QUALITY AND RELIABILITY ENGINEERING INTERNATIONAL
Qual. Reliab. Engng. Int. 2002; **18**: 411–421 (DOI: 10.1002/qre.492)

AUTOCORRELATED PROCESS MONITORING USING TRIGGERED CUSCORE CHARTS

LIANJIE SHU[1], DANIEL W. APLEY[2] AND FUGEE TSUNG[1]*
[1] *Department of Industrial Engineering and Engineering Management, The Hong Kong University of Science and Technology, Kowloon, Hong Kong*
[2] *Department of Industrial Engineering, Texas A&M University, College Station, TX 77843-3131, USA*

SUMMARY

Some of the most widely-investigated control charting techniques for autocorrelated data are based on time series residuals. If the mean shift in the autocorrelated process is a sudden step shift, the resulting mean shift in the residuals is time varying and has been referred to as the fault signature. Traditional residual based charts, such as a Shewhart, CUSUM, or EWMA on the residuals, do not make use of the information contained in the dynamics of the fault signature. In contrast, methods such as the Cuscore chart or Generalized Likelihood Ratio Test (GLRT) do incorporate this information. In order for the Cuscore chart to fully benefit from this, its detector coefficients should coincide with the fault signature. This is impossible to ensure, however, since the exact form of the fault signature depends on the time of occurrence of the mean shift, which is generally not known *a priori*. Any mismatch between the detector and the fault signature will adversely affect the Cuscore performance. This paper proposes a CUSUM-triggered Cuscore chart to reduce the mismatch between the detector and fault signature. A variation to the CUSUM-triggered Cuscore chart that uses a GLRT to estimate the mean shift time of occurrence is also discussed. It is shown that the triggered Cuscore chart performs better than the standard Cuscore chart and the residual-based CUSUM chart. Examples are provided to illustrate its use. Copyright © 2002 John Wiley & Sons, Ltd.

KEY WORDS: fault signature; stationary process; Cuscore; generalized likelihood ratio test; statistical process control

1. INTRODUCTION

Statistical control charts are widely used in industry for process monitoring and quality improvement. In the presence of data autocorrelation, however, traditional control charts may be rendered ineffective due to excessive false alarm rates (see, for example, Bagshaw and Johnson [1], Vasilopoulos and Stamboulis [2], Alwan [3], Harris and Ross [4], Montgomery and Woodall [5], and Apley and Tsung [6]). To deal with autocorrelated data, various control charting techniques have been proposed. The residual-based chart or the special cause chart (SCC) is one of the most widely investigated methods [7–13]. The key idea behind residual-based charts is to fit a time series model to represent the autocorrelation. If the model is adequate, the residuals (i.e. the one-step-ahead prediction errors) are approximately statistically independent. Thus, traditional control charts such as the Shewhart, CUSUM or exponentially weighted moving average (EWMA) charts, can be applied to the residuals.

An autoregressive moving average model, denoted as ARMA(p, q), is often used to represent the autocorrelation of the data. The general ARMA(p, q) process model is

$$x_t = \frac{\Theta(B)}{\Phi(B)} a_t \tag{1}$$

where x_t are observed data, a_t are independent and identically distributed (i.i.d.) normal variables with mean zero and variance σ_a^2, and B is the backshift operator, $Bx_t = x_{t-1}$. $\Phi(B)$ and $\Theta(B)$ are referred to

*Correspondence to: F. Tsung, Department of Industrial Engineering and Engineering Management, The Hong Kong University of Science and Technology, Kowloon, Hong Kong.
Email: season@ust.hk

Contract/grant sponsor: State of Texas Advanced Technology Program; contract/grant number: 000512-0287-2001
Contract/grant sponsor: National Science Foundation; contract/grant number: DMI-0093580
Contract/grant sponsor: RGC Competitive Earmarked Research; contract/grant number: HKUST6073/00E and KHUST6011/01E

