

Algorithmic Statistical Process Control: Concepts and an Application

Scott A. Vander Wiel

William T. Tucker, Frederick W. Faltin, and Necip Doganaksoy

AT&T Bell Laboratories
Murray Hill, NJ 07974

GE Corporate Research and Development Center
Schenectady, NY 12301

The goal of algorithmic statistical process control is to reduce predictable quality variations using feedback and feedforward techniques and then monitor the complete system to detect and remove unexpected root causes of variation. This methodology seeks to exploit the strengths of both automatic control and statistical process control (SPC), two fields that have developed in relative isolation from one another. Recent experience with the control and monitoring of intrinsic viscosity from a particular General Electric polymerization process has led to a better understanding of how SPC and feedback control can be united into a single system. Building on past work by MacGregor, Box, Astrom, and others, the article covers the application from statistical identification and modeling to implementing feedback control and final SPC monitoring. Operational and technical issues that arose are examined, and a general approach is outlined.

KEY WORDS: CUSUM monitoring; Engineering control; Minimum MSE control; Regulation; Statistical monitoring.

Automatic process control and traditional statistical process control (SPC; arguably a misnomer for statistical process monitoring) have developed in relative isolation from one another. Yet both of these distinct, even divergent, methodologies have scored significant successes in the drive for quality improvement. In recent years several authors have sought to bring both approaches into better perspective by delineating more clearly the contexts to which each is best suited.

Even these efforts, however, have tended to assume (at least implicitly) that automatic control and SPC are two sides of an either-or proposition. Recent experience suggests, to the contrary, that substantial improvements to product quality are often best attainable through an integration of techniques from both methodologies, whereby one exploits the benefits of both. Algorithmic statistical process control (ASPC) is our term for an integrated approach to quality improvement—an approach that realizes quality gains through appropriate process adjustment (i.e., process control) and through elimination of root causes of variability signaled by statistical process monitors.

In Section 1, the ASPC concept is described in more detail and some relevant past work is reviewed. In Section 2, an application of ASPC to a polymerization process is described. In Section 3, guidelines to ASPC implementation are presented. In Section 4, a summary is given. Generalizations, detailed application guidelines, and research issues relating to

ASPC are covered in a companion paper by Tucker, Faltin, and Vander Wiel (1991), which gives additional references that pertain to the broader context of ASPC.

1. ALGORITHMIC STATISTICAL PROCESS CONTROL

1.1 SPC and Automatic Control: A Comparison

By SPC we mean a collection of techniques found especially useful in improving product quality by helping an analyst locate and remove root causes of quality variation. Statistical monitoring charts, including Shewhart charts, are emphasized here, but other techniques are envisioned as well. By automatic control we mean a collection of techniques for devising algorithms to manipulate the adjustable variables of a process to achieve the desired process behavior (e.g., output close to a target value).

A comparison of SPC and automatic control reveals the different orientations of the fields in three significant areas:

1. *Philosophy:* Both fields seek to reduce deviations of some characteristic from a target value. In SPC, however, this is accomplished by monitoring a process so as to detect and remove root causes of variability. On the other hand, automatic control seeks to counteract the effects of root causes through continual process adjustment.

2. *Application Context*: Statistical monitoring charts are ordinarily appropriate when it is reasonable to expect successive process measurements to be well modeled as iid and one is concerned with detecting departures from such an ideal. By contrast, automatic control is ineffective on (even harmful to) an iid process. It is most effective in the context of a continually wandering process—for example, a process that could be well modeled by an autoregressive moving average time series.

3. *Traditional Development*: Automatic control is most often used tactically. For example, feedback controllers are typically commissioned to maintain the setpoints of important process parameters. SPC, however, is often allowed a strategic role. Control charts are kept on important quality characteristics, allowing SPC to have a direct impact on the quality of the process output. Of course, there have been occasional exceptions where automatic control has taken a more strategic role. See Astrom (1970, chap. 6), Box and Jenkins (1976, chap. 10), MacGregor and Tidwell (1980), and Roffel, MacGregor, and Hoffman (1989).

Thus, in usual applications, the *ultimate effect* of applying SPC has been to fundamentally *improve* a process by removing sources of variations, whereas that of automatic control has been to optimally *adjust* an existing process.

1.2 Combining SPC and Automatic Control

Tools from either field, used individually, can lead to better product quality—automatic control through process optimization, SPC through process improvement. Ideas from both fields can be used together, however, securing both optimization and improvement. This is the concept of ASPC. Such dual implementation is especially natural in the continuous-process industries, where quality improvement is a key to profitability and where one is apt to find measurements that are correlated over time so that the process appears to wander. Autocorrelation is especially common in these industries because disturbances such as changes in raw materials tend to have immediate, as well as lasting, effects, due, for example, to mixing tanks, recovery and reuse of chemical agents, and the slow response of many systems to control actions. Autocorrelation is not necessarily bad. It does, however, mean that the process is somewhat predictable, and this suggests the possibility of compensation. It is interesting that raw-material mixing, which can cause autocorrelation, is often performed to make incoming stock more homogeneous. In this case the cause of the autocorrelation probably should not be eliminated, and when the raw material is, for example, crude oil, it is often not feasible to eliminate the material causes of variation.

In brief, ASPC reduces predictable *quality* variations using feedback and feedforward techniques and then monitors the *complete* system to detect and remove unexpected root causes of variation. Obviously the philosophy of making compensatory adjustments is quite different from that currently popular in SPC. MacGregor (1988) stated that automatic control strategies "might be thought of as band-aids that hide things that should be improved at the process level" (p. 31). ASPC does not ignore quality improvements attainable through eliminating root causes of variability, however; rather, it advocates that compensatory adjustments be applied *in conjunction with*, rather than *in competition with*, traditional SPC.

1.3 Relevant Past Work

There is a large literature concerning control of stochastic systems that is relevant to the algorithmic part of ASPC. Classic references are Astrom (1970) and Box and Jenkins (1970, 1976). A quite separate body of literature is that of SPC. Early developments in quality monitoring are due to Shewhart (1931), and an often cited textbook is that of Duncan (1986).

One result of the present U.S. emphasis on quality improvement has been a marked interest in a possible marriage of SPC to automatic control. In the discussion following an early article by Box and Jenkins (1962), J. H. Westcott (1962) remarked "Speaking as a control engineer, I . . . welcome this flirtation between control engineering and statistics. I doubt, however, whether they can yet be said to be 'going steady.'"

