Center for Quality and Productivity Improvement
University of Wisconsin-Madison
610 Walnut Street
Madison, Wisconsin 53705

Report No. 42

PROCESS CONTROL FROM AN ECONOMIC POINT OF VIEW

Chapter 1: Industrial Process Control

Tim Kramer*

February, 1990

This research was sponsored by the National Science Foundation under Grant No. DDM-8808138

*Applied Statistician, Hewlett Packard, Avondale, Pennsylvania

This report is part of a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Statistics) at the University of Wisconsin-Madison (1989). Thesis advisor: George E.P. Box.

CENTER FOR QUALITY AND PRODUCTIVITY IMPROVEMENT
UNIVERSITY OF WISCONSIN-MADISON

Report No. 42

PROCESS CONTROL FROM AN ECONOMIC POINT OF VIEW

CHAPTER 1: INDUSTRIAL PROCESS CONTROL

Tim Kramer

February 1990

PRACTICAL SIGNIFICANCE

Control schemes can take many forms. In particular, Shewhart control charts and cumulative sum charts are frequently employed for what is called Statistical Process Control (SPC). By contrast, various forms of feedback, feedforward and feedforward-feedback systems are used for what may be called Automatic Process Control (APC). The people responsible for SPC and those responsible for APC are usually from different departments and different technical backgrounds so it is hardly surprising that occasionally there is controversy and misunderstanding between the two groups. In this report arguments are presented showing how and why each kind of control scheme serves a valuable purpose when used in a suitable context, but can fail and mislead if employed inappropriately.

Key Words: Control, Regulation, SPC, Common Causes, Special Causes

---

This research was sponsored by the National Science Foundation under Grant No. DDM-8808138.

This report is part of a thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Statistics) at the University of Wisconsin-Madison (1989). Thesis advisor: George E.P. Box.

The Center for Quality and Productivity Improvement cares about your reactions to our reports. Please send comments (general or specific) to: Report Feedback, Center for Quality and Productivity Improvement, 610 Walnut Street, Madison, WI 53705. All replies will be forwarded to the authors.

PROCESS CONTROL FROM AN ECONOMIC POINT OF VIEW

CHAPTER 1: INDUSTRIAL PROCESS CONTROL

TABLE OF CONTENTS

CHAPTER 1: INDUSTRIAL PROCESS CONTROL

1.1 Introduction

Control schemes can take many forms. In particular, Shewhart control charts and cumulative sum charts are frequently employed for what is called Statistical Process Control (SPC). By contrast, various forms of feedback, feedforward, and feedforward-feedback systems are used for what may be called Automatic Process Control (APC). The people responsible for SPC and those responsible for APC are usually from different departments and different technical backgrounds so it is hardly surprising that occasionally there is controversy and misunderstanding between the two groups. One reason for this is that there are different purposes for which control schemes are used and, consequently, different meanings for the words *process control*. On the one hand, the principal purpose may be *to learn about* the process and to eliminate assignable causes of disturbance so as to produce and maintain the process in a ''state of control''. On the other hand, the principal purpose may be *to regulate* the process so as to maintain it as close as possible to some desired target value. Although SPC is principally concerned with the first purpose and APC with the second, these goals are not mutually exclusive and may be pursued concurrently.

This chapter presents arguments showing how and why each kind of control scheme serves a valuable purpose when used in a suitable context, but can fail and mislead if employed inappropriately.

1.2 Some Control Schemes for Detection and Removal of Special Causes

Shewhart control

In Lewis Carroll's *Through the Looking Glass*, the White Queen says to Alice, "Now, *here*, you see, it takes all the running *you* can do to keep in the same place". This remark applies to the real world, not just Looking Glass Land. In the parts industry, for example, good quality requires that we manufacture the same thing consistently. It would be nice if this could be done by once and for all adjusting and setting the machine and then letting it run. Unfortunately, this would rarely result in the manufacture of uniform parts, at least not for any length of time. Alternatively, consider some routine task in a hospital such as the taking of blood pressure. Once a best way for taking blood pressure is decided, we would like it done that way consistently. Experience shows, however, that this consistency will not happen unless extraordinary precautions are taken to ensure it. The truth is that we live in a nonstationary world, that is, a world in which external factors never stay still. Indeed, the ideas of stationarity and stability in which things stay put over time is, in reality, a purely conceptual and imaginary concept. It is useful, only because it supplies a standard against which the real unstable, nonstationary world can be judged.

The manufacture of parts is an operation involving machines and people. The parts of a machine are not fixed entities; they wear out, change their dimensions, and lose their adjustment. The people who run the machines differ in their behavior. A single operator forgets things over time and may fail to communicate with others, and when many operators are involved opportunities to vary from the standard procedure are multiplied. Thus, stationarity, stability, and uniformity over time is not the norm, as many statistics texts would have us believe, but is an unnatural state that we must work hard to achieve. The Shewhart chart is a very important tool in this endeavor. Its tremendous virtue is the simple fact that successive observations are plotted in time order, so that time patterns of normal and abnormal behavior of the process can be clearly seen and mulled by a viewer who is familiar with the process. To aid with the assessment of what might be and what might not be normal process behavior, limit (control) lines are often placed at $\pm 3\sigma$ about the target line T. Because many processes tend to remain stable over short periods of time, a measure of this short term standard deviation, σ, is usually employed to judge normal behavior. This standard deviation is determined from short lengths, called rational subgroups, of normal operation of the process. Control lines are plotted under the assumption that, under stable operating conditions, the data would vary in a roughly $N(T, \sigma^2)$ distribution, i.e., a normal distribution having mean equal to the target value T with standard deviation σ.

The process is said to be in a state of control when its usual operation is in conformance with the $N(T, \sigma^2)$ model. The occasional occurrence of a suspicious pattern which is unlikely under this model may then produce a search for what Shewhart called an ''assignable cause''. Deming calls this a *special* cause and he attributes the normal variation of the process about the target value to *common* causes. Thus the effort expended in process improvement can be classified into two categories. Once the process is in a state of control, the occasional occurrence of suspicious patterns can initiate "find and fix" investigations associated with special causes. These investigations remove "bugs" from the process but do not basically change it. On the other hand, fundamental management-induced system changes are needed to deal with common causes. Such changes may improve quality by producing a change in the mean level, μ, or reducing the standard deviation, σ, or both. For example, because of management introducing a new machine for painting cars a higher gloss may be obtained or a more uniform gloss or both.

