable since it is equivalent to employing $\tilde{y}_t$ as a *forecast* made at time t for y_{t+1} and arranging that the total adjustment X_t just cancels this forecasted deviation. The effect of doing this is to replace the deviation y_t that would occur at the output if there were no control, by its error of forecast $y_{t+1} - \tilde{y}_t = e_{t+1}$.

(C) The use of the adjustment equation (1.1) may be formally justified using the following assumptions:

(i) The uncontrolled output y_t may be represented by the integrated moving average (IMA) time series model

$$y_t - y_{t-1} = a_t - \theta\ a_{t-1}, \qquad (1.4)$$

where $\{a_t\}$ is a "white noise" sequence; that is, a series of independent and identically distributed (iid) random variables having mean 0 and standard deviation σ_a.

(ii) The system is responsive, that is to say that the dynamics are such that the adjustment x_t produces its full effect at the output in one time period.

With these assumptions, the output errors will have zero mean and, if G is set equal to $1 - \theta \equiv \lambda$ (so that $H = \theta$), the adjustment scheme will produce minimum mean squared error (MMSE) at the output and so will minimize the output variance var (e_t). In that case, the variance of the adjustments will satisfy g^2 var $(x_t) = G^2$ var (e_t) with var $(e_t) = \sigma_a^2$.

In practice, it is frequently advantageous to employ a value for G less than λ (i.e., $H > \theta$) for which the needed adjustments will be smaller in magnitude than those yielding MMSE (e.g., see Box and Luceño 1995). This yields a constrained adjustment scheme which minimizes var (e_t) subject to a given reduction in g^2 var (x_t) = G^2var (e_t). Such constrained schemes are attractive since large reductions in var (x_t) can frequently be obtained at the cost of only small increases in var (e_t). For example, suppose that the system gain $g = 1$ and $\theta = .6$ in Equation (1.4) so that $\lambda = 1 - \theta = .4$ and the MMSE scheme is $x_t = -.4\ e_t$, but if instead of setting $G = .4$ we used the value

$G = .2$ ($H = .8$), then it is easy to show that the standard deviation of x_t is reduced by 52.7% at the expense of an increase in the standard deviation of e_t of only 5.4%.

In favor of the assumption that the uncontrolled series may, to an acceptable approximation, be represented by the IMA of equation (1.4) are the following:

(i) formal time series analysis of production data has frequently shown this to be an adequate model,

(ii) with this model the variance of the difference between observations m steps apart is a linear function of m so that the variogram for this time series model increases linearly with m (see Box and Luceño 1997, chap. 12), and thus can be explained by the intuitively reasonable concept of "sticky innovations" discussed by Box and Kramer (1992),

(iii) Adjustment using equation (1.3) appears to be extremely robust, that is to say, it still can provide effective control in conditions widely differing from the ideal. Specifically, Box and Luceño (1997) demonstrate such robustness when contrary to assumption: there is inertia in the system response approximated by first-order dynamics, and/or there are one or two intervals of pure delay in the system response, or the distribution of errors deviates in likely ways from normality, or the correct time series model is an autoregressive process rather than an IMA process (see also Luceño 1998).

1.2 Allowance for Expected Step Changes

In this paper we consider the commonly occurring problem of how to compensate for expected step changes in the output that typically occurs in a manufacturing process whenever new batches of feedstock material are introduced. Five methods of adjustment for compensating such step changes are:

Scheme 1: Feedforward adjustment alone.

Scheme 2: Feedback adjustment alone.

Scheme 3: Feedback-Feedforward adjustment.

Scheme 4: Feedback and indirect feedforward to increase the sensitivity of feedback.

Scheme 5: Feedback with both direct and indirect feedforward.

1.3 Transient Produced by Feedback Control

As a preliminary it is of interest to determine the time that would elapse before the feedback adjustment equation (1.1) would reduce a step change to a negligible value.

For discrete data, a unit step change made at time 0 may be represented by a series consisting of "zeros" for $t < 0$ followed by "ones" for $t \geq 0$. Adjustment of such a series by Equation (1.1) produces for $t \geq 0$ the geometric series 1, H, H^2, ...

This corresponds in continuous time to the exponential decay of the impulse-response function $w(t) = (g/\tau)\exp(-t/\tau)$ of a first-order linear dynamic system, where the constant τ is called the time constant of the system. This is the time at which the effect of a step change is reduced to a proportion $e^{-1} = .368$ (i.e., 36.8%) of its initial value. The smoothing constant of the first-order system sampled at equispaced unit times satisfies $H^{\tau} = e^{-1}$ so that $\tau = -1 / \ln H$. In particular, 3τ is a convenient benchmark because this is the approximate time at which the response would decay to about 5% of its initial value. For example, for $G = .4$, $\tau = 1.96$ and $3\tau = 5.87$. The closest integer is 6 and gives $H^6 = .047$ (i.e., 4.7%), which is close to the desired value of 5%.

For the control schemes described below, we will assume that at time t the deviation from target y_t that would occur if there were no interventions can be represented by the IMA time series model (1.4). We suppose that a batch of feedstock lasts for T intervals and that an event occurring in the i th interval of the h th batch is indexed by $t = i + hT$. Interest centers on the size of the output errors e_t (measured as deviations from target after control) and the corresponding adjustments x_t that are needed, which are measured as deviations from zero. We consider the mean squared deviations of each of these quantities. When the target values and the means coincide, then mean squared deviations are the variances var (e_t) and var (x_t). By setting $\sigma_a^2 = 1$, *all* variances can be represented as multiples of the noise variance σ_a^2.