The surge of activity relating SPC to automatic control in the early 60s was not sustained, and although the general idea is not new, the specific suggestion to superimpose statistical process monitoring on a closed-loop system appears to be quite recent and certainly opens up new lines of research in the area of quality improvement. MacGregor (1988) was apparently the first to suggest to the SPC community that SPC charts be used to monitor the performance of a controlled system. He only briefly reviewed basic concepts from stochastic control as well as so called "on line" SPC methods, however, pointing out similarities and overlap and citing several reasons for the lack of interface between the fields. (See also MacGregor and Harris 1990). Box and Kramer (1992) gave overview descriptions of both fields, delineating their similarities and differences.

One interesting point of overlap occurs when (a) control actions have their full effect on process output in the immediately succeeding period, (b) process noise is modeled as a first-order integrated moving average, (c) a fixed cost is associated with taking any nonzero control action, and (d) additional costs are assessed in proportion to the squared deviation of the output from target. In this case, the minimum

expected cost strategy is to adjust the process at the signal of a (control modified) exponentially weighted moving average (EWMA) chart by an amount that will offset the EWMA predictor. This so-called machine-tool problem is one of two discussed by Box and Jenkins (1963). The problem demonstrates a point of overlap where an optimal control scheme can be implemented with a standard SPC chart. The aim of ASPC, however, is to use these two methodologies to perform the separate functions of control and monitoring. Thus, from the perspective of the machine-tool problem, one would still have SPC charts in place operating in conjunction with the SPC-like control rule. This emphasis on integration, as applied to quality improvement, appears to be quite new to both SPC and engineering control.

A general result encompassing the machine-tool problem as a special case was proved by Bather (1963) in a foundational article linking optimal control and SPC. An even earlier article in this area is that of Barnard (1959), who was among the first to suggest that the usual practice of making process adjustments at the signals of control charts, although simple, can be improved on under a reasonable model for a wandering industrial process. He linked the problem of optimal control to that of estimating the current process mean and suggested that it may be useful to view the primary function of a control chart as providing an estimate of that mean.

More recently, exact calculations of the control limits for the machine-tool problem were given by Crowder (1992). Extensions to this problem have been studied by Adams and Woodall (1989), Jensen (1989), Kramer (1989), and Vander Wiel (1991). Taguchi (1985) described a method for on-line process control under similar conditions to those of the machine-tool problem.

2. A BATCH POLYMERIZATION EXAMPLE

2.1 Process Description

In a commercial scale process that produces a polymer resin used in many familiar consumer products, polymerization occurs in five batch reaction lines, consisting of three standard reactors in one group and two larger reactors in another group, running in parallel and sharing common raw materials. The reactors run asynchronously, with each batch cycle consisting of the following steps: Charge the reactor; run the reaction; empty the reactor contents into a holding tank for subsequent processing; engage cleaning procedures; and recharge to begin the next batch. Intrinsic viscosity, a key quality characteristic of the polymer, is measured on completing each batch. The measurement process introduces modest analytical error. Turnaround time is such that the viscosity

measurement from the most recent batch produced in a given reactor is usually, but not always, available when that reactor is prepared for a new batch.

The objective of this study is to minimize viscosity variation about a target level of 100 (coded) viscosity units. To this end, process measurements from previous batches can be used to adjust the amount of catalyst added to future batches. Normally, 50 (coded) gallons are needed to attain the target viscosity. Traditionally, the exact amount of catalyst added to a particular batch was determined by experienced operators, based on their observations and good judgment, together with some general guidelines provided by the responsible manufacturing engineer. More recently, however, a Shewhart-chart approach has been used.

Under the Shewhart scheme, stepped adjustments to catalyst were made in response to out-of-control signals from an X chart (a plot of viscosity vs. batch number) kept for each reactor. The amount of catalyst added was left unchanged until an out-of-control condition was signaled by the chart, at which time the catalyst level was increased or decreased by one gallon depending on whether the viscosity was below or above the target value of 100. One-gallon increments to the catalyst level were loosely based on a rule of thumb stating that every additional gallon of catalyst added should increase viscosity by .75 units.

On occasion, the Shewhart adjustment scheme was preempted by the operators. This happened, for example, when converting production from standard viscosity (a target of 100 units) to low or high viscosity (97 or 103 units) or when a very extreme viscosity measurement was reported by the quality laboratory.

Discussions with the manufacturing engineer pointed to two types of correlation likely to be present among viscosity measurements—autocorrelation over time on a given reactor and cross-correlation among reactors. These correlations stem from several factors. First, mechanical considerations prevent a reactor from being fully emptied between batches. The presence of leftover "heel" material makes it reasonable to expect intrinsic viscosity to exhibit serial correlation on each reactor. Moreover, economic and environmental considerations necessitate the recovery and reuse of certain reaction components (e.g., solvents). Introducing recycled raw materials into the storage tanks has the potential to induce a measure of long-term process memory. In addition, the fact that all reactors draw from large common sources of raw materials suggests that cross-correlations will exist between batches produced by different reactors and autocorrelations will exist among batches produced by a common reactor. Finally, empirical evidence suggests that there are almost always hidden factors such as environmental conditions and main-

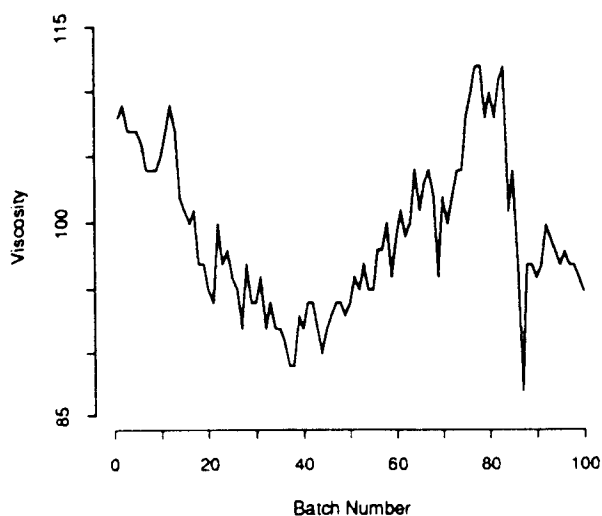


Figure 1. Viscosity Versus Batch Number With Catalyst Held Constant. Viscosity measurements are adjusted to show variation in viscosity when catalyst is held fixed at 50 gallons. A sharp drop in viscosity begins with period 84.

tenance schedules, which further promote such correlations.