Control and improvement

There should of course be no conflict between process control and process improvement. Indeed, as is emphasized by Imai [23], quality control is essential to "lock in" quality improvement. If this is not done, the history of an improvement that initially increased the mean for some quality characteristic might resemble the graph in figure 1.1(a). This is because carefully adjusted machines can become untuned, people can become less careful and supervision can become lax. To prevent such regression, Imai emphasizes that it is necessary to build a quality control system so that, as illustrated in figure 1.1(b), quality is locked in after each new improvement.

Roles of management and worker

Deming and Juran have pointed out that more than 85% of problems are concerned with the nature of the system itself and so are the responsibility of management whereas less than 15% are related to deviations from the system attributable to the workforce. This has sometimes led to the false conclusion that

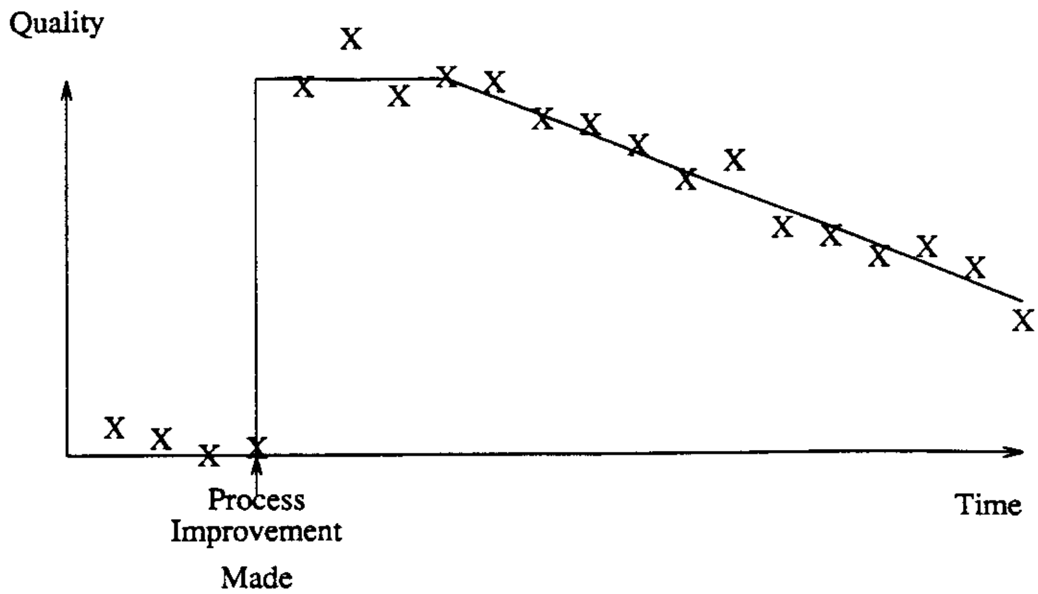


Figure 1.1(a): Quality Improvement Not Locked In

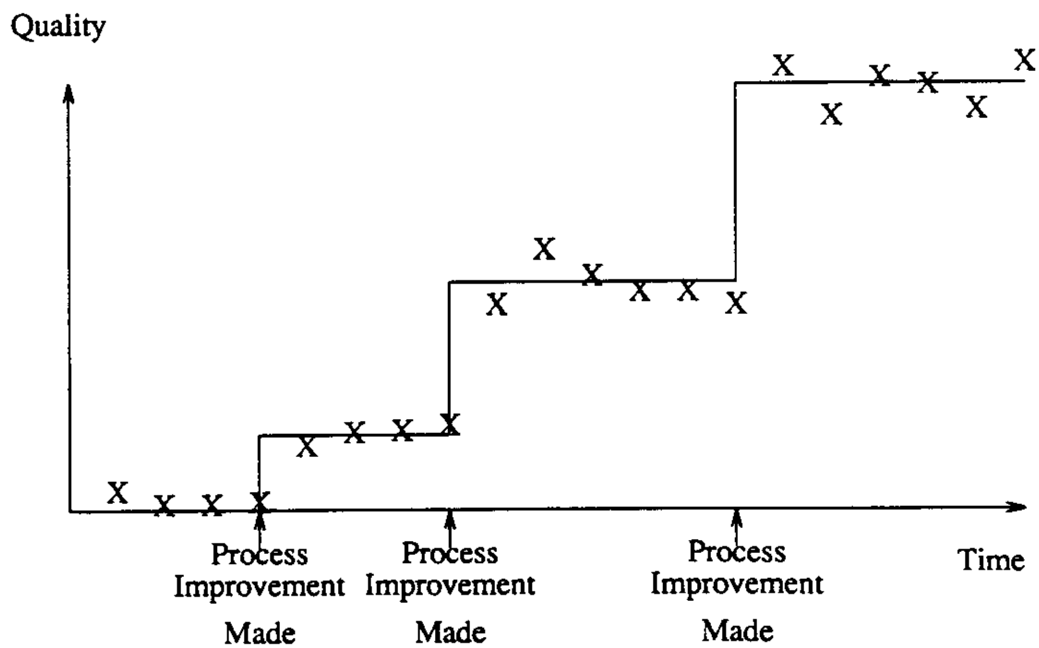


Figure 1.1(b): Quality Improvement Locked In

worker participation is unimportant since it can address only 15% of the problems. However, problem solving teams in which workers participate can work with management in the solution of system problems as well as nonconformance problems. In such team efforts, the manager's role becomes that of a leader and coach.

