14. Cuscore Statistics: Directed Process Monitoring for Early Problem Detection

This chapter presents the background to the Cuscore statistic, the development of the Cuscore chart, and how it can be used as a tool for directed process monitoring. In Sect. 14.1 an illustrative example shows how it is effective at providing an early signal to detect known types of problems, modeled as mathematical signals embedded in observational data. Section 14.2 provides the theoretical development of the Cuscore and shows how it is related to Fisher's score statistic. Sections 14.3, 14.4, and 14.5 then present the details of using Cuscores to monitor for signals in white noise, autocorrelated data, and seasonal processes, respectively. The capability to home in on a particular signal is certainly an important aspect of Cuscore statistics. however, Sect. 14.6 shows how they can be applied much more broadly to include the process model (i. e., a model of the process dynamics and noise) and process adjustments (i. e., feedback control). Two examples from industrial cases show how Cuscores can be devised and used appropriately in more complex monitoring applications. Section 14.7 concludes the chapter with a discussion and description of future work.

The traditional view of statistical process control is that a process should be monitored to detect any aberrant behavior, or what *Deming* [14.1] called "special causes" that are suggested by significant patterns in the data that point to the existence of systematic signals. The timing, nature, size, and other information about the signals can lead to the identification of the signaling factor(s) so that it can (ideally) be permanently eliminated. Conventional Shewhart charts are designed with exactly this philosophy, where the signal they detect is an unexpected spike change in white noise.

Many situations occur, however, where certain process signals are anticipated because they are characteristic of a system or operation. The cumulative score (Cuscore) chart can be devised to be especially sensitive to deviations or signals of an expected type. In general, after working with a particular process, engineers and operators often know – or at least have a belief – about how a process will potentially falter. (Unfortunately, the problem seldom announces its time and location in advance.) For example, consider a process where a valve is used to maintain pressure in a pipeline. Because the valve will experience wear over time, it must be periodically replaced. However, in addition to the usual wear, engineers are concerned that the value may fatigue or fail more rapidly than normal. The Cuscore chart can be used to incorporate this working knowledge and experience into the statistical monitoring function. This concept often has a lot of intuitive appeal for industry practitioners.

After laying the background and theoretical foundation of Cuscores this chapter progresses through signal detection in white noise, autocorrelated data, and seasonal data. Two examples from actual industry settings show how Cuscores can be devised and used appropriately in more complex monitoring applications. The final section of the chapter provides a discussion on how Cuscores can be extended in a framework to include statistical experiments and process control.

14.1 Background and Evolution of the Cuscore in Control Chart Monitoring

Statistical process control (SPC) has developed into a rich collection of tools to monitor a system. The first control chart proposed by *Shewhart* [14.2] is still the most widely used in industrial systems [14.3]. As observational data from the system are plotted on the chart, the process is declared "in control" as long as the points on the chart stay within the control limits. If a point falls outside those limits an "out of control" situation is declared and a search for a special cause is initiated.

Soon practitioners realized that the ability of the Shewhart chart to detect small changes was not as good as its ability to detect big changes. One approach to improve the sensitivity of the chart was to use several additional rules (e.g., *Western Electric rules* [14.4] that signal for a number of consecutive points above the center line, above the warning limits, and so on). Another approach was to design complementary charts that could be used in conjunction with the Shewhart chart but that were better at detecting small changes. *Page* [14.5] and *Barnard* [14.6] developed the cumulative sum (Cusum) chart where past and present data are used in a cumulative way to detect small shifts in the mean. *Roberts* [14.7] and *Hunter* [14.8] proposed the exponentially weighted moving average (EWMA) as another way to detect small changes. This ability comes from the fact that the EWMA statistic can be written as a moving average of the current and past observations, where the weights of the past observations fall off exponentially.

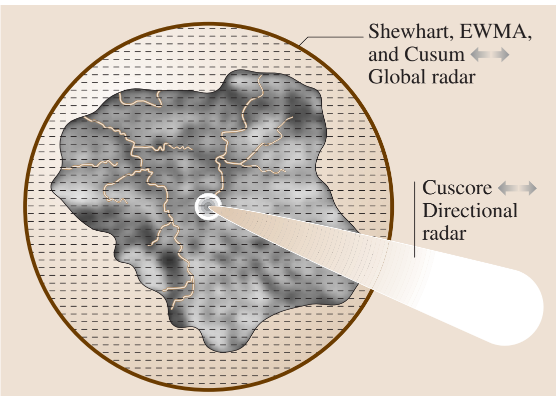

Fig. 14.1 The roles of the Shewhart and Cuscore charts are compared to those of global and directional radar defenses for a small country

Of course the Shewhart, Cusum, and EWMA charts are broadly applicable to many types of process characterizations. Remarkably, the Cuscore chart generalizes the Shewhart, Cusum, and EWMA charts; however, its real benefit is that it can be designed to be a high-powered diagnostic tool for specific types of process characterizations that are not covered by the basic charts. We will develop this result more formally after introducing the Cuscore theory. However, an analogy due to *Box* [14.9] will help to establish the ideas.

Suppose a nation fears aerial attack. As Fig. 14.1 shows, a global radar scanning the full horizon will have a broad coverage of the entire border, but with

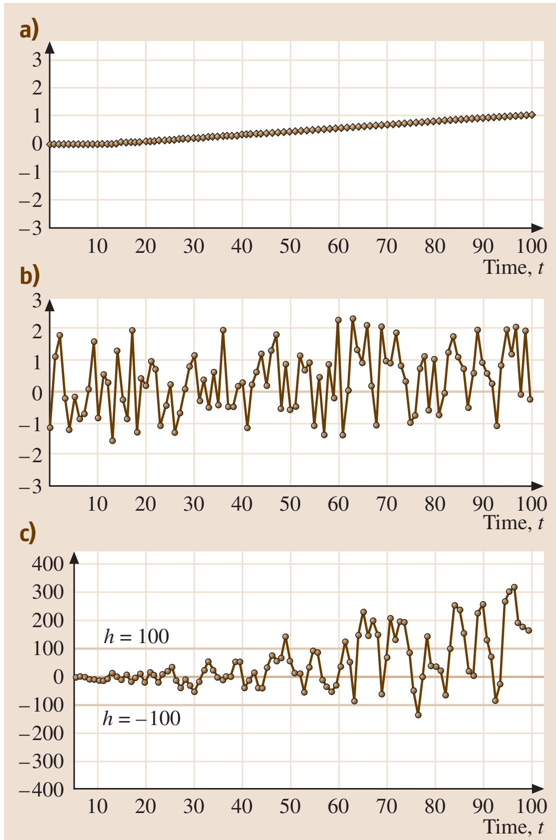

Fig. 14.2a–c Detection of a ramp signal: **(a)** ramp signal beginning at time 10; **(b)** the signal plus white noise consisting of 100 random normal deviates with zero mean and standard deviation $\sigma = 1$; and **(c)** the Cuscore statistic applied to the data of **(b)**

a limited range; this is the analog of the Shewhart, Cusum, and EWMA charts. A directional radar aimed in the direction of likely attack will have a specific zone of the border to cover, but with a long range for early detection; this is the analog of the Cuscore chart.