Both autocorrelation and cross-correlation are easily observed in plots of viscosity data. For example, autocorrelation is seen in Figure 1, a time plot of polymer viscosity from one reactor corrected to show what would have been obtained had the catalyst addition been held constant at the nominal level of 50 gallons. Construction of Figure 1 is explained in Section 2.4.

2.2 Model Formulation

The overview of the polymerization process given in Section 2.1 suggests a tentative model for the measured viscosity of batch t :

$$y_t = \beta u_{t-1} + \varepsilon_t + f_t, \quad (1)$$

where y_t = observed viscosity deviation from 100 units; u_{t-1} = catalyst deviation from nominal (50 gallons) for batch t , determined after completing batch $t-1$; ε_t = autocorrelated intrinsic error and $= \rho\varepsilon_{t-1} + a_t$, where $a_t \sim$ independent $N(0, \sigma_a^2)$; and f_t = batch-to-batch extrinsic error—for example, measurement error—distributed independent $N(0, \sigma_f^2)$; and independent of $\{a_t\}$.

The model describes generation of viscosity measurements on a given reactor. To simplify modeling effort, cross-correlations among the reactors were initially ignored. In words, the model states that viscosity = catalyst effect + autocorrelated intrinsic error + uncorrelated extrinsic error. The model gives consideration to three factors mentioned in the process description: (1) Autocorrelation of intrinsic viscosity, (2) modest measurement error, and (3) a linear relationship between the amount of catalyst added

and viscosity (as suggested by the rule of thumb that also suggests .75 as the value for β).

The autocorrelated term, ε_t , captures the wandering nature of the process. In this application it can be modeled using one of the most elementary time series models, a stationary first-order autoregressive sequence. The correlation structure of ε_t is characterized by the parameter ρ ($|\rho| < 1$). In particular, $\text{corr}(\varepsilon_s, \varepsilon_t) = \rho^{|t-s|}$.

Appendix A shows that Model (1) can be reparameterized as

$$y_t = \beta u_{t-1} + \frac{(1 - \theta B)}{(1 - \rho B)} e_t, \quad (2)$$

where B is the backshift operator ($Bx_t = x_{t-1}$) and $e_t \sim$ independent $N(0, \sigma_e^2)$. This representation is that of a simple autoregressive moving average transfer function (ARMAX) model, which can easily be fit by many standard time series packages including, for example, S-PLUS (Statistical Sciences Inc. 1990), Matlab (1987), SAS/ETS (SAS Institute Inc. 1984), and BMDP (1983). Relationships between the parameters in Representations (1) and (2) are also given in Appendix A.

Model (2) was the first and simplest model proposed to describe fluctuations in viscosity. Residual analyses from this model gave no reason to contemplate a higher order time series model. But since process data is readily available on many variables in addition to viscosity (y_t) and catalyst (u_{t-1}), other models were fit incorporating various supplementary measurements. Some models were designed to capitalize on between-reactor correlation; others used more of the available information from a single reactor such as temperatures at various times throughout the reaction and levels of several chemicals recovered in solution after the reaction. With the data at hand, however, none of the extended models was able to produce a substantial increase in explanatory power over the simple Model (1), so this is the only one considered in the remainder of this article.

2.3 Model Fitting

Viscosity and catalyst measurements were obtained from all five reactors on a total of approximately 450 batches of product dating from May 21 to June 6, 1988. For the reactor considered in this article the specified nominal catalyst level of 50 gallons was not available from engineering considerations but rather represents average catalyst additions during time periods for which the target viscosity was 100 units. Since the average attained viscosity during these periods was very near the target value, a nominal catalyst level of 50 gallons allowed Equation (2) to give an adequate description of the polymer production process even though no intercept term is included.

Parameters in Representation (2) were estimated using the *conditional least squares* algorithm of the ARIMA procedure in SAS/ETS (SAS 1984). The parameter estimates and standard errors (in parentheses) are

$$\hat{\beta} = 1.087 \quad \hat{\rho} = .859 \quad \hat{\theta} = .164 \quad \hat{\sigma}_e = 2.798. \\ (.223) \quad (.071) \quad (.122) \quad (.192).$$

From the relationships (given in Appendix A) between the sets of parameters in Representations (1) and (2), one can obtain estimates and approximate standard errors (via second-order Taylor series expansions) for the parameters in the original process model (1). These are as follows:

$$\hat{\beta} = 1.087 \quad \hat{\rho} = .859 \quad \hat{\sigma}_u = 2.334 \quad \hat{\sigma}_f = 1.222. \\ (.223) \quad (.071) \quad (.385) \quad (.377)$$

The estimate $\hat{\sigma}_f = 1.222$ agrees favorably with an independent study of the laboratory measurement error, which gave $\hat{\sigma}_f = 1.2$. It is also interesting to compare the estimate of β with the rule of thumb that suggested a value of .75. If the "true value" is near 1.1, then the rule of thumb is anticonservative; that is, it would lead to overcontrol.

After the initial model formulation and analysis, several data sets from a six-month period in 1988 were analyzed to validate and refine the initial parameter estimates. Of course, estimates varied somewhat over the different sets of data and for different reactors. Model (1), however, gave a reasonable description of the process over a long period of time. Compromise parameter estimates of β , ρ , and θ used to specify the minimum mean squared error (MSE) control rule are $\hat{\beta} = 1.5$, $\hat{\rho} = .8$, and $\hat{\theta} = .22$. The compromise estimates are weighted averages from the various data sets. Although the compromise $\hat{\beta}$ is larger than the previous estimate, it is more consistent with estimates obtained from the follow-up data. Furthermore, when comparing the various estimates, no significant differences were found (at the .05 significance level). Most important though, Section 2.4 shows the adequacy of the estimated model for the purpose of control.

2.4 Process Control

Optimal Control. An alternative to the Shewhart chart control scheme (Sec. 2.1), which had been used on this process, is to optimize control actions with respect to a given criterion. Optimal schemes often account for costs associated with deviations of the controlled variable from its target value as well as costs incurred because of the control actions. In this study a moderate change in the amount of catalyst added represents negligible cost (or savings) when compared to the cost incurred by a batch of off-target material. For this reason, the mean squared deviation

of viscosity from its target value (MSE) seems to be a reasonable criterion by which to compare control rules.

Hence a reasonable goal is to specify the amount of catalyst to be added in period t in such a way that MSE (y_t) is minimized. Recall from Section 2.1 that, due to laboratory processing time, the viscosity measurement from the most recent batch on a given reactor is not always available when the next catalyst decision is being made. Measurements from two and more periods back, however, are virtually always available. When the most recent measurement is available, we say that the next catalyst decision is a one-step decision. Otherwise it is a two-step decision.