Cusum and other charts based on likelihood ratios

The Shewhart control chart has the advantage that it does not assume that the special causes are associated with any particular pattern. In principle any suspicious pattern might be identified with a Shewhart chart. However, if we were prepared to look for a specific pattern associated with a particular kind of special cause, then, by putting the problem into a hypothesis testing framework, we can compute a best test to detect that pattern. If, for example, we had successfully eliminated a disturbance giving rise to a sine wave of a particular period but feared that it might return, then using the appropriate likelihood ratio (Box and Ramírez [14]), we could employ a test function and a corresponding chart that could detect that particular deviation from randomness more quickly than a Shewhart chart. We could, however, not afford to use such a specialized test alone since it might be useless to detect unexpected forms of disturbance. The problem has been likened (Box [8]) to that faced by a small country wishing to install an early warning radar system against air attack. Very sensitive radar could be used to monitor certain directions known to be likely sources of attack but, for safety, a less sensitive multidirectional screen would have to be used as well. One particular deviant pattern which has received considerable attention occurs when the process mean changes from one value μ_1 to another value μ_2. The optimal likelihood ratio test is then the Wald-Barnard sequential test leading to the use of a cumulative sum (cusum) of the deviations from target. Such is the basis for the commonly used cusum charts (see, for example, Barnard [6], Ewan and Kemp [21], Lucas [29]).

1.3 Some Control Schemes for Regulation

In a familiar prayer one asks for guidance on how "to change what can be changed, to learn to live with what cannot, and to know the difference." Because the outside temperature varies markedly with seasons, most dwellings in the U.S. have a feedback control device–a thermostat driving the furnace to produce an even temperature in the home. Changes in climatic temperature in the U.S. are disturbances that cannot be removed. Similarly, changes in a naturally occurring feedstock, such as crude oil, cannot be removed (although changes may sometimes be smoothed out by mixing).[1]

Regulation becomes necessary for a process where, if left to itself, the quality characteristics would drift about and the cause for such drifting is not initially understood or, if understood, cannot be conveniently eliminated.

---

[1]Notice that in both examples involving ambient temperature and the characteristics of a feedstock, feedforward control is in principle also possible. For example, a sensor outside a house can be used to produce compensatory action by the furnace. Because of the robust qualities of feedback control, the latter is normally used: for example, feedback control can cope with disturbances not bargained for, such as the occasional use of a fireplace or the body heat from a large number of guests.

Feedback control

Feedback control may be employed when there exists a known input variable W_t *which can be manipulated* to cancel deviations of the output Y_t from its target value T. Thus in a dyeing process, a too high value of the dye level Y_t in the dyed cloth might be compensated by a reduction in the dye addition rate W_t at the input. Such a scheme where the deviation from target $e_t = Y_t - T$ is being "fed back" to decide the appropriate degree of compensation W_t is an example of regulation by *feedback control.*

The simplest form of feedback control is *proportional* control where the compensation X_t is proportional to the last error e_{t-1} so that the control equation is of the form

$$W_t = k_0 + k_P e_{t-1} \tag{1.1}$$

where the k's appearing in this and subsequent equations are constants. If there is a trend in the disturbance, this equation produces compensation which persistently lags behind the trend so that an additional term using the cumulative sum of the errors is often added to yield a control equation of the form

$$W_t = k_0 + k_P e_{t-1} + k_I \sum_{i=1}^{t-1} e_i \tag{1.2}$$

This is a discrete version of the control action taken by the familiar *proportional-integral* controller or PI controller as it is often called.

Feedforward control

Feedforward control may be employed when knowledge of the value of some measured but uncontrollable variable U_t may be used to partially cancel deviations of the output Y_t from target value T. Thus if it were known that the measured thickness of the incoming cloth U_t affected the output dye level Y_t then U_t could be used instead of $Y_t - T$ to determine the appropriate dye addition rate W_t. Such an example where the input measured variable "cloth thickness" U_t is being "fed forward" to decide the dye addition rate W_t is an example of regulation by feedforward control. For reasons that are mentioned later it may be desirable to use feedforward and feedback control together in a *feedforward-feedback scheme*.

Simple models to illustrate some important issues

To understand how the factors discussed above influence the choice of a control scheme, we consider some specific models. These models are simple but sufficiently realistic to illuminate further discussion. For the time being, we will *not* explicitly consider the cost of making adjustments and the cost of observing the process so that what is said applies strictly to cases where these costs are negligible or where changes in the scheme do not materially change these costs.

Disturbance models

To represent a nonstationary or drifting disturbance such as ambient temperature, which we have seen is usually compensated in the home by a feedback device, we use a model that can represent a nonstationary system as well as a stationary one.

We will define *the disturbance* as the time series that would be followed at the output "if the system were left to itself". A flexible time series model which can represent both a drifting disturbance such is frequently regulated by feedback control as well as uncorrelated random noise about a fixed mean representing a process in a state of control (see, for example, Box and Jenkins [12] and Astrom [5]) is

$$z_{t+1} - z_t = a_{t+1} - \theta a_t \qquad 0 \le \theta \le 1 \tag{1.3}$$

where $\{a_t\}$ is a white noise sequence of "random shocks", that is, a sequence normally and independently distributed about a zero mean. The model may also be written as