As a first illustration of the Cuscore chart, let us consider it within the framework of looking for a signal in noise. Suppose we have an industrial process where the objective is to control the output Y_t to a target value T. We may conveniently view the target as the specification and record deviations from the target. Suppose that the process may experience a small drift to a new level over time – a ramp signal. Although corrective actions have been taken to hopefully resolve the process, it is feared that the same problem might reoccur. The components of the process are illustrated in Fig. 14.2 which shows: (a) the ramp signal beginning at time $t = 10$; (b) the signal plus white noise consisting of 100 random normal deviates with zero mean and standard deviation $\sigma = 1$; and (c) the appropriate Cuscore statistic

$$Q_t = \sum_{i=1}^{t} (Y_t - T)t \,.$$

The development of the statistic is also shown in Sect. 14.3. We may note that Fig. 14.2b is equivalent to a Shewhart control chart with upper and lower control limits of $+3\sigma$ and -3σ respectively. The Shewhart chart is relatively insensitive to small shifts in the process, and characteristically, it never detects the ramp signal (Fig. 14.2b). With a decision interval $h = 100$, the Cuscore chart initially detects the signal at time 47 and continues to exceed h at several later time periods (Fig. 14.2c). Although tailored to meet different process monitoring needs, the EWMA and Cusum charts would similarly involve a direct plot of the actual data to look for an unexpected signal in white noise. In this case, since we have some expectation about the signal, i. e., that it is a ramp, we incorporate that information into the Cuscore by multiplying the differences (which are residuals) by t before summing.

Similarly, if demanded by the monitoring needs, the Cuscore can be devised to monitor for a mean shift signal in autocorrelated noise, or for a bump signal in nonstationary noise, or for an exponential signal in correlated noise, or any other combination. Indeed, the Cuscore chart can be designed to look for almost any kind of signal in any kind of noise. The theoretical development of the Cuscore statistic will help illuminate this idea.

14.2 Theoretical Development of the Cuscore Chart

Consider a model of the output of a process determined by adding the process target to an autoregressive integrated moving average (ARIMA) time-series model:

$$Y_t = T + \frac{\theta(B)}{\phi(B)} a_{t0} \,, \tag{14.1}$$

where B is the backshift operator such that $B^k X_t = X_{t-k}$; $\phi(B)$ and $\theta(B)$ are the autoregressive (AR) and moving average (MA) polynomials parameterized as $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$ $[(1-B)\phi(B)$ can be used to difference the process]; the a_t values are independent and identically distributed $N(0, \sigma_a^2)$(i. e., white noise). However, in the model the zero in a_t is added to indicate that the a_{t0} values are just residuals; they are not white-noise innovations unless the model is true. This model is referred to as the null model: the in-control model assuming that no feared signal occurs.

Now assume that an anticipated signal that could appear at some time where γ is some unknown parameter of the signal, and $f(t)$ indicates the nature of the signal:

$$Y_t = T + \frac{\theta(B)}{\phi(B)} a_t + \gamma f(t) \,. \tag{14.2}$$

This model is referred to as the discrepancy model and is assumed to be true when the correct value for γ is used.

Box and *Ramírez* [14.10, 11] presented a design for the Cuscore chart to monitor for an anticipated signal. It is based on expressing the statistical model in (14.2) in terms of white noise:

$$a_i = a_i(Y_i, X_i, \gamma) \quad \text{for } i = 1, 2, \ldots, t \,, \tag{14.3}$$

where X_i are the known levels of the input variables. The concept is that we have properly modeled the system so that only white noise remains when the signal is not present. After the data have actually occurred, for each choice of γ, a set of a_i values can be calculated from (14.3). In particular, let γ_0 be some value, possibly different from the true value γ of the parameter. The sequential probability ratio test for γ_0 against some other value γ_1 has the likelihood ratio

$$LR_t = \prod_{i=1}^{t} \exp\left\{\frac{1}{2\sigma_a^2}\left[a_i^2(\gamma_0) - a_i^2(\gamma_1)\right]\right\}.$$

After taking the log, this likelihood ratio leads to the cumulative sum

$$S_t = \frac{1}{2\sigma_a^2}\sum_{i=1}^{t}\left[a_i^2(\gamma_0) - a_i^2(\gamma_1)\right].$$

Expanding $a_i^2(\gamma)$ around γ_0, letting $\eta = (\gamma_1 - \gamma_0)$, and $d_i = -\frac{\partial a_i}{\partial \gamma}|_{\gamma=\gamma_0}$ we have

$$\begin{aligned} S_t &= \frac{1}{2\sigma_a^2}\sum_{i=1}^{t}\left[2\eta a_i(\gamma_0)d_i(\gamma_0) - \eta^2 d_i^2(\gamma_0)\right] \\ &= \frac{\eta}{\sigma_a^2}\sum_{i=1}^{t}\left[a_i(\gamma_0)d_i(\gamma_0) - \frac{\eta}{2}d_i^2(\gamma_0)\right]. \end{aligned}$$

The quantity

$$Q_t = \sum_{i=1}^{t}\left[a_i(\gamma_0)d_i(\gamma_0) - \frac{\eta}{2}d_i^2(\gamma_0)\right] = \sum_{i=1}^{t} q_i \tag{14.4}$$

is referred to as the Cuscore associated with the parameter value $\gamma = \gamma_0$ and d_i is referred to as the detector. The detector measures the instantaneous rate of change in the discrepancy model when the signal appears. *Box* and *Luceño* [14.12] liken its operation to a radio tuner because it is designed in this way to synchronize with any similar component pattern existing in the residuals. Accordingly, it is usually designed to have the same length as the anticipated signal series. The term $\frac{\eta}{2}d_i^2(\gamma_0)$ can be viewed as a reference value around which $a_i(\gamma_0)d_t(\gamma_0)$ is expected to vary if the parameter does not change. The quantity $a_i(\gamma_0)d_t(\gamma_0)$ is equal to Fisher's score statistic [14.13], which is obtained by differentiating the log likelihood with respect to the parameter γ. Thus

$$\begin{aligned} \frac{\partial}{\partial \gamma}\left[\ln p\left(a_i|\gamma\right)\right]\Big|_{\gamma=\gamma_0} &= \frac{\partial}{\partial \gamma}\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{t} a_i^2\right)\Bigg|_{\gamma=\gamma_0} \\ &= \frac{1}{\sigma^2}\sum_{i=1}^{t} a_i(\gamma_0)d_i(\gamma_0), \end{aligned}$$

where $p(a_i|\gamma)$ is the likelihood or joint probability density of a_i for any specific choice of γ and the $a_i(\gamma_0)$ values are obtained by setting $\gamma = \gamma_0$ in (14.3). Since the q_is are in this way a function of Fisher's score function, the test procedure is called the Cuscore. The Cuscore statistic then amounts to looking for a specific signal $f(t)$ that is present when $\gamma \neq \gamma_0$.