Section 2 compares the characteristics of the five schemes of Section 1.2. In Sections 3 and 4 we consider the optimal choice of the damping factor(s). Concluding remarks are given in Section 5.

2. CHARACTERISTICS OF THE SUGGESTED SCHEMES

For the uncontrolled process over infinite time, the variance of y_t in (1.4) is infinite. However, for a finite period of length T, the average value is

$$M_u = T^{-1}\sum_{i=1}^{T}\operatorname{var}(y_i) = 1 + (T-1)\lambda^2/2.$$

2.1 Scheme 1: Feedforward Alone (FF)

We will denote by μ_t the step change in the t th batch and we suppose that, from batch to batch, μ_t has mean 0 and variance σ_μ^2. We also assume that the step changes μ_t can be estimated by pretest with measurement errors ε_t having mean 0 and variance σ_ε^2. We denote by $m_t = \mu_t + \varepsilon_t$ the observed value of μ_t and by σ_m^2 the variance of m_t.

Then the total disturbance z_t resulting from the combination of Equation (1.4), for the background disturbance y_t, and the step change μ_t is

$$z_t - z_{t-1} = a_t - \theta a_{t-1} + \mu_t, \tag{2.1}$$

where $\mu_t = 0$ for any $t \neq 0, T, 2T, 3T, \ldots$. The uncontrolled process would now have variances

$$M_0 = T^{-1}\sum_{i=1}^{T}\operatorname{var}(z_i) = 1 + (T-1)\lambda^2/2 + \sigma_\mu^2$$

and the effect of the feedforward control will be to reduce this to

$$M_1 = 1 + (T-1)\lambda^2/2 + \sigma_\varepsilon^2.$$

Because only one adjustment ($g x_{t-1} = -m_t$) to compensate for m_t is made for each batch of length T, the required adjustments satisfy $V_1 = T^{-1}\sum_{i=1}^{T}\operatorname{var}(g x_i) = \sigma_m^2/T$.

2.2 Scheme 2: Feedback Alone (FB)

Suppose now that the process is controlled using only the feedback adjustment equation (1.1) with $G \le \lambda$. If there are no changes in mean from batch to batch, the errors at the output would satisfy $e_t - H\, e_{t-1} = a_t - \theta a_{t-1}$ (e.g., see Box and Luceño 1997, sec. 12.11) so that var (e_t) would be given by

$$\sigma_e^2 = \frac{1 - 2\theta H + \theta^2}{1 - H^2} \tag{2.2}$$

and $\text{var}(g x_t) = G^2 \sigma_e^2$. In particular, if $G = \lambda$ (or, equivalently, $H = \theta$), then var (e_t) = $\sigma_a^2 = 1$, the smallest value obtainable.

Suppose now that the current batch contributes a deviation in mean μ_t lasting for T intervals and that T is large compared with the time constant of the system (so that $H^{2T} \cong 0$). Then the total disturbance z_t will be given by (2.1), and the error at the output e_t at time t will be the sum of the total disturbance z_t and the accumulated effects of the compensations $g\, X_{t-1}$ performed up to time $t - 1$, that is,

$$e_t = z_t + g X_{t-1}. \tag{2.3}$$

Combining Equations (1.1), (2.1), and (2.3), one obtains $e_t - H\, e_{t-1} = a_t - \theta a_{t-1} + \mu_t$, so that $E\,(e_t) = 0$ and the long-run mean squared error at the output (MSEO) is given by

$$M_2 = \frac{1}{T} \Sigma_{i=1}^{T} \text{var}(e_i) = \sigma_e^2 + \frac{\sigma_\mu^2}{T\left(1 - H^2\right)},$$

where σ_e^2 is given by (2.2). To obtain these errors at the output, adjustments according to (1.1) must be made at the input so that the long-run input variance is then $V_2 = T^{-1} \Sigma_{i=1}^{T} \text{var}(g x_i) = G^2 M_2$.

2.3 Scheme 3: Feedback Plus Feedforward (FB+FF)

If we now use feedforward to compensate the batch mean m_t observed at time $t - 1$ with a measurement error having variance σ_ε^2, the adjustment equation (1.1) should be replaced by $g x_{t-1} = -G e_{t-1} - m_t$, where $m_t = 0$ for $t \neq 0, T, 2T, 3T, \ldots$. Consequently, the errors at the output follow the recursive relation $e_t - H\, e_{t-1} = a_t - \theta a_{t-1} + \varepsilon_t$, the expected error is null, and the MSEO is given by

$$M_3 = \sigma_e^2 + \frac{\sigma_\varepsilon^2}{T\left(1 - H^2\right)}.$$

The long-run variance at the input is $V_3 = G^2 M_3 + \sigma_m^2 / T$.