$$z_{t+1} = \bar{z}_t + a_{t+1} \tag{1.4}$$

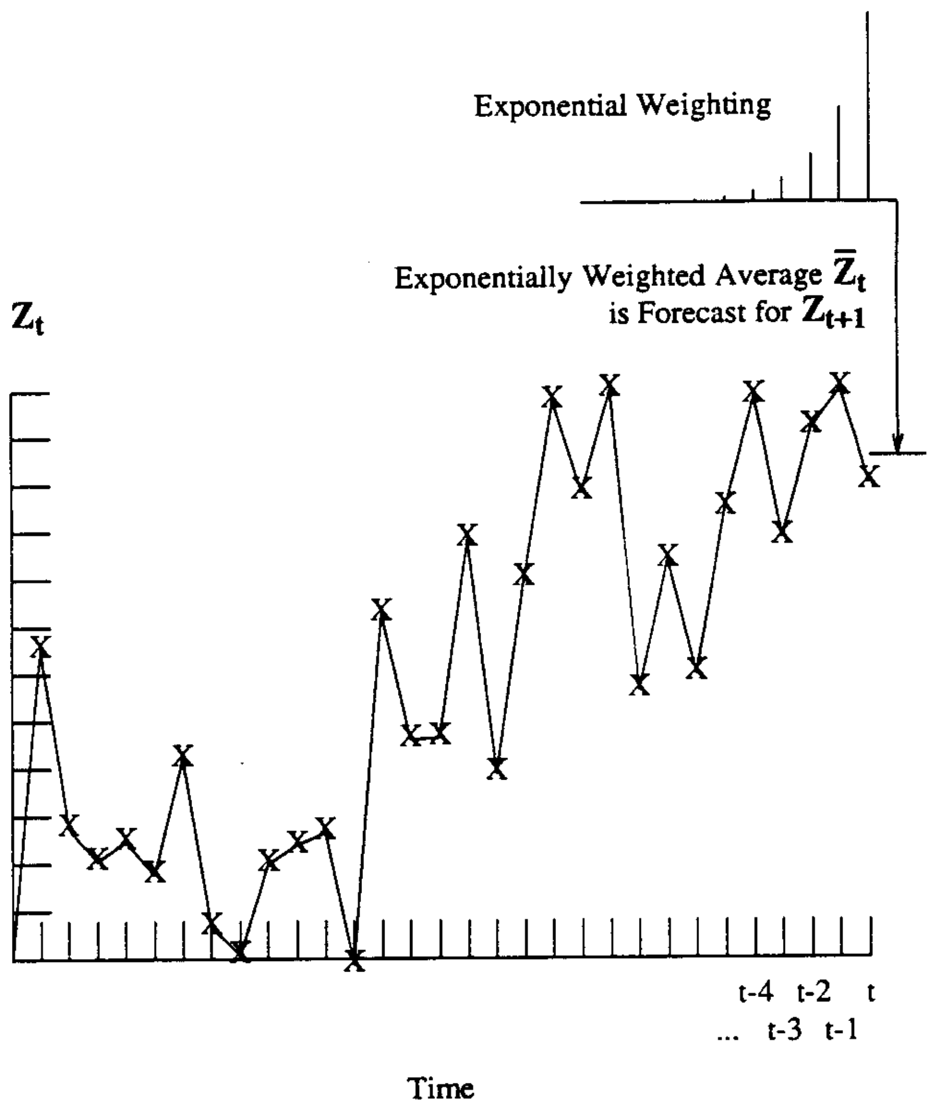
When $0 < \theta < 1$, $\bar{z}_t$ is an *exponentially weighted moving average* (EWMA) of past data

$$\bar{z}_t = w_0 z_t + w_1 z_{t-1} + \cdots \quad \text{with} \quad w_i = (1-\theta)\theta^i \tag{1.5}$$

and θ is sometimes called the *smoothing constant.* This is a *nonstationary* model in which each new observation is generated by an exponentially weighted average of the past observations plus a random shock. Thus, for example, the drifting disturbance shown in figure 1.2 is generated by equation (1.3) with $\theta = 0.5$. The exponential weight function is illustrated by a bar chart above the series. It is easily shown that conditional on knowledge up to time t, the exponentially weighted average $\bar{z}_t$ in (1.4) is the minimum mean square error *forecast* of the next observation z_{t+1}.

When $\theta = 1$, (1.4) becomes the familiar *stationary* model where the errors are IID (independent and identically distributed) about a fixed mean μ

Figure 1.2: A Time Series Generated by the Model of Equation (1.3) With θ = 0.5 (γ = 0.5) Showing Exponential Weight Funtion and One-Step Ahead Forecast



$$z_{t+1} = \mu + a_{t+1} \tag{1.6}$$

for, as θ tends to unity, the weights are spread more and more evenly over the remote past and $\bar{z}_t$ tends to $E(z) = \mu$ where, in the limiting case, each new observation is generated by the mean μ plus a random shock a_t.

When $\theta = 0$, (1.4) becomes the random walk model

$$z_{t+1} = z_t + a_t \tag{1.7}$$

Here all the weight is applied to the last observation so that $\bar{z}_t = z_t$ and each new observation is generated from the previous observation plus a random shock.

As θ approaches unity, (1.4) behaves more and more like the stationary model (1.6) and as θ approaches zero, the process becomes less and less stable and closer and closer to the random walk. Thus, in general, the model can be thought of as an interpolation between the IID model and the random walk model.

An alternate form for (1.4) is

$$z_{t+1} = \bar{z}_1 + \gamma\sum_{i=1}^{t} a_i \tag{1.8}$$

where $\gamma = 1 - \theta$. When $\theta = 0$ ($\gamma = 1$) equation (1.8) gives the random walk (1.7) in the form

$$z_{t+1} = \bar{z}_1 + \sum_{i=1}^{t} a_i \tag{1.9}$$

More elaborate models of this ARIMA class (see, for example, Box and Jenkins [12]) may be used to model more complex situations but this model will suffice for the present discussion.

Models for dynamics and delay

The time series model (1.3) is an example of a *stochastic difference* equation (i.e., a difference equation ''driven'' by the random process a_t). Difference equation models may also be used to characterize inertia and delay in a system. The effect on an output variable X_t of an adjustment in an input variable W_t may be immediate or it may not take effect for b units of time. In addition, it may take time for the full effect to develop due to the inertia of the system. Thus, if we use x_t to represent the change $X_t - X_{t-1}$ at the output for a change in level $w_t = W_t - W_{t-1}$ in the input, then the relationship between x and w when there are b units of delay and g is the gain in the system may be represented as

$$x_t = c_0 + g w_{t-b} \qquad (1.10)$$

where c_0 is a constant. The slightly more general difference equation model

$$x_{t+1} = c_0 + \delta x_t + g(1-\delta) w_{t-b} \qquad 0 < \delta < 1 \qquad (1.11)$$

can take account of inertia in the system as well as delay where, in general, $1 - \delta$ is the proportion of the eventual output response that occurs in the first time interval after the step change in the input.

Figure 1.3(a) shows the response X_t at the output to an input step change W_t of one unit for a delayed system described by equation (1.10) with $g = 2$ and $b = 4$. Figure 1.3(b) shows the response X_t to a unit step input for a system described by equation (1.10) with $\delta = 0.5$, $g = 2$, and $b = 4$.