To use the Cuscore operationally for process monitoring, we can accumulate q_i only when it is relevant for the decision that the parameter has changed and reset it to zero otherwise. Let Q_t denote the value of the Cuscore procedure plotted at time t, i. e., after observation t has been recorded. Let $Q_t^+$and Q_t^-denote the one-sided upper and lower Cuscores respectively as follows:

$$Q_t^+ = \max(0, Q_{t-1}^+ + q_t), \tag{14.5a}$$

$$Q_t^- = \min(0, Q_{t-1}^- + q_t), \tag{14.5b}$$

where the starting values are $Q_0^+ = Q_0^- = 0$. The one-sided Cuscore is preferable when the system has a long period in the in-control state, during which Q_t would drift and thus reduce the effectiveness of the monitoring chart.

If either Q_t^+or Q_t^- exceed the decision interval h, the process is said to be out of control. *Box* and *Ramírez* [14.10] showed that an approximation to h can be obtained as a function of the type-I error, α, the magnitude of the change in the parameter $\gamma = (\gamma_1 - \gamma_0)$, and the variance of the as:

$$h = \frac{\sigma_a^2 \ln(1/\alpha)}{\gamma}. \tag{14.6}$$

For simpler models, we could also develop control limits for the Cuscore chart by directly estimating the standard deviation of the Cuscore statistic. For more complex models, control limits may be obtained by using simulation to evaluate the average run length associated with a set of out of control conditions.

14.3 Cuscores to Monitor for Signals in White Noise

Let us now consider the Cuscore statistics for the basic case of monitoring for signals in white noise, which is the assumption underlying the traditional Shewhart, EWMA, and Cusum charts. We will develop them without the reference value in (14.4), but the reference value will help to improve the average run-length performance of the chart when used in practice. We can write the white-noise null model using (14.1) where the ϕ and θ parameters are set equal to zero, i. e.,

$$Y_t = T + a_{t0}.$$

Writing a_{t0} on the left makes it clear that each residual is the difference between the output and the target:

$$a_{t0} = Y_t - T\ . \tag{14.7}$$

If the model is correct and there is no signal, the result will be a white-noise sequence that can be monitored for the appearance of a signal. When the signal does show up, the discrepancy model is thus

$$Y_t = T + a_t + \gamma f(t)$$

which can be equivalently written with the white noise quantity a_t on the left as

$$a_t = Y_t - T - \gamma f(t)\ .$$

The form of the signal will determine the form of the detector and hence the form of the Cuscore.

The Shewhart chart is developed under the assumption of white noise and that the signal for which the chart detects efficiently is a spike signal:

$$f(t) = \begin{cases} 0 & t \neq t_0 \\ 1 & t = t_0\ . \end{cases} \tag{14.8}$$

For the spike signal in the discrepancy model, the appropriate detector d_t is

$$d_t = -\left.\frac{\partial a_t}{\partial \gamma}\right|_{\gamma=\gamma_0} = 1\ . \tag{14.9}$$

By (14.4), (14.7), and (14.9) the Cuscore statistic is

$$Q_t = \sum_{i=1}^{t} a_{i0} d_i$$
$$= a_{t0}\ ,$$

where the last equality follows since the detector for the spike is only for one period (i.e, the current one) given that the signal series and detector series have the same length. Hence, the Cuscore tells us to plot the current residual, which is precisely the design of the Shewhart chart.

The EWMA chart is developed under the assumption of white noise and that the signal that the chart is designed to detect is an exponential signal with parameter γ:

$$f(t) = \begin{cases} 0 & t > t_0 \\ 1 + \gamma_{t-1} + \gamma_{t-2}^2 + \gamma_{t-3}^3 + \cdots & t \leq t_0\ . \end{cases}$$

For the exponential signal in the discrepancy model, the appropriate detector d_t is

$$d_t = -\left.\frac{\partial a_t}{\partial \gamma}\right|_{\gamma=\gamma_0} = 1 + \gamma_{t-1} + \gamma_{t-2}^2 + \gamma_{t-3}^3 + \cdots\ . \tag{14.10}$$

By (14.4), (14.7), and (14.10) the appropriate Cuscore statistic is

$$Q_t = \sum_{i=1}^{t} a_{i0} d_i$$
$$= a_{t0} + \gamma a_{t0-1} + \gamma^2 a_{t0-2} + \gamma^3 a_{t0-3} + \cdots\ .$$

Here the Cuscore tells us to sum the current and past residuals, applying an exponentially discounted weight to the past data, which is the design of the EWMA chart.

The Cusum chart is developed under the assumption of white noise and that signal to detect is a step change or mean shift given by

$$f(t) = \begin{cases} 0 & t < t_0 \\ 1 & t \geqslant t_0\ . \end{cases} \tag{14.11}$$

In this case, the discrepancy model and the detector are the same as for the spike signal. However, since the signal remains in the process, the detector is applied over all periods to give the Cuscore statistic

$$Q_t = \sum_{i=1}^{t} a_{i0}\ .$$

Here the Cuscore tells us to plot the sum of all residuals, which is precisely the design of the Cusum chart.

A variation of the step change is one that lasts only temporarily, which is called a bump signal of length b

$$f(t) = \begin{cases} 1 & t_{0-b+1} \leqslant t \leqslant t_0 \\ 0 & \text{otherwise}\ . \end{cases}$$

When this signal appears in white noise, the detector is applied only as long as the bump, giving the Cuscore statistic

$$Q_t = \sum_{i=1}^{t} a_{i0-b-1}\ .$$

This is equivalent to the arithmetic moving-average (AMA) chart, which is frequently used in financial analysis (e.g., see TraderTalk.com or Investopedia.com).

The ramp signal that may start to appear at time t_{0-r} where r is the duration of the ramp with a final value m is modeled by

$$f(t) = \begin{cases} \frac{m}{r}t & t_{0-r} \leqslant t \leqslant t_0 \\ 0 & \text{otherwise} . \end{cases}$$

The discrepancy model is the same as with the Shewhart chart, but for this signal the detector is given by

$$d_t = -\left.\frac{\partial a_t}{\partial \gamma}\right|_{\gamma=\gamma_0} = t .$$

The Cuscore is hence

$$Q_t = \sum_{i=1}^{t} a_{i0} d_i = \sum_{i=1}^{t} a_{i0} t = \sum_{i=1}^{t} (Y_t - T)\, t$$

as the example in the introduction shows.