Minimum MSE Feedback Control (no laboratory delay). In this section, a minimum MSE control rule is given for the process model (2) under the assumption that the most recent viscosity measurement is available; that is, a function g is specified so that, if one sets $u_{t-1} = g(u_{t-2}, u_{t-3}, \dots, y_{t-1}, y_{t-2}, \dots)$, then the resulting y_t will have the smallest possible MSE.

The standard textbooks by Astrom (1970) and Box and Jenkins (1970, 1976) gave solutions to the minimum MSE feedback problem for general ARMAX systems. If laboratory delays were never encountered, then the following so-called pure one-step adjustment rule would be the minimum MSE rule for Model (2) or equivalently for the original Model (1):

$$u_{t-1} = \rho u_{t-2} - \left[\frac{\rho - \theta}{\beta} \right] y_{t-1}. \quad (3)$$

The adjustment is a simple linear function of the catalyst added to the previous batch and the previous viscosity measurement. Substituting (3) into (2) reveals that this control action results in the closed-loop process $y_t = e_t$.

Using the compromise parameter estimates the pure one-step rule is $u_{t-1} = .8u_{t-2} - .4y_{t-1}$. That is, the catalyst addition should deviate from its nominal value by 80% of the last deviation minus .4 times the last viscosity deviation. This control algorithm has at least two advantages over the Shewhart-chart approach described in Section 2.1. It is at least as simple, and it should reduce viscosity variability even further than the former approach.

Minimum MSE Feedback Control When Measurements Are Delayed. The minimum MSE rule (3) may be obtained by setting the best output predictor to 0 and solving for the control action. This property is known as certainty equivalence. Bar-Shalom and Tse (1974) gave necessary and sufficient conditions for certainty equivalence to hold for a large class of control systems. In the present context,

certainty equivalence can be used to prescribe the minimum MSE control policy when measurements are possibly delayed.

If a measurement is delayed, (3) cannot be applied. Nevertheless, it is possible to minimize the output MSE with respect to the available data as outlined in the following. (See Appendix B for further details.) Let $\hat{y}_{t|t-i}$ ($i = 1, 2$) be the i -step minimum MSE forecast of y_t . Then

$$\hat{y}_{t|t-i} = \beta u_{t-1} + \hat{N}_{t|t-i}, \quad (4)$$

where $\hat{N}_{t|t-i}$ is the i -step minimum MSE forecast of the model noise term

$$N_t = \frac{(1 - \theta B)}{(1 - \rho B)} e_t = y_t - \beta u_{t-1}.$$

By certainty equivalence the i -step minimum MSE feedback rule is found by setting $\hat{y}_{t|t-i} = 0$, resulting in

$$u_{t-1} = -\hat{N}_{t|t-i}/\beta. \quad (5)$$

When y_{t-1} becomes available, the update equation for $\hat{N}_{t|t-1}$ is

$$\hat{N}_{t|t-1} = \theta \hat{N}_{t-1|t-2} + (\rho - \theta)(y_{t-1} - \beta u_{t-2}). \quad (6)$$

When needed (for two-step adjustments), $\hat{N}_{t|t-2}$ can be computed from

$$\hat{N}_{t|t-2} = \rho \hat{N}_{t-1|t-2}. \quad (7)$$

In the absence of laboratory delays, this procedure is equivalent to using (3). Importantly, adjusting (as previously) to negate the forecast of N_t minimizes the output MSE in period t regardless of what control policy was used in previous periods. For example, suppose the process has been adjusted using a two-step minimum MSE rule through period $t - 1$, but beginning in period t laboratory measurements are no longer delayed. Simply switching to the one-step rule (3) will not minimize the output MSE in period t ; the derivation of (3) tacitly assumes that the rule will be used in each period. Using the intermediate quantity $\hat{N}_{t|t-1}$ avoids this difficulty.

Comparison of Control Strategies. If catalyst additions have an effect on only the immediately succeeding viscosity measurement with linear gain β , then historical data can be used to determine the viscosity deviation from target y_t^* that would have been obtained had the catalyst deviation from nominal been held at some value u_{t-1}^* rather than its actual value u_{t-1} . The calculation is simply

$$y_t^* = y_t + \beta(u_{t-1}^* - u_{t-1}). \quad (8)$$

Figure 1 was constructed using (8) with $u_{t-1}^* = 0$ on the assumption that the estimated gain $\hat{\beta} = 1.5$ is correct. Similarly, Figure 2 displays what would have resulted from applying the estimated pure one-step

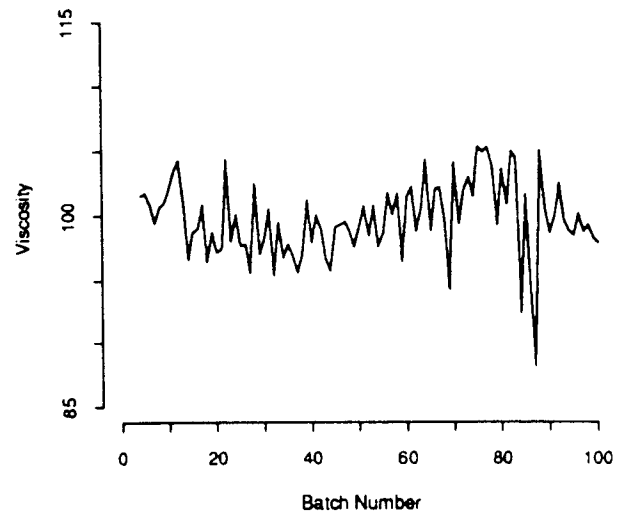


Figure 2. Viscosity Versus Batch Number Under One-Step Control. Viscosity measurements are adjusted to show what would have resulted had catalyst been set using the estimated one-step minimum MSE policy. The sharp drop in Figure 1 beginning with period 84 now appears as a filtered change.

minimum MSE policy (5) [or (3)] over the same period. Table 1 gives a numerical comparison of root mean squared errors (RMSE's) obtained under four types of control—no control, actual control by a skilled operator, estimated pure one-step control, and estimated pure two-step control. The row labeled 1988 refers to the data set of Figures 1 and 2 collected at the early stages of this study. The row labeled 1989 refers to a data set collected after implementing algorithmic control. Estimated control rules were constructed using the compromise parameter estimates.