Published online 13 August 2002

Received 1 November 2001
Revised 28 February 2002

as the AR and MA polynomial, and are parameterized as $\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$ and $\Theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$, respectively. Suppose there is a deterministic shift, which we refer to as a fault, in the process at some time τ. The process data to be monitored may be represented as

$$y_t = x_t + \mu f_{t-\tau}$$

where f_t indicates the nature of the fault, and μ is the fault magnitude. For a step mean shift, for example, $f_t = 1$ for $t \geq 0$ and 0 otherwise. For a sudden spike at a single time period, $f_t = 1$ for $t = 0$ and 0 otherwise. It is assumed that the model in (1) is invertible, in which case the residuals can be obtained by filtering y_t with the inverse filter, $\Phi(B)/\Theta(B)$. That is,

$$e_t = \frac{\Phi(B)}{\Theta(B)} y_t = a_t + \mu \tilde{f}_{t-\tau} \qquad (2)$$

where $\tilde{f}_{t-\tau} = (\Phi(B)/\Theta(B)) f_{t-\tau}$ is referred to as the fault signature [14]. Thus, the residuals are uncorrelated with time-varying mean $\mu \tilde{f}_{t-\tau}$ and variance σ_a^2. The value of $\tilde{f}_t$ depends on the ARMA model and, hence, on the autocorrelation structure of the data.

The information in the dynamics of the fault signature can be useful for detecting faults. Traditional residual-based charts do not make use of this information, however. In contrast, a generalized likelihood ratio test (GLRT) or a cumulative score (Cuscore) chart can take this information into account. Apley and Shi [14] demonstrated that the GLRT performs far better than a residual-based Shewhart or CUSUM chart for many models. The Cuscore test, which is a concept originally proposed by Fisher [15], was later developed by Bagshaw and Johnson [16], Box and Ramírez [17], Box and Luceño [18], and Ramírez [19]. The Cuscore test is intended to detect changes in the parameters of a statistical model, a natural application of which is control charting. In some sense the Cuscore chart is more easily implemented than the GLRT, due to its resemblance in form to the popular CUSUM control chart.

The one-sided Cuscore statistic is (see Appendix A)

$$Q_t = \max\{Q_{t-1} + r_t(e_t - k), 0\}, \qquad t = 1, 2, \ldots \quad (3)$$

where r_t and k are referred to as the detector and reference value, respectively. If Q_t exceeds a decision interval, h, it is concluded that a fault has occurred in the process. Suppose it is known *a priori* that *if* a fault is to occur, it will occur at some known time τ. It is shown in Appendix A that the detector should take the same value as the fault signature, i.e. $r_t = \tilde{f}_{t-\tau}$. In this case, as the $r_t(e_t - k)$ terms in (3) accumulate over time, they provide some measure of how closely the residuals correlate with the fault signature. If the fault does occur at the presumed time τ, the residuals are likely to correlate closely with the fault signature, leading to a greater likelihood of a signal in the Cuscore chart. Note that since $\tilde{f}_{t-\tau} = 0$ for $t < \tau$, using the detector $r_t = \tilde{f}_{t-\tau}$ is equivalent to initializing the Cuscore at observation τ.

In practice, however, the fault time-of-occurrence τ would generally not be known *a priori*. One potential strategy is to set $r_t = \tilde{f}_{t-1}$, which initializes the Cuscore on the first observation. This is in some sense equivalent to assuming that if a mean shift occurs, it will occur on the first observation $\tau = 1$. If the mean shift actually occurs at some time other than $\tau = 1$, the detector will be a time-shifted version of the fault signature, and may not correlate well with the residuals.

To illustrate this mismatch between the detector and the fault signature, two ARMA(1, 1) models with substantially different fault signatures are considered. Model 1 is chosen with $\phi_1 = 0.9$ and $\theta_1 = 0.5$, and Model 2 is chosen with parameters $\phi_1 = 0.45$ and $\theta_1 = -0.5$. Figures 1(a) and 1(b) show the detectors for detecting step mean shifts in Models 1 and 2, respectively, for Cuscore charts that are initialized on observations 1 and 11 (or, equivalently, assuming the mean shift time-of-occurrence is $\tau = 1$ and $\tau = 11$). Also shown are the residuals for a mean shift that actually occurred at observation $\tau = 11$. Clearly, the detector $r_t = \tilde{f}_{t-11}$ for the Cuscore chart initialized at the correct observation correlates much more closely with the residuals than does the detector $r_t = \tilde{f}_{t-1}$ for the Cuscore initialized at the first observation. It is reasonable to presume the Cuscore initialized at observation 11 would detect the shift much quicker than the Cuscore initialized at observation 1. Likewise, the opposite would be true if the mean shift actually occurred at observation $\tau = 1$. These arguments will be substantiated in subsequent sections of this paper.