14.4 Cuscores to Monitor for Signals in Autocorrelated Data

In many real systems, the assumption of white-noise observations is not even approximately satisfied. Some examples include processes where consecutive measurements are made with short sampling intervals and where quality characteristics are assessed on every unit in order of production. Financial data, such as stock prices and economic indices are certainly not uncorrelated and independent observations. In the case of autocorrelated data the white-noise assumption is violated. Consequently the effectiveness of the Shewhart, Cusum, EWMA, and AMA charts is highly degraded because they give too many false alarms. This point has been made by many authors (e.g., see *Montgomery* [14.14] for a partial list).

Alwan and *Roberts* [14.15] proposed a solution to this problem by modeling the non-random patterns using ARIMA models. They proposed to construct two charts: 1) a common-cause chart to monitor the process, and 2) a special-cause chart on the residuals of the ARIMA model. Extensions of the these charts to handle autocorrelated data have been addressed by several authors. *Vasilopoulos* and *Stamboulis* [14.16] modified the control limits. *Montgomery* and *Mastrangelo* [14.17] and *Mastrangelo* and *Montgomery* [14.18] used the EWMA with a moving center line (MCEWMA). However, when signals occur in autocorrelated data, there is a pattern in the residuals that that the residuals-based control charts do not use. The Cuscore, on the other hand, does incorporate this information through the detector. As we have seen, the detector plays an important role in determining Cuscore statistics but this role is attenuated for autocorrelated data.

As in the previous section, we can use the reference value in practice, but will develop the main result without it. Assuming the null model in (14.1) is invertible, i. e., $|\theta| < 1$, it can be written in terms of the residuals as

$$a_{t0} = (Y_t - T)\frac{\phi(B)}{\theta(B)} . \tag{14.12}$$

The discrepancy model in (14.2) can be equivalently written with the white-noise quantity a_t on the left as

$$a_t = [Y_t - T - \gamma f(t)]\frac{\phi(B)}{\theta(B)} . \tag{14.13}$$

We see that to recover the white-noise sequence in an autocorrelated process, both the residuals and the signal must pass through the inverse filter $\phi(B)/\theta(B)$. Hence, the residuals have time-varying mean $\gamma f(t)\{[\phi(B)/\theta(B)]\}$ and variance σ_a^2. Using (14.13), the detector d_t is

$$d_t = -\left.\frac{\partial a_t}{\partial \gamma}\right|_{\gamma=\gamma_0} = f(t)\frac{\phi(B)}{\theta(B)} . \tag{14.14}$$

By (14.4), (14.13), and (14.14) the Cuscore statistic is

$$\begin{aligned} Q_t &= \sum_{i=1}^{t} a_{i0} d_i \\ &= \sum_{i=1}^{t} \left[(Y_i - T)\frac{\phi(B)}{\theta(B)}\right] f(t)\frac{\phi(B)}{\theta(B)} . \end{aligned}$$

Hu and *Roan* [14.19] mathematically and graphically showed the behavior of the detector for several combinations of signals and time-series models. Their study highlights that the behavior is different for different values of ϕ and θ determined by the stability conditions, the value of the first transient response, and the value of the steady-state response.

As an example, suppose we have the ARMA (1,1) noise model

$$(Y_t - T) - \phi_1(Y_{t-1} - T) = a_{t0} - \theta_1 a_{t0-1}$$

or

$$a_{t0} = (Y_t - T)\frac{1-\phi_1 B}{1-\theta_1 B}\ . \quad (14.15)$$

If the step signal in (14.11) occurs at time t_0, using (14.14) we can determine that a change pattern is produced:

$$d_t = f(t)\frac{\phi(B)}{\theta(B)} = \begin{cases} 0 & t < t_0 \\ 1 & t = t_0 \\ (\theta_1 - \phi_1)\theta_1^{t-(t_0+1)} & t \geqslant t_0 + 1\ . \end{cases} \quad (14.16)$$

Then the Cuscore statistic is the sum of the product of (14.15) and (14.16).

However, we can see an important issue that arises in autocorrelated data, which is how the time-varying detector is paired with the current residuals. For example, if we assume that we know the time of the step signal or mean shift, there is a match between the residuals and the detector and we use t_0 in the calculation of d_t for the Cuscore. When we do not know the time of the mean shift, there is a mismatch between the residuals and the detector; in this case we make the estimate $\hat{t}_0$ and write the detector as $d_{\hat{t}}$. (When $\hat{t}_0 = t_0$ then $d_{\hat{t}} = d_t$.) The match or mismatch will affect the robustness of the Cuscore chart, as considered for limited cases in *Shu* et al. [14.20] and *Nembhard* and *Changpetch* [14.21]. There is an opportunity to increase the understanding of this behavior through additional studies.

Yet another issue is to determine over how many periods the detector should be used in the case of a finite signal such as a step or a bump. On this point, *Box* and *Luceño* [14.12] use equal lengths for both white-noise and autocorrelated-noise models. Although such an assumption seems intuitive for white-noise models, on open question is whether a longer detector would improve the efficiency of the Cuscore chart in the case of autocorrelated data.

14.5 Cuscores to Monitor for Signals in a Seasonal Process

In this section, we present the first example of a Cuscore application in an industry case. One of the major services of the Red Cross is to manage blood platelet inventory. Platelets are irregularly-shaped colorless bodies that are present in blood. If, for some unexpected reason, sudden blood loss occurs, the blood platelets go into action. Their sticky surface lets them, along with other substances, form clots to stop bleeding [14.22]. *Nembhard* and *Changpetch* [14.21] consider the problem of monitoring blood platelets, where the practical goal is to detect a step shift in the mean of a seasonal process as an indicator that demand has risen or fallen. This information is critical to Red Cross managers, as it indicates a need to request more donors or place orders for more blood with a regional blood bank. A distinction of this problem is that the step shift, although a special cause, is a characteristic of the system. That is, from time to time, shifts in the mean of the process occur due to natural disasters, weather emergencies, holiday travel, and so on. Given the structure of characteristic shifts in this application, directed monitoring is a natural choice.