The 1988 row of Table 1 shows that before implementing the feedback algorithm operators were able to reduce viscosity deviations from target by about 30% from what would have resulted had catalyst been fixed at its nominal level. Under the pure one-step rule, however, one could expect nearly a 30% further reduction. Since occasionally viscosity measurements are not available from the immediately preceding batch, the one-step results may be slightly optimistic. The RMSE for a pure two-step rule gives a worst-case estimate of what to expect from the control algorithm of (5)–(7). Note that even this is

Table 1. Root Mean Squared Errors Under Four Types of Control

Data set	Sample size	None	Actual	Pure one-step	Pure two-step
1988	106	6.00	4.15	2.96	3.66
1989	213	8.62	3.48	3.37	6.59

NOTE: Actual for 1988 was under operator control. Actual for 1989 was under estimated minimum MSE feedback.

better than the actual RMSE achieved by skilled operators.

The results of the first (1988) row of Table 1 indicate that the feedback algorithm holds potential at reducing variability. Further studies validated this conclusion. The ultimate proof of the scheme's effectiveness, however, was only seen after it was installed.

The second row (1989) of Table 1 shows a comparison of RMSE's under various types of control several months after the algorithmic adjustment policy was implemented. Although the underlying variability appears to have increased, the adjustment scheme has nevertheless substantially reduced viscosity deviations from what would occur if no catalyst adjustments had been made. The actual RMSE is also smaller than it was before installing the adjustment algorithm. In this case the actual RMSE differs slightly from the one-step RMSE only because in some periods it was necessary to use the two-step rule due to delays in the analytical laboratory.

Assuming the continued appropriateness of (8) with $\beta \approx 1.5$ the 1989 follow-up data set confirms the effectiveness of the adjustment scheme several months after the model fitting and design stage of this study.

2.5 Putting the SPC in ASPC

Section 2 has thus far been concerned with the *algorithmic* part of ASPC as applied to a polymerization process, and most of the preceding activity represents a fairly straightforward implementation of techniques from the fields of time series analysis and stochastic control. Nevertheless, the work is somewhat novel in that the application is concerned with product quality in contrast to more usual control applications dealing with process variables such as feedrate and oxygen concentration. Section 1 stressed that ASPC seeks to integrate ideas from both automatic control and traditional SPC. In ASPC, the role of statistical monitoring is to detect and signal when operation of the closed-loop process is not consistent with the estimated model and control algorithm. This section presents a cumulative sum (CUSUM) monitoring chart developed for the polymerization process operating under algorithmic control.

A Process-Step Change. Returning to Figure 1, which shows viscosity measurements that would have resulted from holding catalyst constant, one might wonder whether the abrupt downward shift beginning at period 84 in the sequence is evidence of a fundamental change in the polymerization process. A change of this nature could, for example, be due to resupplying the large holding tanks of raw material feeding the reactors. With the new material, perhaps

the nominal amount of catalyst required to produce the target viscosity has changed.

Two significant questions in monitoring a closed-loop process are evident: (1) How is information about underlying changes in the process reflected in the sequences of control actions and process output? (2) How can this information be best used to detect process changes? To illustrate, the apparent step change in the "no control" viscosity sequence plotted in Figure 1 appears as a filtered change under one-step minimum MSE control as shown in Figure 2. Obviously, the effect of control actions needs to be taken into account by an effective monitoring scheme.

For purposes of discussion, suppose that the process is given by (2) with known parameter values equal to the compromise estimate; however, suppose that, beginning with period 84, the mean has shifted by an amount δ ; that is, the correct model is

$$y_t = \delta I[t \geq 84] + \beta u_{t-1} + \frac{(1 - \theta B)}{(1 - \rho B)} e_t, \quad (9)$$

where

$$I[t \geq 84] = 0, \quad \text{if } t < 84 \\ = 1, \quad \text{if } t \geq 84.$$

Assume also that laboratory measurements are always available after one period so that the pure one-step minimum MSE control rule (3) is always used. In this case (3) shows that it is possible to reconstruct the output sequence $\{y_t\}$ from the control sequence $\{u_{t-1}\}$. For practical purposes, the converse also holds, since u_{t-1} has an infinite moving average representation in terms of $\{y_t\}$. In other words, since the control rule is known, each series contains exactly the same information, and hence it suffices to consider only one of them. In what follows, only the output series y_t is used as data. Account is made, however, for the fact that the process is operating under a known control algorithm.

Substituting the pure one-step minimum MSE control rule (3) into (9) and simplifying shows that the closed-loop process output is given by

$$y_t = d_{t-84} + e_t, \quad (10)$$

where the deterministic sequence of means d_t is

$$d_t = \delta \frac{(1 - \rho B)}{(1 - \theta B)} I[t \geq 0]. \quad (11)$$

Hence Figures 1 and 2 demonstrate two ways of looking at a step change in the process. In Figure 1 the mean of the plotted points changes at period 84 from 100 to a constant new level $100 + \delta$. The points plotted, however, are serially correlated with known covariance structure—that of an autoregressive moving average process of order (1, 1). In Figure 2, the plotted points are serially independent with mean

100 before batch 84 and with a patterned mean after batch 84. Independence is obtained because the process is operated under minimum MSE control; the patterned mean arises because the step change is passed through the inverse filter of the process noise, as Equation (11) shows.

Equation (10) shows that the mean of the (pure one-step) minimum MSE controlled process follows the solution d_t to the deterministic difference Equation (11). It is easy to verify that the solution is

$$d_t = 0, \text{ if } t < 0$$

$$= \delta \left[\left(\frac{1 - \rho}{1 - \theta} \right) + \theta^t \left(\frac{\rho - \theta}{1 - \theta} \right) \right] \text{ if } t \geq 0. \quad (12)$$

Note that $d_0 = \delta$, and thereafter the process mean decays exponentially to a new level $d_x = \delta(1 - \rho)/(1 - \theta)$. Using the compromise parameter estimates produces $d_x = .256\delta$. Thus an underlying step change in the process level is partially, though not completely, compensated for by the (one-step) minimum MSE controller.

Often, conventional practices of engineering control would use the potential for step changes to justify an integral term in the controller [e.g., replacing ρ with 1 in Eq. (3)] to give long-run compensation for a level shift. If the estimated model is correct and no step changes occur, however, any controller with integral action would produce an output MSE greater than necessary. Since we desire to detect and hope to remove any step change, we (at least initially) do not incorporate integral action in the controller, and we leave the detection of step changes to the monitoring system discussed next.