In light of this, this article proposes what we refer to as the triggered Cuscore chart. The basic idea is to track some auxiliary 'trigger' statistic that is used purely to initialize or 'trigger' the Cuscore chart. In essence, the trigger statistic yields an estimate $\hat{\tau}$ of the mean shift time-of-occurrence. The Cuscore chart is then initialized at this time with detector $r_t = \tilde{f}_{t-\hat{\tau}}$. The expectation is that if $\hat{\tau}$ is close to the actual time-of-occurrence, the triggered Cuscore chart should outperform the standard Cuscore chart, initialized at the first observation.

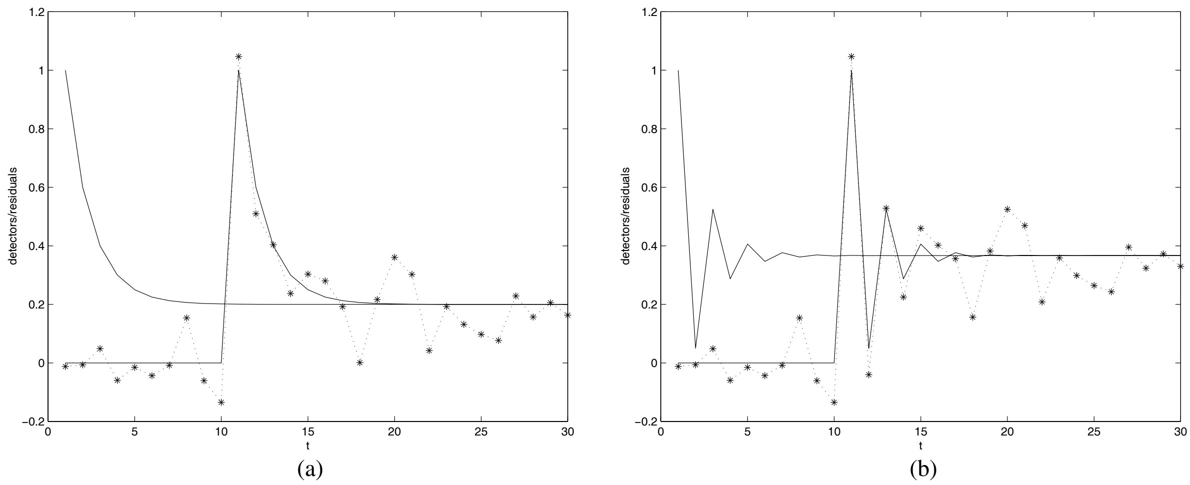

Figure 1. Two Cuscore detectors (shown as solid lines), initialized at timestep 1 and timestep 11, for step mean shifts in Model 1 (a) and Model 2 (b). The dashed line shows the residuals for a mean shift that actually occurred at timestep 11

The remainder of the article is outlined as follows. Section 2 describes the CUSUM-triggered Cuscore chart, which uses a CUSUM as the trigger statistic. In Section 3, the performance of the CUSUM-triggered Cuscore chart is compared with the standard Cuscore and residual-based CUSUM charts. An extension of the CUSUM-triggered Cuscore chart that uses a GLRT to estimate the mean shift time-of-occurrence discussed in Section 4. In Section 5, an example is used to illustrate the performance of the triggered Cuscore chart. Section 6 contains concluding remarks.

2. THE TRIGGERED CUSCORE CHART

As argued in the previous example, a mismatch between the Cuscore initialization time and the mean shift time-of-occurrence can severely affect the Cuscore chart performance. If the Cuscore chart is to benefit from the information represented by the fault signature dynamics, it should be triggered as closely as possible to the actual mean shift time-of-occurrence. A number of different methods for triggering the Cuscore were investigated. What seemed to be the most effective (and also one of the simplest) was a standard CUSUM on the residuals, of the form

$$S_t = \max\{0,\ S_{t-1} + e_t - k\}, \quad t = 1, 2, \ldots, t_{\text{trig}} \tag{4}$$