2.4 Scheme 4: Feedforward to Increase the Sensitivity of the Feedback System (FB+FFS)

Suppose now that when a new batch is to be initiated there is no direct feedforward. Instead, the information concerning the *time* when a change will occur is fedforward to temporally increase the sensitivity of the feedback system. Specifically, the sensitivity of the feedback control is temporally increased by using a larger damping factor $\tilde{G}$ in the feedback adjustment equation during the first k intervals for each new batch, and $k << T$ is chosen so that after k intervals the effect of the transient has decayed to a negligible value. Then the adjustment equations for the batch starting at time $t = T$ are $g x_t = -G e_t$ (for $t < T$), $g x_t = -\tilde{G} e_t$, (for $t = T, \ldots, T + k - 1$), and $g x_t = -G e_t$, (for $t = T + k, \ldots, 2T - 1$). The errors at the output will follow the recursive relations

$$e_t - H\, e_{t-1} = a_t - \theta a_{t-1}, \quad (t < T) \tag{2.4a}$$

$$e_T - H\, e_{T-1} = a_T - \theta a_{T-1} + \mu_T, \tag{2.4b}$$

$$e_t - \tilde{H}\, e_{t-1} = a_t - \theta a_{t-1}, \quad (t = T+1, \ldots, T+k) \tag{2.4c}$$

$$e_t - H\, e_{t-1} = a_t - \theta a_{t-1}, \quad (t = T+k+1, \ldots, 2T-1) \tag{2.4d}$$

$$e_{2T} - H\, e_{2T-1} = a_{2T} - \theta a_{2T-1} + \mu_{2T}, \tag{2.4e}$$

and so on, where $\tilde{H} = 1 - \tilde{G}$. Again $E(e_t) = 0$. Assuming that T is large enough so that $H^{2T} \cong 0$, $\tilde{H}^{2T} \cong 0$, and $k << T$, the MSEO and input variances are given by

$$M_4 = \frac{k}{T}\tilde{\sigma}_e^2 + \frac{T-k}{T}\sigma_e^2 + \frac{\sigma_\mu^2}{T}\left(\frac{\tilde{H}^{2k}}{1-H^2} + \frac{1-\tilde{H}^{2k}}{1-\tilde{H}^2}\right) + \frac{\xi}{T}\left(\frac{1}{1-H^2} - \frac{1}{1-\tilde{H}^2}\right)$$

and

$$V_4 = \frac{k}{T}\tilde{\sigma}_e^2\tilde{G}^2 + \frac{T-k}{T}\sigma_e^2 G^2 + \frac{\sigma_\mu^2}{T}\left(\frac{G\tilde{H}^{2k}}{2-G} + \frac{\tilde{G}\left(1-\tilde{H}^{2k}\right)}{2-\tilde{G}}\right) + \frac{\xi}{T}\left(\frac{G}{2-G} - \frac{\tilde{G}}{2-\tilde{G}}\right),$$

respectively, where σ_e^2 is given by (2.2),

$$\tilde{\sigma}_e^2 = \frac{1-2\theta\tilde{H}+\theta^2}{1-\tilde{H}^2}, \tag{2.5}$$

and

$$\xi = \left(\tilde{\sigma}_e^2 - \sigma_e^2\right)\left(1-\tilde{H}^{2k}\right). \tag{2.6}$$

The Appendix provides methods to compute these long-run MSEO and input variances that apply for responsive and first-order dynamics.

2.5 Scheme 5: Feedback With Both Direct and Indirect Feedforward (FB + FF + FFS)

In addition to feedback, feedforward is used to cancel the measured batch mean and also to temporally increase the sensitivity of the feedback adjustment. The adjustment equations for the batch starting at time $t = T$ are as in Section 2.4 excepting that now $g x_{T-1} = -G e_{T-1} - m_T$, $g x_{2T-1} = -G e_{2T-1} - m_{2T}$, and so on.

Consequently, noting that $\mu_T = m_T - \varepsilon_T$ in Equation (2.4b), the term containing m_T disappears from (2.4b) and hence the effect of m_T on the MSEO, M_5, is zero. The variation associated with μ_T is passed to the input variance V_5, which now contains a new term of magnitude σ_m^2 / T. Consequently, for large T,

$$M_5 = \frac{k}{T}\tilde{\sigma}_e^2 + \frac{T-k}{T}\sigma_e^2 + \frac{\sigma_\varepsilon^2}{T}\left(\frac{\tilde{H}^{2k}}{1-H^2} + \frac{1-\tilde{H}^{2k}}{1-\tilde{H}^2}\right) + \frac{\xi}{T}\left(\frac{1}{1-H^2} - \frac{1}{1-\tilde{H}^2}\right)$$

and

$$V_5 = \frac{k}{T}\tilde{\sigma}_e^2\tilde{G}^2 + \frac{T-k}{T}\sigma_e^2 G^2 + \frac{\sigma_\varepsilon^2}{T}\left(\frac{G\tilde{H}^{2k}}{2-G} + \frac{\tilde{G}\left(1-\tilde{H}^{2k}\right)}{2-\tilde{G}}\right) + \frac{\sigma_m^2}{T} + \frac{\xi}{T}\left(\frac{G}{2-G} - \frac{\tilde{G}}{2-\tilde{G}}\right),$$

where σ_e^2, $\tilde{\sigma}_e^2$, and ξ are given by (2.2), (2.5), and (2.6), respectively.

2.6. Relative Performance of the Five Schemes.

Comparison of the various schemes is made easier by considering close approximations to the exact expressions derived earlier. These are shown in Table 1.

It is important to notice that the terms in the table for M_i (i = 1, …, 5) that change from one scheme to another can be of order T, T^0, or T^{-1} (the terms of order T^0 are enclosed by square brackets in Table 1). Thus,

a) Unless T is small, the schemes 2, 3, 4, and 5 that include feedback will yield markedly smaller values of M_i than the pure feedforward scheme 1 because they eliminate the nonstationary term of order T that is present in M_1.

b) By augmenting feedback control with feedforward as in scheme 3 compared with scheme 2, or in scheme 5 compared with scheme 4, the batch to batch variance σ_μ^2 is replaced by the measurement variance σ_ε^2; however the terms involved are only of order T^{-1}.

c) By appropriate choice of k, G, and $\tilde{G}$, the MSEO in schemes 4 and 5 can always be made smaller than those in schemes 2 and 3, respectively.