CUSUM Monitoring Chart. Equation (10) shows that even during a step change in the underlying process the closed-loop output is serially independent when pure one-step minimum MSE control is used. Pure one-step control is the usual mode of operation, since only occasionally are viscosity measurements delayed for two periods in the analytical laboratory. Following a shift, there is a short transient period after which the output mean stabilizes at a new level d_x . Aside from the transient period and occasional departures from pure one-step minimum MSE control, this is precisely the scenario in which CUSUM charts perform well. Since plant personnel were already familiar with these charts, it was natural to introduce a CUSUM monitoring scheme for the algorithmically controlled process.

Initial CUSUM implementation was consciously aimed at detecting shifts in the process mean because operational experience (as well as our own analyses conducted during verification studies performed after implementing algorithmic control) had provided evidence of the presence of such changes. A more com-

plete system, of course, would incorporate provisions for monitoring more general kinds of process characteristics. One important candidate would be the innovation standard deviation, σ_e , which could be tracked via, say, a moving range or moving standard deviation chart. Likewise, other equally valid charting alternatives, such as an EWMA chart or a Shewhart chart with run rules, might have been used for monitoring the process level. For the particular characteristics of this application, however, CUSUM monitoring of the process mean has proved to be an effective and successful tool.

Changes in process level were hypothesized to be due to any of several sources, including raw-materials effects and seasonal factors affecting heat-exchange effectiveness. The net effect of such a shift could be equivalently viewed either as a change in viscosity of the product or as a change in the nominal catalyst level needed to produce material of the target viscosity. The latter conceptualization has come to be the more popular among plant engineers and operators. One purpose of the CUSUM system was to detect such shifts as quickly as possible when they occurred to resolve which of the conjectured mechanisms (if any) was in fact responsible.

The principal practice issue to be addressed was the granularity with which the process should be charted; that is, should separate CUSUM charts be maintained for each of the several reactors, data from all reactors incorporated into a single chart, or some option in between? Each alternative posed advantages and disadvantages, as we believe there was no intrinsically "right" answer. Aggregating data from multiple sources on a single chart suffers from the obvious drawback that the effect of an excursion in one reactor is diluted by data from others, which may still be performing "on nominal." Indeed, opposing shifts in reactors that contribute data to the same chart might even have a cancellation effect, allowing both to perform at their new levels indefinitely without being detected. Variance monitoring could prove to be of great value in such an instance. On the other hand, using separate charts for each reactor means that data arrive less frequently on any given chart, thus allowing common shifts in multiple reactors to go undetected for an appreciably longer elapsed time, with a correspondingly detrimental impact on product quality—perhaps even the production of scrap. One is tempted to try an "all of the above" approach, but charting both individual reactors and groups of reactors, though viable, raises complex simultaneous inference issues among correlated charts.

In the present context, practical considerations guided our choice. Design efficiency had dictated that systems for charging raw materials, emptying reacted polymer, and so forth should be shared among

as many reactors as capacity would permit. Therefore, the reactors had been built in two largely separate groups, with the reactors in each group sharing these "overhead" resources. There also existed modest differences in technology and capacity between the two groups. Production scheduling also tended to be handled by reactor group.

Experience indicated that, although process changes affecting a single reactor were not unheard of, the linkage that existed within groups resulted more often in changes that were manifested by the reactor group, or even plantwide (if both groups were affected). Moreover, although solo excursions were a concern, their impact was mitigated somewhat by the practice of scheduling most or all of the reactors in a group to be in production simultaneously; thus not only the data but also any off-target product would be diluted by the presence of other reactors that were still on target.

These considerations suggested that the most suitable way to construct the CUSUM system would be to chart data by reactor group. Chart properties desired for each group varied slightly, since differing capacities and numbers of reactors in each group meant that changes in the product needed to be detected within a different number of batches to constitute a comparable volume of material produced—which is the physically meaningful measure of time to detection in the present problem. The same quality criterion (i.e., the change in viscosity that needed to be detected) was applied to both reactor groups, however.

From this point, CUSUM chart design followed the recommendations of Lucas (1976), who used lower and upper statistics defined, respectively, as $L_t = \max[0, -y_t - k + L_{t-1}]$ and $H_t = \max[0, y_t - k + H_{t-1}]$. A signal is given if either L_t or H_t exceeds a bound, h . One of the groups, for example, used CUSUM chart parameters $h = 9.45$ and $k = 1.4$, selected to rapidly detect a *one standard deviation* (σ_e) shift of size $\Delta = 2.8$ in the output mean; this requires some interpretation, however. As shown by Equation (12), a shift of size δ in the underlying process results in a steady-state shift in the closed-loop process of size $\Delta = \delta(1 - \rho)/(1 - \theta)$. For example, by employing the compromise parameter estimates in this result, a steady-state shift of $\Delta = 2.8$ would result from a shift in the underlying process of size $\delta = 10.9$, a dramatic change in viscosity.

The CUSUM system has been successful in identifying process changes as a trigger for root cause determination and continuous improvement. We shall comment further on this.

2.6 Implementation and Benefits

The progression of activity in the polymerization application may be summarized as follows: After

forming a process model from discussions with knowledgeable process engineers, the model was reparameterized into an ARMAX form as given in (2). Since the model fit actual data well, a minimum MSE control algorithm was developed as if the estimated model were correct. Algorithm development was complicated by the fact that viscosity measurements from laboratory analyses may be delayed by either one or two periods. The control algorithm was applied to historical process data, and this cross-validated its effectiveness. Next, a CUSUM monitoring chart was developed for use on the process running under algorithmic control. The chart is intended to detect a step change in the nominal catalyst level necessary to attain the target viscosity.

Initially, the minimum MSE adjustment algorithm was implemented on only one reactor for a trial period. The results were so favorable, however, that management immediately requested that the algorithm be applied to all reactors. CUSUM monitoring charts were installed somewhat later after observing that changes in the process could cause the adjustment algorithm to persistently underestimate (or overestimate) the amount of catalyst necessary to produce on-target viscosity.

Since installing the optimal control policy and the CUSUM monitoring chart, a 35% reduction in the standard deviation of viscosity has been realized at the reaction stage of this process. A great reduction has also been experienced down line in the fraction of product (now near 0) that fails to meet specifications. Plant personnel attribute much of the improvement to the installation of the ASPC system and its operational compatibility with parallel engineering initiatives.

The CUSUM system has successfully notified plant staff of shifts in the nominal catalyst level needed to maintain on-target performance. To date, root-cause analyses have turned up two important findings in this regard. The first was the discovery of a faulty catalyst feed valve, which allowed residual catalyst to remain in the charge line, resulting in a viscosity "spike" when certain reactors were brought on or off line.