where k is the reference value, and t_{trig} is the observation which S_t first exceeds some decision interval H. We point out that the CUSUM is not used here to signal a mean shift. Rather, it is only used to trigger the Cuscore chart. Consequently, it will be referred to as the trigger CUSUM. When the trigger CUSUM first signals at time t_{trig}, the CUSUM statistic is traced backwards to the time period when it most recently rose above zero. This time period, denoted $\hat{\tau}$, is then taken to be a 'likely' time-of-occurrence for the mean shift and the Cuscore chart is initialized (retroactively) at time $\hat{\tau}$. From that point onwards, the trigger CUSUM is discontinued and the Cuscore

$$Q_t = \max\{0,\ Q_{t-1} + r_t(e_t - k)\}, \quad t = \hat{\tau}, \hat{\tau} + 1, \ldots \tag{5}$$

takes over. When Q_t exceeds its decision interval h, an out-of-control signal occurs.

To design the CUSUM-triggered Cuscore chart, three parameters must be specified: the reference value k (the same value is suggested for both the trigger CUSUM and the Cuscore), the decision interval H for the trigger CUSUM, and the decision interval h for the Cuscore chart. For a traditional CUSUM on i.i.d. data, a common rule-of-thumb is to select $k = \mu/2$, where μ is some mean shift magnitude of particular interest. H is then selected to provide a desired in-control average run length (ARL). This choice of k is nearly optimal, in terms of minimizing the out-of-control ARL. Since the time series residuals of an autocorrelated process have a time-varying mean, this rule-of-thumb is somewhat ambiguous in this case. In analogy with this guideline, however, we suggest using $k = \mu \tilde{f}_{\text{SS}}/2$, where

$$\tilde{f}_{\text{SS}} = \frac{1 - \phi_1 - \phi_2 - \cdots - \phi_p}{1 - \theta_1 - \theta_2 - \cdots - \theta_q}$$

is the steady-state magnitude of the fault signature after its initial transients have died out (refer to Hu and Roan [20] for details). Once the k value is determined, the decision interval H for the trigger CUSUM can be selected using standard methods to provide a desired in-control trigger ARL. Since the trigger CUSUM only initializes the Cuscore and does not actually signal a mean shift, the desired trigger ARL should be chosen much smaller than the overall ARL. In the examples of the subsequent section having overall in-control ARLs of 500, trigger ARLs of 30 to 50 were used. The final design step is to select the decision interval h for the Cuscore, which can be chosen to provide a desired overall in-control ARL.

3. SIMULATION RESULTS

In this section, simulations were performed to compare the ARL performance of three residual-based charts for detecting step mean shifts. The three charts are the standard Cuscore chart (3) initialized on the first observation (i.e. $r_t = \tilde{f}_{t-1}$), the standard residual-based CUSUM chart (4) used to directly signal a mean shift, and the CUSUM-triggered Cuscore chart (5) with the trigger CUSUM (4) for initializing the Cuscore. In the simulations, two cases are considered: (i) $\tau = 1$, and (ii) $\tau > 1$. In case (i) the detector for the standard Cuscore initialized on the first observation will match the fault signature, whereas in case (ii) it will not.

In all cases, Monte Carlo simulations with 25 000 replicates were used to investigate the ARL, which served as the performance measure. All charts were designed to have an in-control ARL of 500. For case (ii), the out-of-control ARL that was calculated is actually a conditional ARL. The conditional ARL is defined as the expected number of samples between when the mean shift first occurs and when it is first signaled by the chart, *given* that the chart did not signal prior to the occurrence of the mean shift. In other words, Monte Carlo replicates for which the chart signaled before the occurrence of the mean shift are discarded. One additional subtlety that was used in calculating the ARL for the CUSUM-triggered Cuscore should be pointed out. The time $\hat{\tau}$ at which the Cuscore is initialized will often be less than t_{trig}. Consequently, since the Cuscore chart is begun retroactively at time $\hat{\tau}$, it may also signal retroactively before time t_{trig}. In this event, the actual Cuscore signal time was taken to be t_{trig}.