(Please place Tables 1 and 2 near here.)

For given estimates of σ_μ, σ_ε, and λ, using these expressions, the values of M_i and V_i for the various schemes may be calculated and compared. The results should be interpreted in the light of practical necessities and in particular the convenience of running the different schemes. To illustrate these calculations, we consider the following example. Suppose that the values of the parameters are $T = 100$, $\theta = .8$ ($\lambda = .2$), $\sigma_a = 1$, $\sigma_\mu = 2$, $\sigma_\varepsilon = .5$, $G = .2$, $\tilde{G} = .8$, and $k = 3$. Thus there is a 25% error in the measurements of the batch means. From Table 2 we see that

(i) Because the background disturbance is nonstationary, without feedback control to eliminate this nonstationary disturbance the MSEO, M_1, will be large.

(ii) The MSEO M_2 is greatly reduced because the nonstationarity is eliminated by feedback adjustment. However, it still contains a component depending on σ_μ because of the transient in the response to feedback. The input variances are almost the same as in scheme FF; the main difference is that scheme FB makes small adjustments every time, whereas only one big adjustment at each time of batch change is called for by scheme FF.

(iii) A further reduction in MSEO (from M_2 to M_3) is obtained by eliminating the measured batch changes by feedforward. However, a transient due to measurement errors still remains. The input variance V_3 is almost twice V_2.

(iv) The addition of indirect (rather than direct) feedforward to increase feedback sensitivity produces in this case a MSEO M_4 smaller than M_2 but larger than M_3. Because the damping factor $\tilde{G}$ has not been chosen carefully, the required input variance V_4 is larger than V_2 and V_3.

(v) The addition of direct and indirect feedforward produces MSEO M_5 and input variance V_5 larger than M_3 and V_3, respectively, again because of the choice of G and $\tilde{G}$.

3. OPTIMAL VALUES FOR THE DAMPING FACTORS

In Table 2, the fact that M_5 is larger than M_3 deserves further consideration. In scheme 3, there is one free parameter (G) to be chosen arbitrarily, whereas in scheme 5 there are two such parameters (G and $\tilde{G}$). When the free parameters G and, possibly, $\tilde{G}$ are chosen to minimize M_i, appropriate choices can always be made to yield a value of M_5 smaller than M_3. Table 3 illustrates the relative performance of the schemes previously considered. The values of the parameters in Table 3 are again $T = 100$, $\theta = .8$ ($\lambda = .2$), $\sigma_a = 1$, $\sigma_\mu = 2$, $\sigma_\varepsilon = .5$, and $k = 3$, but G and $\tilde{G}$ are now chosen to minimize the MSEO. We see that

(i) The MSEOs may be reduced somewhat by choosing G and $\tilde{G}$ adequately.

(ii) The MSEO produced by feedback control M_2 may be reduced to M_3 by adding direct feedforward, without increasing the average input variance appreciably.

(iii) Indirect feedforward is less efficient that direct feedforward in reducing the MSEO.

(iv) The scheme using direct and indirect feedforward provides values for M_5 and V_5 that are slightly smaller than M_3 and V_3 for direct feedforward, but the latter may be simpler to use.

(Please place Tables 3 and 4 near here.)

4. COMPARISON OF OUTPUT MEAN SQUARED ERRORS WITH INPUT VARIANCE FIXED

A difficulty in comparing schemes 2, 3, 4, and 5 is that we cannot be guided by the values of M_i alone for, as we see in Tables 2 and 3, these can be associated with very different values of the required input variances V_i. To overcome this difficulty, constrained feedback adjustment schemes may be explored that minimize the MSEO with respect to G, and possibly $\tilde{G}$, subject to the restriction of the input variance INPVAR $\leq \gamma$, where γ is some chosen constant. (For the particular cases in which γ is very large, unconstrained adjustment schemes such as those of Table 3 are obtained.)

The input variances may be reduced considerably by using constrained adjustment schemes provided that a modest increase in the MSEO may be accepted. In Table 4 therefore we have calculated the values of M_i when V_i have the same fixed bounds. Three examples are given with the input variances bounded by $\gamma = .06$, .0525, and .045. We see that the schemes FB+FF and FB+FF+FFS provide almost the same overall performance, which is better than the performance of the other schemes when $\gamma = .06$. However, as the constraint in the input variance is more important (i.e., as γ decreases), the performance of the scheme FB+FFS improves in comparison with the other schemes and is the best for $\gamma = .045$.

5. CONCLUDING REMARKS

The comparison among the five schemes of Section 1.2 may be summarized as follows:

(i) Because feedforward adjustment alone cannot eliminate nonstationary disturbances, and as in any case based on dead reckoning, any reasonable form of feedback control is likely to do much better.

(ii) If scheme 2 employing simple feedback is used as a basis for comparison, some improvement in mean squared error at the output is possible by use of direct and/or indirect feedforward to supplement the feedback scheme.

(iii) For unconstrained schemes, the use of direct feedforward to complement feedback is better than simple feedback alone but, for highly constrained schemes it does less well.

(iv) For constrained schemes, the sensitized feedforward scheme to complement feedback outperforms the other schemes.