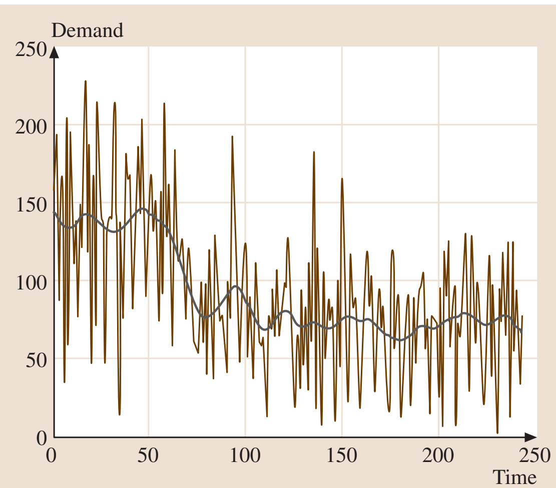

Fig. 14.3 The time-series plot and smooth curve for the quantity of blood platelets ordered from the Red Cross

Figure 14.3 shows the actual time-series data of the demand for platelets from the Red Cross from January 2002 to August 2002 and the smooth curve of the data. The smooth curve suggests that mean of the series has shifted down during the data-collection period. It is easy to visually identify the mean shift in this series. However, it is difficult to conclude that it is a mean shift as it is unfolding. This is the main issue: we want to detect the mean shift as soon as possible in real time. To detect a mean shift in seasonal autocorrelated data, we must use an appropriate time-series model of the original data. Following a three-step model-building process of model identification, model fitting, and diagnostic checking (*Box*, *Jenkins*, and *Reinsel* [14.23]), we find that an appropriate null model of the data is the ARIMA $(1,0,0)\times(0,1,1)_7$ seasonal model given by

$$\begin{aligned} a_{t0} &= Y_t\frac{\phi(B)}{\theta(B)} = Y_t\frac{(1-B^7)(1+0.281B)}{(1-0.833B^7)} \\ &= Y_t + 0.281Y_{t-1} - Y_{t-7} - 0.281Y_{t-8} \\ &\quad + 0.833a_{t0-7} \,. \end{aligned} \tag{14.17}$$

The discrepancy model is

$$\begin{aligned} a_t &= [Y_t - \gamma f(t)]\frac{\phi(B)}{\theta(B)} \\ &= [Y_t - \gamma f(t)]\frac{(1-B^7)(1+0.281B)}{(1-0.833B^7)} \\ &= Y_t + 0.281Y_{t-1} - Y_{t-7} - 0.281Y_{t-8} - \gamma f(t) \\ &\quad - 0.281\gamma f(t-1) + \gamma f(t-7) + 0.281\gamma f(t-8) \\ &\quad + 0.833a_{t-7} \,. \end{aligned}$$

The detector for the model is

$$\begin{aligned} d_t &= -\left.\frac{\partial a_t}{\partial \gamma}\right|_{\gamma=\gamma_0} \\ &= f(t) + 0.281f(t-1) - f(t-7) \\ &\quad - 0.281f(t-8) + 0.833d_{t-7} \,. \end{aligned} \tag{14.18}$$

Using (14.16) and (14.17) in the one-sided Cuscore statistic of (14.5b) and using a reference value with $\eta = \sigma_a = 31.58$ yields the results shown in Fig. 14.4. The figure also shows that control limits are approximately 196 and -196, which are based on (14.6) with $\alpha = 1/500$. Here the Cuscore chart signals a negative mean shift at observation 67, just two time periods later than the actual occurrence.

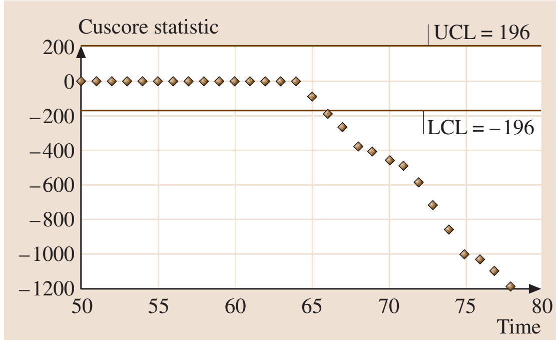

Fig. 14.4 A Cuscore chart for the Red Cross data

This example follows the best-case scenario, which is to predict the time of the occurrence of the mean shift at exactly the time that it really occurs, that is $\hat{t}_0 = t_0$. In such a case, there will be a match between the residuals and the detector, making the use of the Cuscore straightforward. In reality, we are unlikely to have prior information on when the mean shift will occur or, in terms of this application, when there will be a difference in the level of platelets ordered. Consequently, in the determination of the Cuscore statistic there will be a mismatch between the detector and the residuals. The mismatch case is considered fully for this application in *Nembhard* and *Changpetch* [14.21].

14.6 Cuscores in Process Monitoring and Control

As a second example of Cuscore in industry, we now consider a case from *Nembhard* and *Valverde-Ventura* [14.24] where cellular window blinds are produced using a pleating and gluing manufacturing process. Cellular shades form pockets of air that insulate windows from heat and cold. These shades start as 3000-yard rolls of horizontally striped fabric. On the machines, the fabric winds over, under and through several rollers, then a motorized arm whisks a thin layer of glue across it and a pleater curls it into a cell. When the process goes as planned, the crest of the pleat is in the center of the stripe and the finished product is white on the back and has a designer color on the front. When something goes wrong, defects can include a color that bleeds through to the other side, a problem known as "out of registration."

Using a high-speed camera, position data are acquired on the fabric every 20 pleats then a computer compares the edge of the colored band with the target position and measures the deviation (Fig. 14.5). If the two lines match then the deviation is zero and the blind is said to be "in-registration." If the lines do not match, a feedback controller is used to adjust the air cylinder pressure. Unfortunately, as can be seen from the displacement measurements in Fig. 14.5, the feedback controller performed very poorly.

To address this problem, we can use the Box–Jenkins transfer function plus noise and signal model in Fig. 14.6

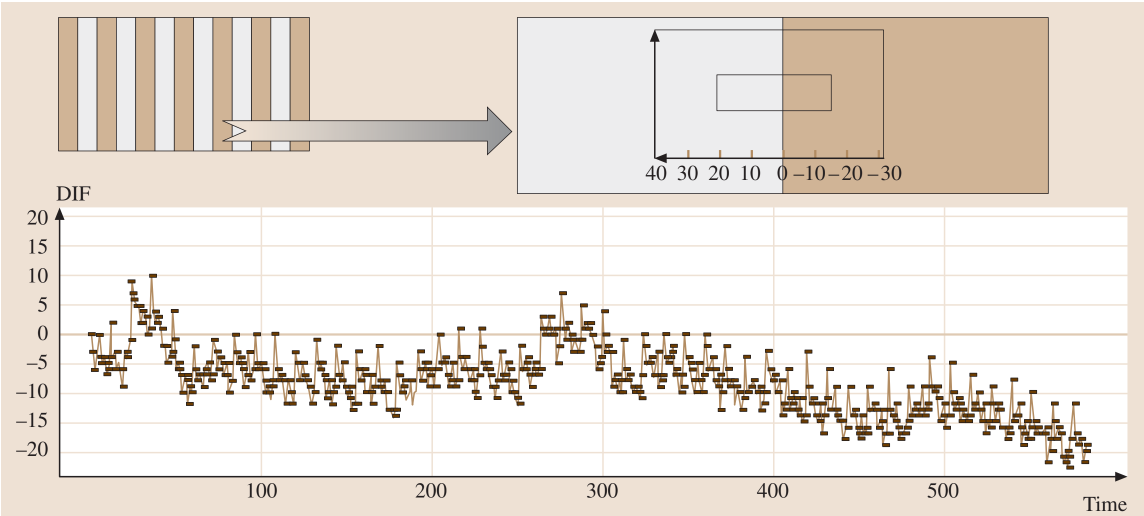

Fig. 14.5 Representation of the measurement of the displacement of the leading edge of the fabric with respect to a fixed point