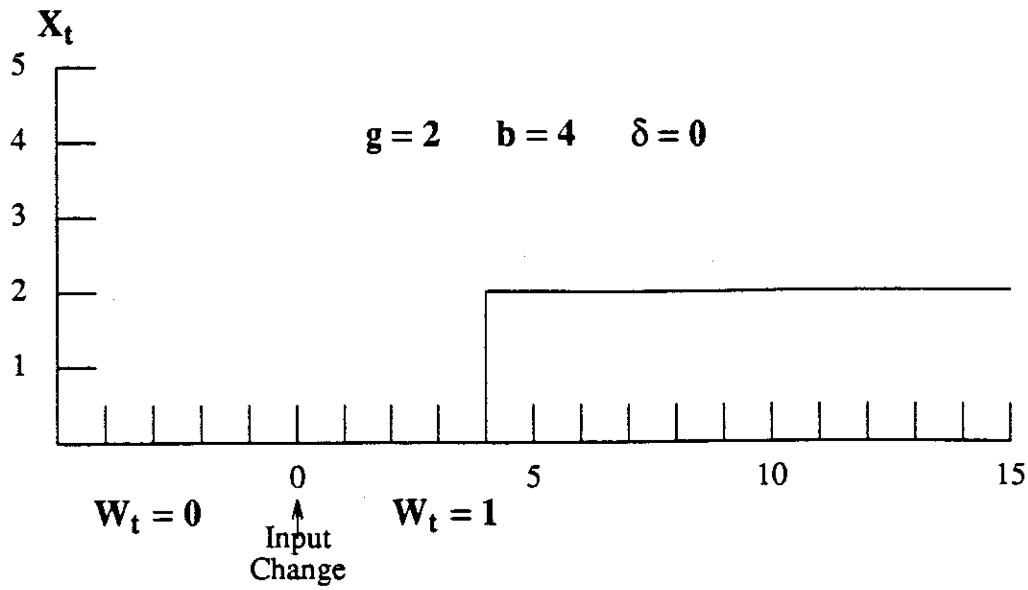


Figure 1.3(a): Output for A One Unit Change in Input at t = 0

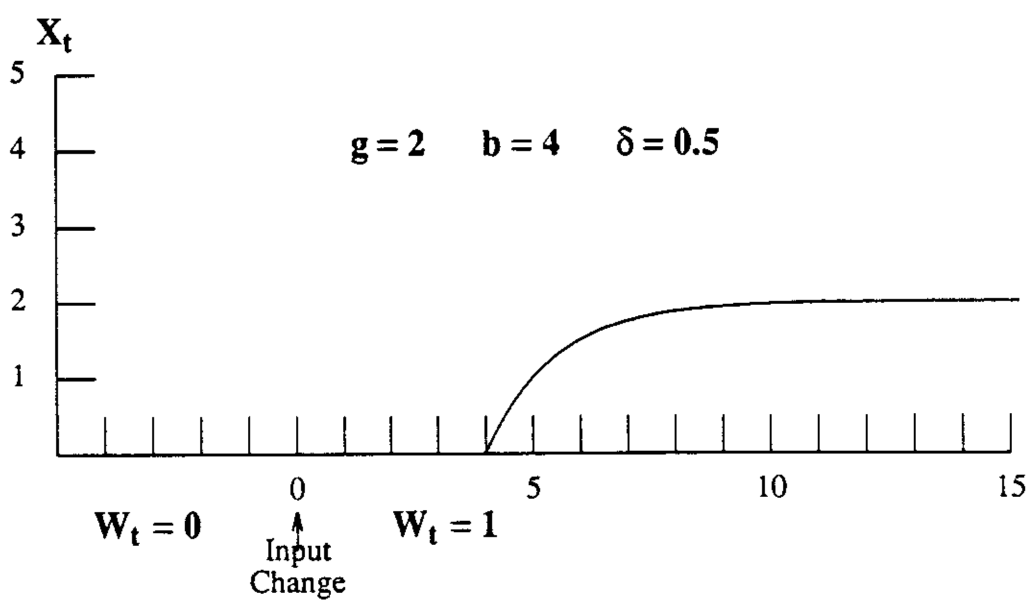


Figure 1.3(b): Output for A One Unit Change in Input at t = 0

A simple example of a feedback control scheme

Consider again the example of feedback control applied to the dyeing process of §1.2. Suppose the dynamics and delays are represented by equation (1.10) with one step of delay ($b = 1$) and the disturbance of the output is to be regulated by equation (1.3). The error at the output is $e_t = Y_t - T = z_t - X_t$ and it is easily shown (Box and Jenkins [12]) that the following controller (control equation) minimizes the mean square error, $E(e_t^2)$.

$$W_t = k_0 + k_P e_t + k_I \sum_{i=1}^{t} e_i \tag{1.12}$$

where the constants k_P and k_I depend on the gain g in the system, the constant of inertia δ, and the time series constant θ in the following way:

$$k_P = \frac{\gamma}{g(1-\delta)} \quad \text{and} \quad k_I = \frac{\gamma}{g} \tag{1.13}$$

where $\gamma = 1 - \theta$. Thus the level W_t at which the dye addition rate should be set at time t depends on the last deviation from target, $e_t = Y_t - T$, and on the sum of the deviations up to time t. This is seen to be the discrete equivalent of the familiar proportional plus integral (PI) controller[2] given earlier in equation (1.2).

It may also be shown that this scheme is equivalent to forecasting the deviation from target, $e_{t+1} = z_{t+1} - X_{t+1}$ from the origin t with an appropriate EWMA of past values of the disturbance z_t and then manipulating W_t so that X_{t+1} cancels this predicted deviation. The error made at the output is then the one step ahead forecast

[2]Controllers where the compensation is a linear combination of a term involving control proportional to the last error and the integral of all previous errors are called proportional plus integral (PI) controllers. A modified form in which an additional term involving the first derivative with respect to time of the error is included is called a PID controller. In this paper, we will be referring to the discrete

error, $z_{t+1} - z_t = a_{t+1} = e_{t+1}$. Notice that minimization of mean square error at the output is equivalent to minimization of *quadratic* off-target cost. Thus, if we could assume the cost at time t of being off-target was proportional to $e_t^2 = (y_t - T)^2$ and if there were no adjustment cost and no monitoring cost, then this feedback scheme would minimize the total cost. Note also that we have assumed that the object is to regulate the process and not to discover the cause of the disturbance.