The second, more sweeping, finding was that CUSUM indications corresponded on a preponderance of occasions with identifiable thermal events—supporting an early conjecture regarding the origin of nominal shifts. A subsequent engineering study confirmed that reactions were sometimes heat-transfer limited. This study resulted in two corrective responses. The first was a reprogramming of the closed-loop temperature control during reaction; this by itself resulted in further significant reductions in viscosity variability. The other was the decision to proceed with substantive modifications to the plant heat-exchange system—a major capital improvement

project nearing completion as of this writing. These changes are expected to eliminate once and for all the effect of heat-transfer limitations as a source of variability in polymer production.

3. APPLICATION GUIDELINES

The application of ASPC described in Section 2 followed general guidelines that were developed by Tucker et al. (1991). To place the polymerization application in a broader context, this section outlines a procedure for implementing ASPC on typical industrial processes. Whereas the polymerization application exemplifies this methodology, it was justified by Tucker et al.

There are two essential characteristics a process must possess if the algorithmic (i.e., control) portion of ASPC is to be successful: First, it must be possible to use past process data and covariates to construct a good predictor of future process performance. Second, there must be available a compensatory variable whose adjustment will have a predictable effect on the performance property of interest. These characteristics are clearly present in the process underlying Model (1), since past values of y_t and u_{t-1} can be used to construct a good predictor of intrinsic viscosity. Furthermore, using the amount of catalyst as a compensatory variable was an obvious and effective choice.

The general four-step procedure followed in the application is as follows:

1. Develop a time series transfer-function model for the process output including the effect of past performance, control actions, and other relevant process characteristics. This involves identifying process orders and delays and estimating model parameters. In the application, this step resulted in Model (2).

2. Next, based on pertinent costs, design a control rule for the estimated model. For example, recall the variable delay scheme given by Equations (4)–(7).

3. Along with installing the control rule, put in place SPC charts to monitor the closed-loop process. The SPC charts should signal if the process and controller are no longer operating as expected from the identification and estimation stage. Section 2.5 described our realization of this step in terms of the CUSUM monitoring chart.

4. When a monitoring signal occurs, search for an assignable cause and, if feasible, remove it. If no cause is found or the cause cannot realistically be removed, it may be necessary to reestimate system parameters or even reidentify the process form and orders. If the process was reestimated or reidentified, return to step 2. Responses to monitoring signals in the application were discussed in Section 2.6.

In our experience, each stage of the foregoing procedure requires careful planning and analysis. At least in our case, however, the result has been a system that operators and engineers have been able to use successfully in their drive for continuous quality improvement.

4. SUMMARY

ASPC represents a proactive approach to quality improvement in which concepts from automatic control are joined with ideas from SPC. Variations in product quality are then reduced in two ways, (1) through algorithmic compensation for predictable quality deviations and (2) through elimination of root causes of variability as signaled by statistical monitoring charts. Engineering control and SPC have for the most part developed in isolation from one another. We, however, advocate integrating tools from both fields so as to yield quality improvements *both by removing sources of variability and by compensating for predictable process deviations from target.*

Application of ASPC to a polymerization process has resulted in a 35% reduction in viscosity variability at the reaction stage of the process and virtual elimination of off-spec material from this source. In this study, changes to the chief quality characteristic, viscosity, were made by adjusting a compensatory variable, the amount of catalyst. A minimum MSE control algorithm was developed for this process, and then the closed-loop output was monitored by a CUSUM chart.

ACKNOWLEDGMENTS

We thank the referees for their exceptionally helpful and insightful comments, which improved the content and presentation of this article. In particular, comments by one referee led to the simple form of the adjustment algorithm with variable delays. Another referee supplied several references in which automatic control has been applied to product quality variables.

APPENDIX A: EQUIVALENCE OF MODELS (1) AND (2)

The equivalence follows from Box and Jenkins (1976, A4.4), who implicitly appealed to a representation theorem like that of Fuller (1976, theorem 2.6.3). To relate the parameters of Models (1) and (2), write (1) as

$$y_t - \rho y_{t-1} = \beta(u_{t-1} - \rho u_{t-2}) + \delta_t, \quad (\text{A.1})$$

where $\delta_t = a_t + f_t - \rho f_{t-1}$, so $E(\delta_t) = 0$ and

$$\begin{aligned} E(\delta_t \delta_{t-k}) &= \sigma_a^2 + (1 + \rho^2)\sigma_f^2, & k = 0 \\ &= -\rho\sigma_f^2, & k = 1 \\ &= 0, & k = 2, 3, \dots \end{aligned} \quad (\text{A.2})$$

Following Box and Jenkins, model equivalence is obtained by representing δ_t as

$$\delta_t = e_t - \theta e_{t-1},$$

$$e_t, \text{ independent } N(0, \sigma_e^2), \quad (\text{A.3})$$

so

$$E(\delta_t \delta_{t-k}) = (1 + \theta^2)\sigma_e^2, \quad k = 0$$

$$= -\theta\sigma_e^2, \quad k = 1$$

$$= 0, \quad k = 2, 3, \dots \quad (\text{A.4})$$

Setting (A.2) equal to (A.4) and solving relates (σ_f^2, σ_a^2) to $(\theta, \rho, \sigma_e^2)$ by $\sigma_f^2 = \theta\sigma_e^2/\rho$ and $\sigma_a^2 = [1 + \theta^2 - \theta(\rho^{-1} + \rho)]\sigma_e^2$.

APPENDIX B: MINIMUM MSE CONTROL WHEN MEASUREMENTS ARE DELAYED

Equations (6) and (7) for forecasting N_t were given by Box and Jenkins (1976, chap. 5). Furthermore,

$$N_t = \hat{N}_{t|t-1} + e_t$$

$$= \hat{N}_{t|t-2} + e_t + (\rho - \theta)e_{t-1}. \quad (\text{B.1})$$

Since e_t is independent of $(y_{t-1}, y_{t-2}, \dots)$ and of $\hat{N}_{t|t-1}$ and similarly (e_t, e_{t-1}) is independent of $(y_{t-2}, y_{t-3}, \dots)$ and of $\hat{N}_{t|t-2}$, Equation (B.1) together with (2) implies that $\text{MSE}(y_t)$ is minimized by negating the most up-to-date forecast of N_t as in (5).