for process representation. In this model, the output Y_t is the combination of the disturbance term that follows an ARIMA process, as we had in (14.1) and (14.2), plus an input (or explanatory) variable X_t, that is controllable but is affected by the process dynamics S_t. In this case, the combined model of the output in the presence of a signal is:

$$Y_t = \frac{L_2(B)}{L_1(B)} X_{t-k} + a_t \frac{\theta(B)}{\phi(B)} + \gamma\ ,\ f(t)$$

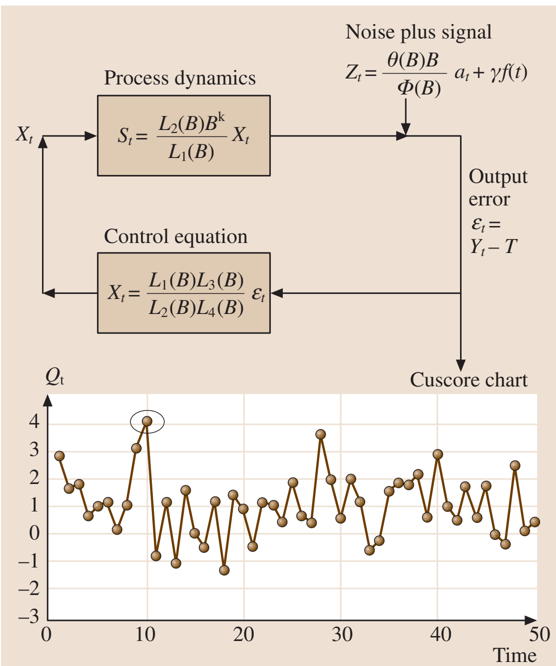

Fig. 14.6 A block diagram showing the input, output, and noise components and the relationship between feedback control and Cuscore monitoring of an anticipated signal

where $L_1(B)$ and $L_2(B)$ are the process transfer function polynomials. The control equation tells us how to change X_t over time based on the observed error ε_t. In addition to the process transfer polynomials, the control equation contains the polynomials $L_3(B)$, which describes the noise plus signal Z_t in terms of white noise, and $L_4(B)$, which describes the error ε_t in terms of white noise.

Assuming that minimum variance [or minimum mean-square error (MMSE)] control is applied, we have the null model

$$a_{t0} = \frac{1}{L_4(B)} \varepsilon_t\ . \tag{14.19}$$

The discrepancy model is

$$a_t = \frac{1}{L_4(B)} \varepsilon_t - \gamma f(t) \frac{\phi(B)}{\theta(B)}\ . \tag{14.20}$$

(See *Nembhard* and *Valverde-Ventura* [14.24] for a complete derivation of the null and discrepancy models.)

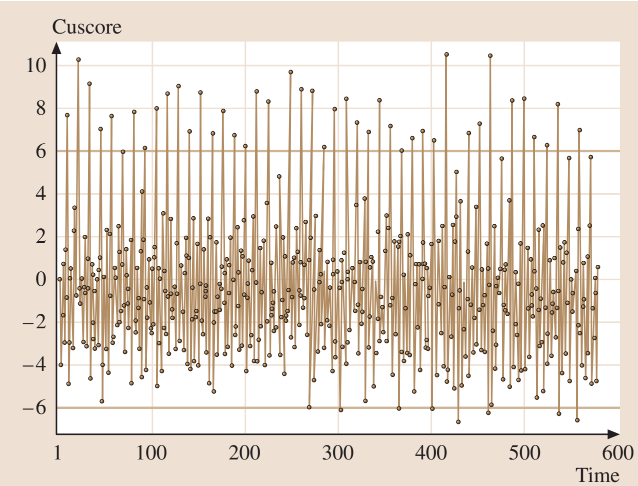

Fig. 14.7 Cuscore chart detects spike signals at every twelfth pleat

Notice that the noise disturbance and signal are assumed to occur after and independently of the process control.

Using (14.20), the detector is

$$d_i = -\frac{\partial a_t}{\partial \gamma}\bigg|_{\gamma=0} = f(t)\frac{\phi(B)}{\theta(B)} \,. \qquad (14.21)$$

Finally, using (14.4), the Cuscore statistic for detecting a signal $f(t)$ hidden in an ARIMA disturbance in an MMSE-controlled process (and omitting the reference value) is given by summing the product of equations (14.19) and (14.21):

$$Q_t = \sum \frac{1}{L_4(B)}\varepsilon_t \frac{\phi(B)}{\theta(B)} f(t) \,. \qquad (14.22)$$

For the special case when $k = 1$ (i. e., a responsive system), and the disturbance is white noise, (14.22) simply reduces to the output error, ε_t, which is equivalent to using a Shewhart chart. However, in this pleating and gluing process $k = 2$ and the spike is hidden in an integrated moving-average (IMA) (1, 1) disturbance. The appropriate Cuscore for this case is

$$Q_t = \frac{1}{1+0.84B}\varepsilon_t \,. \qquad (14.23)$$

We constructed the Cuscore chart in Fig. 14.7 using (14.23). In this application, during the null operation (i. e., when there is no signal) the Cuscore chart displays observations normally distributed with a mean of zero, and standard deviation $s\sigma_a$. At the moment the spike appears, the corresponding observation belongs to a normal distribution with mean of s^2 and standard deviation of $s\sigma_a$. This mean of s^2 gives the ability for us to observe the spike in the chart.

Note that the Cuscore chart identifies spike signals at pleat numbers 8, 20, 32, etc. In tracking down this problem, it appeared that the printing cylinder used by the supplier to print the fabric was the cause. In that process, the printing consists of passing the fabric over a screen roll with 12 channels. However, one of the twelve stripes had a different width, probably because the printing cylinder was not joined properly at the seam.