A simple example of a feedforward control scheme

Suppose now that in a feedforward control scheme cloth thickness U_t is measured as the cloth enters the machine and it takes one unit of time to reach the output so that the contribution at the output of U_t is given by $y_t = c + hU_{t-1}$ where c and h are constants. Suppose further that the dynamics and delays linking the compensating variable W_t and the output are again represented by equation (1.11) with $b = 1$. Then the control action w_t which exactly compensates (at least theoretically) for cloth thickness is readily shown to be (Box and Jenkins [12]):

$$w_t = c_0 + c_1 U_t + c_2 U_{t-1} \qquad (1.14)$$

with

$$c_1 = -\frac{h}{g(1-\delta)} \quad \text{and} \quad c_2 = \frac{h\delta}{g(1-\delta)} \qquad (1.15)$$

Feedback and feedforward control each has advantages and disadvantages which may be of greater or lesser importance in different applications. In particular, feedback control has the advantage that it is much less dependent on exact knowledge

---

analog of these controllers, and in particular, to the PI controller.

of the model but it has the disadvantage that the corrective signal could be greatly delayed by process dynamics. Notice that a mixture of feedback and feedforward control can combine the merits of both. Thus, in the dye level example, it is likely that only partial compensation could be achieved by adjusting for cloth thickness whereas feedback control could compensate for other drifting disturbances. Notice that in this feedforward scheme as in the feedback scheme we have tacitly supposed that no account need be taken of adjustment cost or observation cost.

1.4 Criticisms of SPC and APC

Because the aims and assumptions of SPC and APC are different, it is not unusual to find misunderstandings between those responsible for SPC and those responsible for APC. The SPC practitioners may criticize APC for (as they see it)

i) compensating disturbances rather than removing them,
ii) concealing assignable causes rather than revealing them, and
iii) overcompensating for disturbances resulting in increased variation. In particular, in the parts industry, rumors of improvement when automatic controllers are disconnected are common.

On the other hand, those responsible for APC may argue that, for example,

i) SPC is inefficient for *regulating* a process,
ii) SPC takes no account of the dynamics of the system, and
iii) the supposition that a process can always be brought to a state of control by standardizing conditions is often unrealistic. Feedback or feedforward compensation is often necessary to deal with such disturbances as ambient temperature and quality of feedstock.

The objectives of *process improvement* on the one hand and *process regulation* on the other should, of course, not be regarded as rivals; both are necessary but some tension between these objectives is to be expected. As we shall show, however, these objectives can be pursued concurrently. We now address these specific issues in more detail.

Can we remove disturbances?

When processes are affected by nonstationary variation such as the quality of feedstock or ambient temperature we have two alternatives: we may try to eliminate the cause of variation or we can compensate for it. Thus, a switch to an alternative supplier may result in a more uniform feedstock. However, particularly if a feedstock is a naturally occurring one such as a metallic ore, crude oil or lumber from a forest, achievement of approximate uniformity may not be possible or too expensive to contemplate and we must instead try to compensate for differences by feedback or feedforward control. Again, if we find that the temperature variation in Wisconsin from winter to summer is too extreme, some of us may move to California, but if for other reasons we wish to stay in Wisconsin we should compensate for the cold weather by using a furnace controlled by a thermostat supplying appropriate feedback control.

Are disturbances concealed by automatic control?

As often practiced, feedback and feedforward control conceal compensated disturbances and thus reduce the chance of removing the causes for them. However, this concealing of disturbances does not need to happen. If the equation governing a particular control scheme, such as (1.11) is known and a record of the actions taken

$\{W_t\}$ and the deviations observed $\{e_t\}$ is available, then the nature of the disturbance that has been affecting the system is easily back calculated. By inspection of this reconstructed disturbance the nature of long term fluctuations, suspicious patterns, and unusually large individual deviations is revealed and can be studied. Thus, the pattern of the reconstructed disturbance over time might suggest the identity of a previously unknown explanatory variable. Process improvement could then result. In some cases it might be possible in future to hold this variable fixed; alternatively it might be measured and fed forward. In either case this would result in a major change in the *system*. In addition, individual deviations from the overall trend could indicate assignable causes leading to "find and fix" initiatives exactly parallel to those used with a Shewhart control chart.

Overcompensation?

The equations (1.12) and (1.13) for a simple feedback system and (1.14) and (1.15) for a simple feedforward system were derived under specific assumptions about the nature of the disturbance equation (1.3) and dynamics of the represented system, equation (1.11). If these assumptions are inappropriate the control achieved may be poor and in some instances may be worse than no control. In particular it may take the form of overcompensation.

An extreme example could occur when an improperly tuned feedback controller was applied to a process having a purely random disturbance and already in a state of control. This would result in increasing rather than reducing variation. Notice, however, that such a difficulty would not occur if the system had been properly identified. Consider, for example, the control equation (1.12) applied to a process which is already in a state of control so that the disturbance was in fact white noise corresponding to the model of equation (1.3) with $\theta = 1$. Now, if we substitute the value $\theta = 1$ ($\gamma = 0$) in equation (1.13), we obtain *zero values* for both constants k_P and k_I which determine the amount of proportional and integral control. So the correct feedback control system to be applied in this case is *no control.* But this *is* the control system which would be obtained if we took the trouble to correctly identify the disturbance.