It is important to understand that the usual steady-state fixed-delay minimum MSE control rules giving u_{t-1} as a fixed linear combination of a finite number of past u_t 's and y_t 's [like (3) and its two-step version] are not appropriate when the delay may vary (and, more generally, when the *actual* adjustments are not always as given by a fixed delay control rule). For example, suppose that the two-step minimum MSE rule $u_{t-1} = -\hat{N}_{t|t-2}/\beta$ has been applied through period $t-1$. This implies (e.g., see Box and Jenkins 1976)

$$u_{t-2} = -\frac{\rho(\rho - \theta)}{\beta(1 - \rho\beta)} e_{t-3} \quad (\text{B.2})$$

and

$$y_{t-1} = e_{t-1} + (\rho - \theta)e_{t-2}. \quad (\text{B.3})$$

Next suppose that in period t laboratory measurements are no longer delayed and the control action is given by the *steady-state* one-step rule (3). Substituting (B.2) and (B.3) into (3) gives

$$\beta u_{t-1} = -\frac{\rho^2(\rho - \theta)}{(1 - \rho\beta)} e_{t-3}$$

$$- (\rho - \theta)[e_{t-1} + (\rho - \theta)e_{t-2}].$$

Substituting this into (2) and simplifying gives $y_t = e_t + \theta(\rho - \theta)e_{t-2}$, whereas the correct adjustment

$u_{t-1} = -\hat{N}_{t|t-1}/\beta$ results in $y_t = e_t$, which has a smaller MSE (for $\rho \neq \theta$).

The forecasting idea embodied in (4)–(7) can be generalized to higher order ARMAX systems including systems with adjustment dynamics.

[Received January 1990. Revised December 1991.]

REFERENCES

- Adams, B. M., and Woodall, W. H. (1989). "An Analysis of Taguchi's On-line Process Control Procedure Under a Random Walk Model." *Technometrics*, 31, 401–413.
- Astrom, K. J. (1970). *Introduction to Stochastic Control Theory*. New York: Academic Press.
- Barnard, G. A. (1959). "Control Charts and Stochastic Processes." *Journal of the Royal Statistical Society, Ser. B*, 21, 239–271.
- Bar-Shalom, Y., and Tse, E. (1974). "Dual Effect, Certainty Equivalence, and Separation in Stochastic Control." *IEEE Transactions on Automatic Control*, 19, 494–500.
- Bather, J. A. (1963). "Control Charts and the Minimization of Costs." *Journal of the Royal Statistical Society, Ser. B*, 25, 49–80.
- BMDP Statistical Software (1983). *BMDP Statistical Software, 1983 Printing With Additions*. Los Angeles: Author.
- Box, G. E. P., and Jenkins, G. M. (1962). "Some Statistical Aspects of Adaptive Optimization and Control." *Journal of the Royal Statistical Society, Ser. B*, 24, 297–343.
- (1963). "Further Contributions to Adaptive Quality Control: Simultaneous Estimation of Dynamics: Non-Zero Costs." *Bulletin of the International Statistics Institute*, 34, 943–974.
- (1970). *Time Series Analysis Forecasting and Control*. Oakland, CA: Holden-Day.
- (1976). *Time Series Analysis Forecasting and Control* (rev. ed.). Oakland, CA: Holden-Day.
- Box, G. E. P., and Kramer, T. T. (1992). "Statistical Process Monitoring and Feedback Adjustment—A Discussion." *Technometrics*, 34, 251–267.
- Crowder, S. V. (1992). "An SPC Model for Short Production Runs: Minimizing Expected Cost." *Technometrics*, 34, 64–73.
- Duncan, A. J. (1986). *Quality Control and Industrial Statistics* (5th ed.). Homewood, IL: Richard D. Irwin.
- Fuller, W. A. (1976). *Introduction to Statistical Time Series*. New York: John Wiley.
- Jensen, K. L. (1989). "Optimal Adjustment in the Presence of Process Drift and Adjustment Error." unpublished Ph.D. dissertation, Iowa State University, Dept. of Statistics.
- Kramer, T. T. (1989). "Process Control From an Economic Point of View." unpublished Ph.D. dissertation, University of Wisconsin-Madison, Dept. of Statistics.
- Lucas, J. M. (1976). "The Design and Use of V-Mask Control Schemes." *Journal of Quality Technology*, 8, 1–12.
- MacGregor, J. F. (1988). "On-line Statistical Process Control." *Chemical Engineering Progress*, 84, 21–31.
- MacGregor, J. F., and Harris, T. J. (1990). Discussion of "Exponentially Weighted Moving Average Control Schemes: Properties and Enhancements." by J. M. Lucas and M. S. Saccucci. *Technometrics*, 32, 23–26.
- MacGregor, J. F., and Tidwell, P. W. (1980). "Modeling and Control of Continuous Industrial Polymerization Reactors." in *Computer Applications to Chemical Engineering* (ACS Symposium Series 124), eds. R. G. Squires and G. V. Reklaitis. Washington, DC: American Chemical Society, pp. 251–268.
- Matlab (1987). *Pro-Matlab*. Sherborn, MA: The Mathworks.
- Roffel, J. J., MacGregor, J. F., and Hoffman, T. W. (1989). "The Design and Implementation of a Multivariable Internal Model

- Controller for a Continuous Polybutadiene Polymerization Train." in *Proceedings of the IFAC Symposium on Dynamics and Control of Chemical Reactors, Distillation Columns, and Batch Processes*, ed. C. McGreavy, Oxford, U.K.: Pergamon Press.
- SAS Institute Inc. (1984), *SAS/ETS Users Guide, Version 5 Edition*, Cary, NC: Author.
- Shewhart, W. A. (1931), *Economic Control of Quality of Manufactured Product*, Princeton, NJ: Van Nostrand Reinhold.
- Statistical Sciences Inc. (1990), *S-Plus Users Manual*, Seattle, WA: Author.
- Taguchi, G. (1985), "Quality Engineering in Japan." *Communications in Statistics—Theory and Methods*, 14, 2785–2801.
- Tucker, W. T., Faltin, F. W., and Vander Wiel, S. A. (1991), "Algorithmic Statistical Process Control: An Elaboration," Statistical Research Report 102, AT&T Bell Labs, Murray Hill, NJ.
- Vander Wiel, S. A. (1991), "Optimal Discrete Adjustments for Short Production Runs," Statistical Research Report 101, AT&T Bell Labs, Murray Hill, NJ.
- Westcott, J. H. (1962), Discussion of "Some Statistical Aspects of Adaptive Optimization and Control," by G. E. P. Box and G. M. Jenkins, *Journal of the Royal Statistical Society, Ser. B*, 24, 340–341.