14.7 Discussion and Future Work

This chapter focuses on the development and application of Cuscore statistics. Since *Box* and *Ramírez* [14.10, 11] presented a design for the Cuscore chart, other work has been done to use them in time series. For example, *Box* and *Luceño* [14.12] suggested monitoring for changes in the parameters of time-series models using Cuscores. *Box* et al. [14.25] and *Ramirez* [14.26] use Cuscores for monitoring industry systems. *Luceño* [14.27] and *Luceño* [14.28] considered average run-length properties for Cuscores with autocorrelated noise. *Shu* et al. [14.20] designed a Cuscore chart that is triggered by a Cusum statistic and uses a generalized likelihood ratio test (GLRT) to estimate the time of occurrence of the signal. These statistical aids help the Cuscore to perform better. *Runger* and *Testik* [14.29] compare the Cuscore and GLRT. *Graves* et al. [14.30] considered a Bayesian approach to incorporating the signal that is in some cases equivalent to the Cuscore. *Harrison* and *Lai* [14.31] develop a sequential probability ratio test (SPRT) that outperforms the Cuscore for the limited cases of data similar to the t-distribution and distributions with inverse polynomial tails.

Although the statistical foundation can be traced back to *Fisher*'s efficient score statistic [14.13], it still needs further development to realize its true potential as a quality engineering tool. Accordingly, *Nembhard* and *Valverde-Ventura* [14.24] developed a framework that may help to guide the development and use of Cuscore statistics in industry applica-

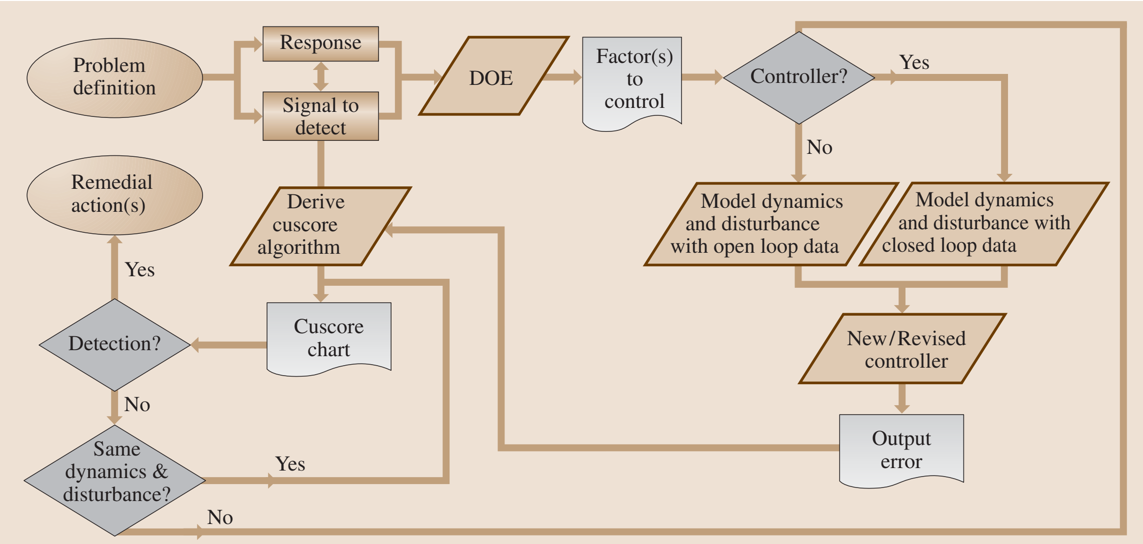

Fig. 14.8 Framework for using Cuscores with DOE and process control

tions, as shown in Fig. 14.8. This framework parallels the define, measure, analyze, improve, and control (DMAIC) approach used in Six Sigma (*Harry* and *Schroeder* [14.32]). The problem-definition step closely parallels the define step in DMAIC. Design of experiments (DOE) helps us to measure and analyze the process, the second two DMAIC steps. From DOE we develop an understanding of the factors to control, so we can then adjust and monitor in keeping with the last two DMAIC steps. The monitoring in this case is accomplished using a Cuscore chart.

The Cuscore is a natural fit with the DMAIC approach as it strives to incorporate what we learn about the problem into the solution. Some consideration needs to be given to the system to establish a clear understanding of the response, the expected signal to be detected, and the relationship between the two. More specifically, for the Cuscore to be applicable we should be able to describe how the signal might modify the response and, therefore, the output error.

This framework also recognizes that in many industrial systems, using only SPC to monitor a process will not be sufficient to achieve acceptable output. Real processes tend to drift away from target, use input material from different suppliers, and are run by operators who may use different techniques. For these and many other reasons, a system of active adjustment using engineering process control (EPC) is often necessary. *Box* and *Jenkins* [14.33] pioneered the integration of SPC and EPC to monitor and adjust industrial processes jointly by demonstrating the interrelationships between adaptive optimization, adaptive quality control, and prediction. *Box* and *Kramer* [14.34] revived the discussion on the complementary roles of SPC and EPC. Since then, many other authors have addressed the joint monitoring and adjustment of industrial processes. *Montgomery* and *Woodall* [14.35] give over 170 references in a discussion paper on statistically based process monitoring and control. Others since include *Shao* [14.36]; *Nembhard* [14.37]; *Nembhard* and *Mastrangelo* [14.38]; *Tsung*, *Shi*, and *Wu* [14.39]; *Tsung* and *Shi* [14.40]; *Ruhhal*, *Runger*, and *Dumitrescu* [14.41]; *Woodall* [14.42]; *Nembhard* [14.43]; *Nembhard*, *Mastrangelo*, and *Kao* [14.44]; and *Nembhard* and *Valverde–Ventura* [14.45]. The texts by *Box* and *Luceño* [14.12] and *del Castillo* [14.46] also address the topic.

In addition to those issues addressed in Sect. 14.5 for autocorrelated data, future work that will further advance the area of Cuscore statistics include their integration with suboptimal controllers, which are often used in practice. There is also a great need to expand the understanding of the robustness of Cuscores to detect signals (other than the one specifically designed for), to develop ways to detect multiple signals then identify or classify them once an out-of-control condition occurs, and to develop multivariate Cuscore detection capabilities.