ACKNOWLEDGMENTS

This research was partially supported by the National Science Foundation Grant DMI-9812839 and the Spanish Grant DGESIC-PB97-0555.

APPENDIX. LONG-RUN MEAN SQUARED ERROR AT THE OUTPUT AND INPUT VARIANCE FOR THE SCHEMES OF SECTION 2 UNDER FIRST-ORDER DYNAMICS

Scheme "4" with first-order dynamics. For generality, we consider the evaluation of the long-run mean squared error at the output and input variance under first-order dynamics and proportional integral (PI) control. The PI adjustment equation is

$$g\,x_t = -G\left[(1+P) - PB\right]e_t\,; \tag{A.1}$$

where G and P are constants. The error at the output is

$$e_t = z_t + \frac{g(1-\delta)}{1-\delta B}X_{t-1},$$

where B is the backshift operator such that $B\,X_t = X_{t-1}$, g is the system gain and δ is the inertia. The total disturbance is given by

$$z_t - z_{t-1} = a_t - \theta\, a_{t-1} + \mu_t\,,$$

where $\mu_t = 0$ for any $t \neq T, 2T, 3T, \ldots$.

For $t < T$, the errors at the output satisfy

$$e_t = a_t - \theta_1\, a_{t-1} - \theta_2\, a_{t-2} + \phi_1\, e_{t-1} + \phi_2\, e_{t-2}\,, \tag{A.2}$$

where $\phi_1 = 1 + \delta - G(1+P)(1-\delta)$, $\phi_2 = -\delta + G\,P(1-\delta)$, $\theta_1 = \theta + \delta$, and $\theta_2 = -\theta\delta$. Similarly, for $t = T$, we have

$$e_T = a_T - \theta_1 a_{T-1} - \theta_2\, a_{T-2} + \phi_1\, e_{T-1} + \phi_2\, e_{T-2} + \mu_T\,. \tag{A.3}$$

Equation (A.3) shows that μ_T affect the error at the output e_T at time T so that we temporally increase the sensitivity of the feedback scheme by replacing (A.1) with

$$g\,x_t = -\tilde{G}\left[\left(1+\tilde{P}\right) - \tilde{P}B\right]e_t \quad (t = T, \ldots, T+k-1). \tag{A.4}$$

Then, for $T+1 \leq t \leq T+k$, the errors at the output satisfy

$$e_t = a_t - \theta_1\, a_{t-1} - \theta_2\, a_{t-2} + \tilde{\phi}_1\, e_{t-1} + \tilde{\phi}_2\, e_{t-2}\,,$$

where $\tilde{\phi}_1 = 1 + \delta - \tilde{G}\left(1+\tilde{P}\right)(1-\delta)$ and $\tilde{\phi}_2 = -\delta + \tilde{G}\,\tilde{P}(1-\delta)$. Finally, for $T+k < t \leq 2T$, the underlying PI scheme is resumed so that the errors at the output satisfy (A.2) for $T+k < t < 2T$ and (A.3) with T replaced by $2T$, and so on.

From (A.2), the expressions for $\zeta_0 = \text{var}(e_t)$ and $\zeta_1 = \text{cov}(e_t, e_{t-1})$, for $t < T$, can be obtained by using standard time series methods (e.g., see Luceño 1993). Assuming for the moment that $\mu_T = 0$, Equations (A.2) and (A.3) can be used to obtain

$$\text{var}(e_T) = (1 + \mathbf{v}'\mathbf{A}_T\mathbf{v})\sigma_a^2,$$
$$\text{var}(e_{T+1}) = (1 + \tilde{\mathbf{v}}'\mathbf{A}_{T+1}\tilde{\mathbf{v}})\sigma_a^2,$$
$$\text{var}(e_t) = \left(1 + \tilde{\mathbf{v}}'\tilde{\mathbf{A}}_t\tilde{\mathbf{v}}\right)\sigma_a^2, \quad (T+2 \le t \le T+k)$$
$$\text{var}(e_{T+k+1}) = \left(1 + \mathbf{v}'\tilde{\mathbf{A}}_{T+k+1}\mathbf{v}\right)\sigma_a^2,$$
$$\text{var}(e_t) = (1 + \mathbf{v}'\mathbf{A}_t\mathbf{v})\sigma_a^2, \quad (T+k+2 \le t \le 2T),$$

and so on, where $\mathbf{v}' = (-\theta_1, -\theta_2, \phi_1, \phi_2)$ and $\tilde{\mathbf{v}}' = (-\theta_1, -\theta_2, \tilde{\phi}_1, \tilde{\phi}_2)$; and $\sigma_a^2\mathbf{A}_t$ and $\sigma_a^2\tilde{\mathbf{A}}_t$ stand for the covariance matrix of vector $(a_{t-1}, a_{t-2}, e_{t-1}, e_{t-2})$ depending on whether this matrix is a function of ϕ_1 and ϕ_2 or of $\tilde{\phi}_1$ and $\tilde{\phi}_2$. Thus for $t = T$ or $T + 1$,

$$\mathbf{A}_t = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & \alpha_t & 1 \\ 1 & \alpha_t & \beta_t & \gamma_t \\ 0 & 1 & \gamma_t & \delta_t \end{pmatrix},$$

where $\alpha_T = \phi_1 - \theta_1$, $\delta_T = \zeta_0 / \sigma_a^2$, $\gamma_T = \zeta_1 / \sigma_a^2$, and $\beta_T = \delta_T$, and then

$$\alpha_{T+1} = \phi_1 - \theta_1, \quad \delta_{T+1} = \beta_T, \quad \gamma_{T+1} = -\theta_1 - \theta_2\alpha_T + \phi_1\beta_T + \phi_2\gamma_T, \tag{A.5a}$$
$$\beta_{T+1} = 1 - \theta_1\alpha_{T+1} - \theta_2(-\theta_2 + \phi_1\alpha_T + \phi_2) + \phi_1\gamma_{T+1} + \phi_2(-\theta_2 + \phi_1\gamma_T + \phi_2\delta_T). \tag{A.5b}$$