All mean shifts in the simulations were step mean shifts with

$$f_t = \begin{cases} 0 & t < 0 \\ 1 & t \geq 0 \end{cases}$$

For the ARMA(1, 1) processes that were considered, the invertibility condition is $|\theta_1| < 1$. For a step mean shift in an invertible ARMA(1, 1) process, the fault signature is (Hu and Roan [20])

$$\tilde{f}_t = \begin{cases} 0 & t < \tau \\ 1 & t = \tau \\ 1 + \theta_1 - \phi_1 & t = \tau + 1 \\ \dfrac{1 - \phi_1 - \theta_1^{t-\tau}(\theta_1 - \phi_1)}{1 - \theta_1} & t \geq \tau + 2 \end{cases}$$

and the steady-state value of the fault signature reduces to

$$\tilde{f}_{SS} = \frac{1 - \phi_1}{1 - \theta_1}$$

For Models 1 and 2 the steady-state values of the fault signatures are 0.2 and 0.367, respectively. Without loss of generality, $\sigma_a^2 = 1$ was assumed in all simulations.

Comparisons for ARMA(1, 1) processes

Table 1 gives the ARLs of the three charts for detecting mean shifts in Model 1. All charts were designed to have in-control ARLs of approximately 500. For the CUSUM-triggered Cuscore chart, the trigger CUSUM was designed to have a trigger ARL of 50. A single reference value of $k = 0.15$ ($= 1.5 \sim f_{SS}/2$) was used throughout. Values of τ ranging from 1 to 41 were considered for case (ii). Specifically, for each Monte Carlo replicate, a different τ was generated from a uniform distribution over the range [2, 41], denoted $U[2, 41]$. Several observations can be made from Table 1 as follows.

- When $\tau = 1$, the detector for the standard Cuscore chart exactly matches the fault signature. Thus, the standard Cuscore chart performs the best for the larger mean shifts with magnitude 1.5 or greater. The CUSUM-triggered Cuscore chart performs better than the Cuscore chart for small mean shifts ($\mu \leq 1$). The residual-based CUSUM chart performs the worst.
- When $\tau \sim U[2, 41]$, the detector for the standard Cuscore chart no longer matches the actual fault signature. As a result, the performance of the standard Cuscore chart is adversely affected, although it still performs slightly better than the residual-based CUSUM chart. The CUSUM-triggered Cuscore chart performance is largely unaffected by having $\tau > 1$, and is substantially better than the standard Cuscore when $\tau \sim U[2, 41]$.

Table 2 gives the ARLs of the three charts for detecting mean shifts in Model 2. The reference value

Table 1. ARL's for the Cuscore, residual-based CUSUM, and CUSUM-triggered Cuscore charts for mean shifts at $\tau = 1$ and $\tau \sim U[2, 41]$ in Model 1

	$\tau = 1$			$\tau \sim U[2, 41]$		
μ	Cuscore $h = 2.0125$	CUSUM $H = 9.783$	Triggered cuscore $h = 2.4125$ $H = 4.08$	Cuscore $h = 2.0125$	CUSUM $H = 9.783$	Triggered cuscore $h = 2.4125$ $H = 4.08$
0.0	500.4	500.4	500.1	500.4	500.4	500.1
0.5	141.7	178.0	121.3	167.8	164.0	132.1
1.0	45.7	81.5	43.1	70.2	71.4	45.8
1.5	15.7	45.3	19.2	35.7	37.4	20.9
2.0	5.6	27.9	9.1	20.8	22.1	10.9
2.5	2.4	18.4	4.8	13.2	14.2	6.3
3.0	1.4	12.8	2.9	8.9	9.7	4.0

Table 2. ARLs for the Cuscore, residual-based CUSUM, and CUSUM-triggered Cuscore charts for mean shifts at $\tau = 1$ and $\tau \sim U[2, 41]$ in Model 2

	$\tau = 1$			$\tau \sim U[2, 41]$		
μ	Cuscore $h = 2.5145$	CUSUM $H = 6.827$	Triggered cuscore $h = 2.656$ $H = 3.19$	Cuscore $h = 2.5145$	CUSUM $H = 6.827$	Triggered cuscore $h = 2.656$ $H = 3.19$
0.0	500.3	499.7	499.6	500.3	499.7	499.6
0.5	104.0	104.7	96.8	105.6	105.1	98.2
1.0	33.4	37.0	32.8	37.2	37.3	32.8
1.5	14.5	18.9	15.7	19.0	19.1	16.4
2.0	7.5	12.0	8.9	12.0	12.1	10.0
2.5	4.2	8.4	5.4	8.5	8.6	6.9
3.0	2.5	6.4	3.5	6.5	6.5	5.1