References

14.1 W. E. Deming: *Out of the Crisis* (Center for Advanced Engineering Studies, Cambridge 1986)
14.2 W. A. Shewhart: Quality control charts, Bell Sys. Tech. J. **5**, 593–603 (1926)
14.3 Z. G. Stoumbos, M. R. Reynolds Jr., T. P. Ryan, W. H. Woodall: The state of statistical process control as we proceed into the 21st century, J. Am. Stat. Assoc. **451**, 992–998 (2000)
14.4 Western Electric: *Statistical Quality Control Handbook* (Western Electric Corp., Indianapolis 1956)
14.5 E. S. Page: Continuous inspection schemes, Biometrika **41**, 100–114 (1954)
14.6 G. A. Barnard: Control charts and stochastic processes, J. R. Stat. Soc. B **21**, 239–271 (1959)
14.7 S. W. Roberts: Control chart tests based on geometric moving averages, Technometrics **1**, 239–250 (1959)
14.8 J. S. Hunter: The exponentially weighted moving average, J. Qual. Technol. **18**, 203–210 (1986)
14.9 G. E. P. Box: Sampling and Bayes' inference in scientific modeling and robustness, J. R. Stat. Soc. A **143**, 383–430 (1980)
14.10 G. E. P. Box, J. Ramírez: *Sequential Methods in Statistical Process Monitoring: Sequential Monitoring of Models*, CQPI Report No. 67 (Univ. Wisconsin, Madison 1991)
14.11 G. E. P. Box, J. Ramírez: Cumulative score charts, Qual. Reliab. Eng. Int. **8**, 17–27 (1992). Also published as *Report No. 58* (Univ. Wisconsin, Madison 1992)
14.12 G. E. P. Box, A. Luceño: *Statistical Control by Monitoring and Feedback Adjustment* (Wiley, New York 1997)
14.13 R. A. Fisher: Theory of statistical estimation, Proc. Cambridge Philos. Soc. **22**, 700–725 (1925)
14.14 D. C. Montgomery: *Introduction to Statistical Process Control*, 5th edn. (Wiley, New York 2005)
14.15 L. Alwan, H. V. Roberts: Time-series modeling for statistical process control, J. Bus. Econ. Stat. **6**, 87–95 (1988)
14.16 A. V. Vasilopoulos, A. P. Stamboulis: Modification of control chart limits in the presence of correlation, J. Qual. Technol. **10**, 20–30 (1978)
14.17 D. C. Montgomery, C. M. Mastrangelo: Some statistical process control methods for autocorrelated data, J. Qual. Technol. **23**, 179–204 (1991)
14.18 C. M. Mastrangelo, D. C. Montgomery: SPC with correlated observations for the chemical and process industries, Qual. Reliab. Eng. Int. **11**, 79–89 (1995)
14.19 S. J. Hu, C. Roan: Change patterns of time series-based control charts, J. Qual. Technol. **28**, 302–312 (1996)
14.20 L. Shu, D. W. Apley, F. Tsung: Autocorrelated process monitoring using triggered cuscore charts, Qual. Reliab. Eng. Int. **18**, 411–421 (2002)
14.21 H. B. Nembhard, P. Changpetch: Directed monitoring of seasonal processes using cuscore statistics, Qual. Reliab. Eng. Int. , to appear (2006)
14.22 Franklin Institute Online: *Blood Platelets* (http://www.fi.edu/biosci/blood/platelet.html, 2004)
14.23 G. E. P. Box, G. M. Jenkins, G. C. Reinsel: *Time Series Analysis, Forecasting And Control*, 3rd edn. (Prentice Hall, Englewood Cliffs 1994)
14.24 H. B. Nembhard, R. Valverde-Ventura: A framework for integrating experimental design and statistical control for quality improvement in manufacturing, J. Qual. Technol. **35**, 406–423 (2003)
14.25 G. Box, S. Graves, S. Bisgaard, J. Van Gilder, K. Marko, J. James, M. Seifer, M. Poublon, F. Fodale: *Detecting Malfunctions In Dynamic Systems*, Report No. 173 (Univ. Wisconsin, Madison 1999)
14.26 J. Ramírez: Monitoring clean room air using cuscore charts, Qual. Reliab. Eng. Int. **14**, 281–289 (1992)
14.27 A. Luceño: Average run lengths and run length probability distributions for cuscore charts to control normal mean, Comput. Stat. Data Anal. **32**, 177–195 (1999)
14.28 A. Luceño: Cuscore charts to detect level shifts in autocorrelated noise, Qual. Technol. Quant. Manag. **1**, 27–45 (2004)
14.29 G. Runger, M. C. Testik: Control charts for monitoring fault signatures: Cuscore versus GLR, Qual. Reliab. Eng. Int. **19**, 387–396 (2003)
14.30 S. Graves, S. Bisgaard, M. Kulahci: A Bayes-adjusted cumulative sum. Working paper (2002)
14.31 P. J. Harrison, I. C. H. Lai: Statistical process control and model monitoring, J. Appl. Stat. **26**, 273–292 (1999)
14.32 M. Harry, R. Schroeder: *Six Sigma: The Breakthrough Management Strategy Revolutionizing the World's Top Corporations* (Random House, New York 2000)
14.33 G. E. P. Box, G. M. Jenkins: Some statistical aspects of adaptive optimization and control, J. R. Stat. Soc. B **24**, 297–343 (1962)
14.34 G. E. P. Box, T. Kramer: Statistical process monitoring and feedback adjustment—a discussion, Technometrics **34**, 251–285 (1992)
14.35 D. C. Montgomery, W. H. Woodall (Eds.): A discussion of statistically-based process monitoring and control, J. Qual. Technol. **29**, 2 (1997)
14.36 Y. E. Shao: Integrated application of the cumulative score control chart and engineering process control, Stat. Sinica **8**, 239–252 (1998)
14.37 H. B. Nembhard: Simulation using the state-space representation of noisy dynamic systems to determine effective integrated process control designs, IIE Trans. **30**, 247–256 (1998)

14.38 H. B. Nembhard, C. M. Mastrangelo: Integrated process control for startup operations, J. Qual. Technol. **30**, 201–211 (1998)

14.39 F. Tsung, J. Shi, C. F. J. Wu: Joint monitoring of PID-controlled processes, J. Qual. Technol. **31**, 275–285 (1999)

14.40 F. Tsung, J. Shi: Integrated design of run-to-run PID controller and SPC monitoring for process disturbance rejection, IIE Trans. **31**, 517–527 (1999)

14.41 N. H. Ruhhal, G. C. Runger, M. Dumitrescu: Control charts and feedback adjustments for a jump disturbance model, J. Qual. Technol. **32**, 379–394 (2000)

14.42 W. H. Woodall: Controversies and contradictions in statistical process control, J. Qual. Technol. **32**, 341–378 (2000)

14.43 H. B. Nembhard: Controlling change: process monitoring and adjustment during transition periods, Qual. Eng. **14**, 229–242 (2001)

14.44 H. B. Nembhard, C. M. Mastrangelo, M.-S. Kao: Statistical monitoring performance for startup operations in a feedback control system, Qual. Reliab. Eng. Int **17**, 379–390 (2001)

14.45 H. B. Nembhard, R. Valverde-Ventura: Cuscore statistics to monitor a non-stationary system. Qual. Reliab. Eng. Int., to appear (2006)

14.46 E. Del Castillo: *Statistical Process Adjustment For Quality Control* (Wiley, New York 2002)