Similarly, for $T + 2 \le t \le T + k + 1$,

$$\tilde{\mathbf{A}}_t = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & \tilde{\alpha}_t & 1 \\ 1 & \tilde{\alpha}_t & \tilde{\beta}_t & \tilde{\gamma}_t \\ 0 & 1 & \tilde{\gamma}_t & \tilde{\delta}_t \end{pmatrix},$$

where $\tilde{\alpha}_t$, $\tilde{\delta}_t$, $\tilde{\gamma}_t$, and $\tilde{\beta}_t$ are computed using (A.5) where ϕ_1 and ϕ_2 are replaced by $\tilde{\phi}_1$ and $\tilde{\phi}_2$ and $T + 1$ by t. Finally, the long-run mean squared error at the output (without taking the effect of μ_T into account for the moment) is computed as $M_{41} = T^{-1}\sum_{t=T}^{2T-1}\text{var}(e_t)$, where T is assumed to be large enough for the effect of μ_T to settled down before the next step change.

The corresponding V_{41} computed without taking the effect of μ_T into account (for the moment) may be computed using (A.1) and (A.4). Thus

$$\mathrm{var}(g x_t) = \tilde{G}^2\left[\left(1+\tilde{P}\right)^2 \beta_{t+1} + \tilde{P}^2 \delta_{t+1} - 2\tilde{P}\left(1+\tilde{P}\right)\gamma_{t+1}\right]\sigma_a^2, \quad (t = T, \ldots, T+k-1),$$

$$\mathrm{var}(g x_t) = G^2\left[(1+P)^2 \beta_{t+1} + P^2 \delta_{t+1} - 2P(1+P)\gamma_{t+1}\right]\sigma_a^2, \quad (t = T+k, \ldots, 2T-1),$$

and $V_{41} = T^{-1}\sum_{t=T}^{2T-1} \mathrm{var}(g x_t)$.

The effect of μ_T is given by $M_{42} = \sigma_\mu^2 T^{-1} \sum_{t=0}^{T-1} \pi_t^2$, where $\pi_{-1} = 0$, $\pi_0 = 1$, $\pi_t = \tilde{\phi}_1 \pi_{t-1} + \tilde{\phi}_2 \pi_{t-2}$ $(t = 1, \ldots, k)$, and $\pi_t = \phi_1 \pi_{t-1} + \phi_2 \pi_{t-2}$ $(t > k)$. Correspondingly, $V_{42} = T^{-1}\sum_{t=T}^{2T-1} \mathrm{var}_2(g x_t)$ where

$$var_2(g x_t) = \left[-\tilde{G}\left(1+\tilde{P}\right)\pi_{t-T} + \tilde{G}\tilde{P}\pi_{t-T-1}\right]^2 \sigma_\mu^2, \quad (t = T, \ldots, T+k-1),$$

$$var_2(g x_t) = \left[-G(1+P)\pi_{t-T} + GP\pi_{t-T-1}\right]^2 \sigma_\mu^2, \quad (t = T+k, \ldots, 2T-1).$$

Finally, the long-run mean squared error at the output M_4 is the sum of M_{41} and M_{42}, and the long-run input variance V_4 is the sum of V_{41} and V_{42}.

Scheme "2" with first-order dynamics. The formulas for scheme "2" producing no augmentation of the feedback system can be obtained by substituting (G, P) for $(\tilde{G}, \tilde{P})$. Assuming that T is large enough for the effect of μ_T to settled down before the next step change, we obtain

$$M_2 = \zeta_0 + \frac{\sigma_\mu^2}{T}\left(\frac{1-\phi_2}{1+\phi_2}\right)\frac{1}{(1-\phi_2)^2 - \phi_1^2},$$

and, similarly,

$$V_2 = G^2 \zeta_0 \left[1 + 2P(1+P)(1 - \zeta_1/\zeta_0)\right]$$
$$+ \frac{\sigma_\mu^2}{T}\frac{G^2}{1+\phi_2}\frac{1}{(1-\phi_2)^2 - \phi_1^2}\left\{(1-\phi_2)\left[(1+P)^2 + P^2\right] - 2\phi_1 P(1+P)\right\}$$

Scheme "3" with first-order dynamics. The formulas for scheme "3" using feedforward to compensate directly for the measurable part of the input change can be obtained similarly, so that

$$M_3 = \zeta_0 + \frac{\sigma_\varepsilon^2}{T}\left(\frac{1-\phi_2}{1+\phi_2}\right)\frac{1}{(1-\phi_2)^2 - \phi_1^2},$$

and

$$V_3 = \sigma_m^2\left(1+\delta^2\right)(1-\delta)^2/T + G^2\zeta_0\left[1 + 2P(1+P)(1 - \zeta_1/\zeta_0)\right]$$
$$+ \frac{\sigma_\varepsilon^2}{T}\frac{G^2}{1+\phi_2}\frac{1}{(1-\phi_2)^2 - \phi_1^2}\left\{(1-\phi_2)\left[(1+P)^2 + P^2\right] - 2\phi_1 P(1+P)\right\}$$

Scheme "5" with first-order dynamics. Compute M_4 and V_4 using the methods for scheme "4" but replacing σ_μ^2 by σ_ε^2. Then take $M_5 = M_4$ and $V_5 = V_4 + \sigma_m^2 \left(1+\delta^2\right)/(1-\delta)^2 / T$.