Now the original rationalization for proportional-integral controllers for which equation (1.12) supplies the discrete analog, was not based on system identification but was empirical (see, for example, Mayr [36]). Also, it is known from experience that these controllers tend to be very robust to moderate misspecification. In practice, therefore, a standard PI controller will often simply be hooked up to a system, the characteristics of which have not been identified. These standard controllers can be tuned (i.e., the constants for proportional and integral control can be varied) but this tuning process is not always done very well. Furthermore, in recent years, controllers of the type developed for the process industries have sometimes been transferred to the parts industries without sufficient thought being given to their applicability. This is undoubtedly the basis for stories about *disconnection* of a controller having *reduced* the variance. Formal identification of the disturbance by fitting an appropriate time series model and of the dynamics by carrying out appropriate experiments could avoid these problems but in many cases would be too tedious for routine use. A practical, but less than perfect alternative, is to pay more careful attention to the tuning of the controller using experimental design methods to obtain the optimal setting. For example, if it is decided that the objective function should be the mean square error of Y about the target value, or preferably the logarithm of the mean square error, then response surface experiments (Box and Draper [7]) may be run using the settings of the control constants of the controller as variables. Such a procedure can be used to correctly identify the best settings of the control variables to produce minimum mean square error. In particular, such a procedure will lead to the conclusion that *no* feedback control is needed in appropriate circumstances and so avoid overcompensation. Other criteria, of course, such as immediacy of response can also be optimized in this way.

1.5 Criticism of SPC by APC Practitioners

Shewhart control charts inefficient for regulating a process?

It is the essence of SPC philosophy that you do not react to apparent process changes unless they are established as statistically significant by, for example, a Shewhart chart or cusum chart. For a nonstationary system where the mean is shifting from time to time this can result in sluggish reaction as compared to a feedback system in which adjustments are made at each interval. It can also result in a much larger mean squared deviation from target. Thus, on the assumption of a quadratic loss associated with being off-target, a feedback scheme may produce a much smaller loss than a Shewhart scheme.

However, in the above the tacit assumption is made that it costs nothing to adjust the process. This turns out to be a critical assumption. Suppose instead that to make an adjustment the process must be shut down for a short time, or an expensive tool has to be replaced. Then some cost $\$C_a$ is associated with each adjustment. In that case, as is shown later, minimum cost control schemes look much more like Shewhart control schemes. In particular, if the drifting location is represented by the nonstationary model (1.3), minimum cost is obtained using a chart in which an exponentially weighted moving average (EWMA) of the data is plotted against parallel *action* lines. The EWMA has a smoothing constant θ equal to the parameter θ of the disturbance of equation (1.3). When $\theta = 0$, so that the disturbance is a random walk, then (equation (1.9)) the EWMA is simply the last observation which is plotted between two parallel lines as with the Shewhart chart. In such a scheme, however, the position of these lines is not decided on the basis of 3σ-limits but on the relative costs of adjustment and of being off-target. The point to notice is that with different assumptions about costs, a scheme much like the Shewhart control chart is optimal and any scheme involving straight feedback is more costly. More generally, the relative values of various costs is of great importance in deciding the optimal choice of a control scheme.

Shewhart control does not allow for system dynamics

The inertia of a system is an important determinant of optimal control. Particularly in the process industries, a correction applied to the process at time zero may not be fully effective until some considerable later time. Naive attempts at compensation which ignore this fact can produce very inadequate control. In the parts industries, however, need for such allowance for dynamics is less common and controllers built to deal with dynamics and inappropriately tuned may be ineffective.

State of control model unrealistic?

In the parts industry after a considerable and continuing effort it is often possible to bring and maintain the process in a state of control modelled by

$$y_t = \mathrm{T} + a_t \tag{1.16}$$

where T is the target value and a_t is a white noise process roughly distributed as $N(0, \sigma^2)$. By contrast, in the process industries we are typically dealing with some disturbances which cannot be stabilized and must be compensated for. To represent such disturbances nonstationary time series are usually necessary with optimum control being provided by feedback of feedforward systems.

Costs

The schemes so far discussed are not the only kinds of control schemes that may be appropriate. One important consideration is the relative value of various costs. In particular:

> Off-target cost. Ideally, the output quality characteristic y_t would always be maintained at the target value T. In practice, some deviation must occur and this will have an associated cost.
>
> Minimum mean square error control such as can be achieved with feedback and feedforward schemes in the manner illustrated in section (1.3) will be minimum cost schemes on the assumption that
>
> > a) the cost of being off-target is a quadratic function of the deviation from target.

b) this is the only cost that changes as we change the scheme.

Adjustment cost. For some processes it costs nothing to adjust to a new level. But, for example, when a machine must be stopped and a new tool fitted, an appreciable cost will be associated with adjustment and a minimum cost scheme must take account of both the cost of being off-target as well as the adjustment cost.

Observation cost. Some processes are such that it costs no more to take data frequently than to take them less frequently. In others, particularly where expensive analyses must be made, there is a cost associated with frequency of data taking.

Thus, minimum cost schemes might in general take into account off-target costs, adjustment cost, and observation cost.

1.6 Summary and Preview

Some of the factors which determine the appropriateness of a control scheme are:

a) the purpose of the scheme
b) the nature of the disturbances which affect the output system
c) the dynamics and delays in adjusting the system
d) whether or not there is a particular input W_t that can be manipulated to partially compensate disturbances at the output in a feedback control scheme
e) whether or not the disturbances of the output is partly determined by an input U_t which can be measured and used in a feedforward control scheme
f) the cost of being off-target
g) the cost of making adjustments
h) the cost of observing the process

The effect of these factors in the choice of control schemes is considered in this thesis. In particular this makes it possible to address the following questions

1) How often should data be collected from the process to see how close the process is to a desired target?
2) When should the process be adjusted so as to maintain it near the target?

3) What size of adjustment should be made?

Chapter two considers the above questions when there are fixed costs of monitoring and adjustment and when adjustments are made without delay or dynamics. Chapter three considers these questions when there are system dynamics and when both the mean squared deviation from target and the variance of changes in an input variable are important. Chapter four summarizes the results and introduces some possible extensions to the research.

Bibliography

1. Abraham, B. and G. E. P. Box. 1979. Sampling interval and feedback control. *Technometrics*, 21 (no. 1): 1-8.

2. Adams, B. M. and W. H. Woodall. An analysis of Taguchi's on-line process control procedure under a random walk model. *Technometrics*, to appear.

3. Alwan, L. and H. V. Roberts. 1986. Time series modeling for statistical process control. *Proceedings of the Bus. and Econ. Stat. Section, Amer. Stat. Assoc.*, 315-320.