was again chosen to be $k = 1.5 \sim f_{SS}/2$. Compared with Model 1, the fault signature for Model 2 has a larger steady-state value and oscillating transient dynamics (see Figure 1). The larger the steady-state fault signature, the less severe is the forecast recovery phenomena, and the more effective are residual-based charts at detecting mean shifts [11]. Also because the forecast recovery phenomenon is less severe for Model 2 than for Model 1, the adverse effects of $\tau > 1$ on the standard Cuscore chart are less severe. Regardless of the differences between the two models, the observations for Model 2 are qualitatively similar to those for model 1. When $\tau = 1$, the standard Cuscore chart performs the best for $\mu \geq 1.5$. When $\tau \sim U[2, 41]$, the performance of the standard Cuscore chart is adversely affected, and the CUSUM-triggered Cuscore chart performs the best. In both cases (i) and (ii), the residual-based CUSUM chart performs the worst.

Effect of trigger ARL

In the CUSUM-triggered Cuscore chart, the trigger CUSUM should be designed with a much lower trigger ARL than the overall in-control ARL. To investigate how the trigger ARL affects the performance of the CUSUM-triggered Cuscore chart, the trigger CUSUM is designed with three different ARLs: 50, 40, and 30. For Model 1, the H values that are required for the trigger CUSUM with $k = 0.15$ to have these three trigger ARLs are 4.08, 3.6564, and 3.1466.

Table 3 shows the overall ARLs of the CUSUM-triggered Cuscore charts with different trigger ARLs for Model 1. Rather than generating τ randomly, three specific values of τ were considered: $\tau = 1$, 11, and 41. As would be expected, a shorter trigger ARL results in slightly better detection performance for small τ and slightly worse detection for large τ.

Effect of the reference value

The choice of reference value k also affects the chart performance. Precise guidelines for the optimal reference value for the CUSUM-triggered Cuscore chart would be difficult to determine, given the number of quantities on which they would depend. It is shown below that the optimal reference value

Table 3. ARLs for CUSUM-triggered Cuscore charts with different trigger ARLs for Model 1

	Trigger ARL = 50			Trigger ARL = 40			Trigger ARL = 30		
	$H = 4.08$ $h = 2.4125$			$H = 3.6564$ $h = 2.38$			$H = 3.1466$ $h = 2.34$		
μ	$\tau = 1$	$\tau = 11$	$\tau = 41$	$\tau = 1$	$\tau = 11$	$\tau = 41$	$\tau = 1$	$\tau = 11$	$\tau = 41$
0.0	500.1	500.1	500.1	499.9	499.9	499.9	499.7	499.7	499.7
0.5	121.3	124.7	154.6	119.7	127.0	163.4	120.1	129.4	175.8
1.0	43.1	43.9	52.5	41.7	43.8	56.1	40.2	43.8	61.2
1.5	19.2	20.1	24.1	18.0	19.5	25.4	16.7	19.2	27.9
2.0	9.1	10.2	12.9	8.3	9.8	13.8	7.4	9.4	15.4
2.5	4.8	5.8	7.8	4.2	5.5	8.3	3.5	5.2	9.5
3.0	2.9	3.7	5.0	2.5	3.5	5.5	2.1	3.3	6.3

Table 4. ARLs of CUSUM-triggered Cuscore charts with various reference values for mean shifts in Model 1

μ	$k = 0.125$ $H = 4.312$ $h = 2.603$	$k = 0.15$ $H = 4.08$ $h = 2.4125$	$k = 0.175$ $H = 3.871$ $h = 2.2555$	$k = 0.20$ $H = 3.679$ $h = 2.119$
		$\tau = 1$		
0	500.3	500.1	500.3	500.5
0.5	118.8	121.3	123.7	126.9
1	43.9	43.1	42.8	43.1
1.5	20.2	19.2	18.6	18.2
2	9.8	9.1	8.7	8.4
2.5	5.1	4.8	4.5	4.3
3	3.1	2.9	2.8	2.6
		$\tau = 11$		
0	500.3	500.1	500.3	500.5
0.5	119.9	124.7	130.3	135.4
1	44.1	43.9	44.2	45.1
1.5	21.2	20.1	19.5	19.1
2	11.3	10.2	9.5	8.9
2.5	6.7	5.8	5.3	4.8
3	4.3	3.7	3.3	3.0
		$\tau = 41$		
0	500.3	500.1	500.3	500.5
0.5	142.1	154.6	167.7	178.4
1	50.5	52.5	55.9	59.3
1.5	23.6	24.1	24.1	25.0
2	13.4	12.9	12.6	12.5
2.5	8.2	7.8	7.4	7.1
3	5.5	5.0	4.7	4.4