REFERENCES

Alwan, L. C. and Roberts, H. V. (1988), "Time-Series Modeling for Statistical Process Control," *Journal of Business and Economic Statistics*, 6, 87–95.

Box, G. E. P. and Jenkins, G. M. (1970), *Time Series Analysis, Forecasting and Control*, San Francisco: Holden-Day.

Box, G. E. P., Jenkins, G. M., and Reinsel, G. C. (1994), *Time Series Analysis, Forecasting and Control*, Prentice Hall, Englewood Cliffs, NJ (3 rd edition).

Box, G. E. P. and Kramer, T. (1992), "Statistical Process Monitoring and Feedback Adjustment—A Discussion," *Technometrics*, 34, 251–285.

Box, G. E. P. and Luceño, A. (1995), "Discrete Proportional-Integral Control with Constrained Adjustment," *Journal of the Royal Statistical Society, D*, 44, 479–495.

______ (1997), *Statistical Control by Monitoring and Feedback Adjustment*, New York: John Wiley.

______ (2000), "Six Sigma, Process Drift, Capability Indices, and Feedback Adjustment," *Quality Engineering*, 12, 297–302.

Deming, W. E. (1986), *Out of the Crisis*, Cambridge, MA: Massachusetts Institute of Technology, Center for Advanced Engineering Studies.

Fearn, T. and Maris, P. I. (1991), "An Application of Box-Jenkins Methodology to the Control of Gluten Addition in a Flour Mill," *Applied Statistics*, **40**, 477–484.

Harry, M. (1994), "The Vision of Six Sigma. Roadmap for Breakthrough," *Sigma Publishing Company*, May, 1988, 60 – 64.

Hoerl, R. W. (1998), "Six Sigma and the Future of the Quality Profession," *Quality Progress*, June, 35 – 42.

Jensen, K. L. and Vardeman, S. B. (1993), "Optimal Adjustment in the Presence of Deterministic Process Drift and Random Adjustment Error," *Technometrics*, **35**, 376–389.

Luceño, A. (1993), "A Fast Algorithm for the Repeated Evaluation of the Likelihood of a General Linear Process for Long Series," *Journal of the American Statistical Association,* 88, 229-236.

______ (1998), "Performance of Discrete Feedback Adjustment Schemes With Dead Band, Under Stationary Versus Nonstationary Stochastic Disturbance," *Technometrics*, **40**, 223–233.

______ (2000), "Minimum Cost Dead Band Adjustment Schemes Under Tool-Wear Effects and Delayed Dynamics," *Statistics and Probability Letters*, **50**, 165–178.

Luceño, A., González, F. J., and Puig-Pey J. (1996), "Computing Optimal Adjustment Schemes for the General Tool-Wear Problem," *Journal of Statistical Computation and Simulation*, **54**, 87–113.

MacGregor, J. F. (1972), *Topics in the Control of Linear Processes Subject to Stochastic Disturbances*, Ph. D. Thesis, University of Wisconsin-Madison.

Montgomery, D. C. and Woodall, W. H. (1997) A Discussion of Statistically-Based Process Monitoring and Control. *Journal of Quality Technology*, **29**, 179–193.

Smith, W. (1992), *Presentation at the Case Study Conference*, Center for Quality and Productivity Improvement, University of Wisconsin-Madison, Madison, WI.

Tucker, W. T., Faltin, F. W., and Vander Wiel, S. A. (1993), "Algorithmic Statistical Process Control: An Elaboration," *Technometrics*, **35**, 363–375.

Vander Wiel, S. A., Tucker, W. T., Faltin, F. W., and Doganaksoy, N. (1992), "Algorithmic Statistical Process Control: Concepts and an Application," *Technometrics*, **34**, 286–297.

TABLE 1a

Scheme	Long-run mean squared errors at the output
1) FF	$M_1 = \frac{T-1}{2}\lambda^2 + \left[1+\sigma_\varepsilon^2\right]$
2) FB	$M_2 = \left[\frac{1-2\theta H+\theta^2}{1-H^2}\right] + \frac{\sigma_\mu^2}{T\left(1-H^2\right)}$
3) FB+FF	$M_3 = \left[\frac{1-2\theta H+\theta^2}{1-H^2}\right] + \frac{\sigma_\varepsilon^2}{T\left(1-H^2\right)}$
4) FB+FFS	$M_4 \cong \left[\frac{k}{T}\,\frac{1-2\theta\tilde{H}+\theta^2}{1-\tilde{H}^2} + \frac{T-k}{T}\,\frac{1-2\theta H+\theta^2}{1-H^2}\right] + \frac{\sigma_\mu^2}{T}\left(\frac{\tilde{H}^{2k}}{1-H^2} + \frac{1-\tilde{H}^{2k}}{1-\tilde{H}^2}\right)$
5) FB+FF+FFS	$M_5 \cong \left[\frac{k}{T}\,\frac{1-2\theta\tilde{H}+\theta^2}{1-\tilde{H}^2} + \frac{T-k}{T}\,\frac{1-2\theta H+\theta^2}{1-H^2}\right] + \frac{\sigma_\varepsilon^2}{T}\left(\frac{\tilde{H}^{2k}}{1-H^2} + \frac{1-\tilde{H}^{2k}}{1-\tilde{H}^2}\right)$