4. Åström, K. J and B. Wittenmark. 1984. *Computer Controlled Systems: Theory and Design*. Prentice-Hall, New Jersey.

5. Åström, K. J. 1970. *Introduction to Stochastic Control*. Mathematics in Science and Engineering Series, 70. Academic Press.

6. Barnard, G. A. 1959. Control charts and stochastic processes. *J. Royal Stat. Soc., Series B*, 21 (no. 2): 239-271.

7. Box, G. E. P. and N. R. Draper. 1987. *Empirical Model-Building and Response Surfaces*, Wiley.

8. Box, G. E. P. 1980. Sampling and Bayes' inference in scientific modelling and robustness. *J. Royal Stat. Soc., Series A*, 143 (part 4): 383-430.

9. Box, G. E. P. and G. M. Jenkins. 1963. Further contributions to adaptive optimization and control: simultaneous estimation of dynamics: non-zero costs. *Bulletin of the International Statistical Institute*, 34^{th} Session, Ottawa.

10. Box, G. E. P. and G. M. Jenkins. 1965. Mathematical models for adaptive control and optimisation. *A. I. Chem. E. – I. Chem. E. Symposium Series*, 4: 61-68.

11. Box, G. E. P. and G. M. Jenkins. 1968. Some recent advances in forecasting and control. *J. Royal Stat. Soc., Series C (Applied Statistics)*, 17 (no. 2): 91-109.

12. Box, G. E. P. and G. M. Jenkins. 1976. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.

13. Box, G. E. P., G. M. Jenkins and J. F. MacGregor. 1974. Some recent advances in forecasting and control, part II. *J. Royal Stat. Soc., Series C (Applied Statistics)*, 23 (no. 2): 158-179.

14. Box, G. E. P. and J. G. Ramírez. 1989. Personal communication.

15. Clarke, D. W. and P. J. Gawthorpe. 1975. Self-tuning controller. *Proceedings of the Institution of Electrical Engineers*, 122 (no. 9): 929-934.

16. Clarke, D. W. and R. Hastings-James. 1971. Design of digital controllers for randomly disturbed systems. *Proceedings of the Institution of Electrical Engineers* 118, (no. 10): 1503-6.

17. Crowder, S. V. 1986. *Kalman Filtering and Statistical Process Control*. Ph.D. diss., Iowa State University.

18. Crowder, S. V. 1987. A simple method for studying run-length distributions of exponentially weighted moving average charts. *Technometrics*, 29 (no.4): 401-408.

19. Crowder, S. V. 1987. Average run lengths of exponentially weighted moving average charts. *Journal of Quality Technology*, 19 (no. 3): 161-164.

20. Deming, W. E. 1986. *Out of the Crisis*, MIT, Center for Advanced Engineering Study, Cambridge, Mass.

21. Ewan, W. D. and K. W. Kemp. 1960. Sampling inspection of continuous processes with no autocorrelation between successive results. *Biometrika*, 47 (3 and 4): 363.

22. Hunter, J. S. 1986. The exponentially weighted moving average. *Journal of Quality Technology*, 18 (no. 4): 203-210.

23. Imai, M. 1986. *Kaizen: The Key to Japan's Competitive Success*. Random House.

24. Jowett, G. H. 1955. The comparison of means of sets of observations from sections of independent stochastic series. *J. Royal Stat. Soc., Series B*, 17 (no. 2): 208-227.

25. Jowett, G. H. 1957. Statistical analysis using local properties of smoothly heteromorphic stochastic series. *Biometrika*, 44: 454-463

26. Kartha, P. 1973. Some Statistical and Economic Problems in Quality Control. Ph.D. diss., University of Wisconsin, Madison.

27. Kramer, T. 1986. An extension of the Box-Jenkins approach to process regulation. Speech given at the 3rd Quality and Productivity Research Conference in Oakland, Michigan.

28. Lorenzen, T. J. and L. C. Vance. 1986. The economic design of control chart: a unified approach. *Technometrics*, 28 (no. 1): 3-10.

29. Lucas, J. M. 1976. The design and use of V-mask control schemes. *Journal of Quality Technology*, 8 (no. 1): 1-12.

30. Lucas, J. M. and R. B. Crosier. 1982. Fast initial response for cusum quality-control schemes: give your cusum a head start. *Technometrics*, 24 (no. 3): 199-206.

31. MacGregor, J. F. 1972. Topics in the Control of Linear Processes Subject to Stochastic Disturbances. Ph.D. diss., University of Wisconsin, Madison.

32. MacGregor, J. F. 1976. Optimal choice of the sampling interval for discrete process control. *Technometrics*, 18 (no. 2): 151-160.

33. MacGregor, J. F. 1987. Interfaces between process control and online statistical process control. *Computing and Systems Technology Division Communications*, 10 (no. 2): 9-20.

34. MacGregor, J. F., T. J. Harris and J. D. Wright. 1984. Duality between control of processes subject to randomly occurring deterministic disturbances and ARIMA stochastic disturbances. *Technometrics*, 26 (no. 2): 389-397.

35. MacGregor, J. F. and P. W. Tidwell. 1977. Correspondence: Discrete stochastic control with input constraints. *Proceedings of the Institution of Electrical Engineers*, 124 (no. 8): 732-734.

36. Mayr, O. 1970. *The Origins of Feedback Control.* M.I.T. Press, Cambridge, Massachusetts.

37. Roberts, S. W. 1959. Control chart test based on geometric moving averages. *Technometrics*, 1 (no. 3): 239-50.

38. Roberts, S. W. 1966. A comparison of some control chart procedures. *Technometrics*, 8 (no. 3): 411-430.

39. Shewhart, W. A. 1931. *Economic Control of Quality of Manufactured Product.* D. Van Nostrand Company, Inc.

40. Taguchi, G. 1986. *Introduction to Quality Engineering: Designing Quality Into Products and Processes.* Asian Productivity Organization.

41. Taguchi, G. 1984. Specification value and quality control, part 6. *The International QC Forum.*

42. Taguchi, G. 1984. Specification value and quality control, part 7. *The International QC Forum.*