TABLE 1b

Scheme	Long-run adjustment variances
1) FF	$V_1 = \sigma_m^2 / T$
2) FB	$V_2 = G^2 M_2$
3) FB+FF	$V_3 = G^2 M_3 + \sigma_m^2 / T$
4) FB+FFS	$V_4 \cong \frac{k}{T}\,\frac{1-2\theta\tilde{H}+\theta^2}{1-\tilde{H}^2}\,\tilde{G}^2 + \frac{T-k}{T}\,\frac{1-2\theta H+\theta^2}{1-H^2}\,G^2 + \frac{\sigma_\mu^2}{T}\left(\frac{G\tilde{H}^{2k}}{2-G} + \frac{\tilde{G}\left(1-\tilde{H}^{2k}\right)}{2-\tilde{G}}\right)$
5) FB+FF+FFS	$V_5 \cong \frac{k}{T}\,\frac{1-2\theta\tilde{H}+\theta^2}{1-\tilde{H}^2}\,\tilde{G}^2 + \frac{T-k}{T}\,\frac{1-2\theta H+\theta^2}{1-H^2}\,G^2 + \frac{\sigma_\varepsilon^2}{T}\left(\frac{G\tilde{H}^{2k}}{2-G} + \frac{\tilde{G}\left(1-\tilde{H}^{2k}\right)}{2-\tilde{G}}\right) + \frac{\sigma_m^2}{T}$

TABLE 2

Scheme	output variance	input variance
1) FF	$M_1 = 3.2300$	$V_1 = .04250$
2) FB	$M_2 = 1.1111$	$V_2 = .04444$
3) FB+FF	$M_3 = 1.0069$	$V_3 = .08278$
4) FB+FFS	$M_4 = 1.0594$	$V_4 = .08978$
5) FB+FF+FFS	$M_5 = 1.0204$	$V_5 = .10728$

TABLE 3

Scheme	Optimal G	Optimal $\tilde{G}$	output variance	input variance
1) FF	not used	not used	$M_1 = 3.2300$	$V_1 = .04250$
2) FB	.2702	not used	$M_2 = 1.0961$	$V_2 = .08000$
3) FB+FF	.2055	not used	$M_3 = 1.0069$	$V_3 = .08500$
4) FB+FFS	.2053	.6409	$M_4 = 1.0563$	$V_4 = .07399$
5) FB+FF+FFS	.2011	.3014	$M_5 = 1.0059$	$V_5 = .08491$

TABLE 4

	Scheme	Optimal G	Optimal $\tilde{G}$	output variance	input variance
	1) FF	not used	not used	$M_1 = 3.2300$	$V_1 = .04250$
	2) FB	.2336	not used	$M_2 = 1.0997$	$V_2 = .06000$
a)	3) FB+FF	.1303	not used	$M_3 = 1.0302$	$V_3 = .06000$
	4) FB+FFS	.1868	.5624	$M_4 = 1.0581$	$V_4 = .06000$
	5) FB+FF+FFS	.1280	.1852	$M_5 = 1.0291$	$V_5 = .06000$
	1) FF	not used	not used	$M_1 = 3.2300$	$V_1 = .04250$
	2) FB	.2181	not used	$M_2 = 1.1038$	$V_2 = .05250$
b)	3) FB+FF	.0966	not used	$M_3 = 1.0718$	$V_3 = .05250$
	4) FB+FFS	.1752	.5183	$M_4 = 1.0612$	$V_4 = .05250$
	5) FB+FF+FFS	.0951	.1324	$M_5 = 1.0708$	$V_5 = .05250$
	1) FF	not used	not used	$M_1 = 3.2300$	$V_1 = .04250$
	2) FB	.2013	not used	$M_2 = 1.1105$	$V_2 = .04500$
c)	3) FB+FF	.0436	not used	$M_3 = 1.3162$	$V_3 = .04500$
	4) FB+FFS	.1623	.4721	$M_4 = 1.0667$	$V_4 = .04500$
	5) FB+FF+FFS	.0432	.00526	$M_5 = 1.3157$	$V_5 = .04500$

Table 1. (a) Long-run Mean Squared Errors at the Output for Five Alternative Control Schemes; (b) Corresponding Long-run Adjustment Variances.

Table 2. Relative Performance of the Schemes in Table 1. An example with $T = 100$, $\theta = .8$, $\sigma_a = 1$, $\sigma_\mu = 2$, $\sigma_\varepsilon = .5$, $G = .2$, $\tilde{G} = .8$, and $k = 3$.

Table 3. Relative Performance of the Optimal Unconstrained Schemes of the Types Considered in Section 3. An example with $T = 100$, $\theta = .8$, $\sigma_a = 1$, $\sigma_\mu = 2$, $\sigma_\varepsilon = .5$, and $k = 3$.

Table 4. Relative Performance of the Optimal Constrained Schemes With Input Variance Fixed At (a) 0.06, (b) 0.0525, (c) 0.045. The parameters are $T = 100$, $\theta = .8$, $\sigma_a = 1$, $\sigma_\mu = 2$, $\sigma_\varepsilon = .5$, and $k = 3$.