for the CUSUM-triggered Cuscore chart depends not only on the fault signature (and thus the ARMA model) but also τ and μ. Due to the complexity of determining truly optimal reference values, the simple guideline $k = \mu \sim f_{SS}/2$, analogous to what is used for standard CUSUM charts in i.i.d. data, is recommended. Here, μ denotes a mean shift magnitude of particular interest.

Table 4 gives the ARLs of CUSUM-triggered Cuscore charts for Model 1 when different reference values are used. Once again, the trigger ARL was 50, and the overall in-control ARL was 500. One common result is that, just as for CUSUM charts on i.i.d. data, larger reference values are better for larger mean shifts and smaller reference values for smaller mean shifts. Another observation is that the optimal reference value depends on the exact time-of-occurrence of the mean shift. Consider, for example, a mean shift magnitude of $\mu = 1.0$. When $\tau = 1$, the optimal reference value for $\mu = 1.0$ is roughly 0.175. When $\tau = 11$ and

Table 5. ARLs of CUSUM-triggered Cuscore charts with a GLRT to estimate τ for mean shifts in Model 1

	Trigger ARL = 50			Trigger ARL = 40			Trigger ARL = 30		
	$H = 4.08$ $h = 2.6265$			$H = 3.6564$ $h = 2.5648$			$H = 3.1466$ $h = 2.489$		
μ	$\tau = 1$	$\tau = 11$	$\tau = 41$	$\tau = 1$	$\tau = 11$	$\tau = 41$	$\tau = 1$	$\tau = 11$	$\tau = 41$
0.0	500.5	500.5	500.5	500.1	500.1	500.1	500.1	500.1	500.1
0.5	106.3	106.0	145.2	105.0	108.8	158.4	104.8	112.8	177.2
1.0	37.1	34.4	45.2	35.9	34.3	49.2	34.8	34.3	57.2
1.5	17.6	14.8	19.5	16.2	14.0	21.4	14.9	13.7	25.2
2.0	8.8	7.1	10.1	7.9	6.5	11.2	6.9	6.2	13.4
2.5	4.8	3.7	5.9	4.1	3.4	6.6	3.5	3.1	8.2
3.0	2.9	2.3	3.8	2.6	2.1	4.4	2.1	2.0	5.6

$\tau = 41$, however, the optimal reference values are roughly 0.15 and 0.125, respectively.

4. THE CUSUM-TRIGGERED CUSCORE WITH A GLRT TO ESTIMATE τ

The suggested procedure for estimating the mean shift time-of-occurrence τ in the CUSUM-triggered Cuscore was to trace the trigger CUSUM statistic backwards to the observation on which it most recently rose above zero. Although this procedure is somewhat *ad hoc*, the previous examples demonstrate that the resulting CUSUM-triggered Cuscore performs superior to the standard Cuscore or the residual-based CUSUM. This section demonstrates that even better performance may be achieved if more sophisticated methods are used to estimate τ. In particular, we consider using a GLRT to estimate τ.

To describe the procedure, suppose the trigger CUSUM sounds an alarm at time t_{trig}, and that $\hat{\tau}$ is the estimated time-of-occurrence from tracing the CUSUM statistic backwards to the time at which it most recently rose above zero. In other words, the trigger CUSUM was last reset to zero at time $\hat{\tau} - 1$. The generalized likelihood ratio (GLR) statistic for testing the hypothesis that the mean shift occurred at time τ (for $\tau = \hat{\tau}, \hat{\tau} + 1, \ldots, t_{\text{trig}}$) is (Apley and Shi [14])

$$T(\tau) = \left(\sigma_a^2 \sum_{i=0}^{t_{\text{trig}}-\tau} \tilde{f}_i^2\right)^{-1/2} \sum_{i=0}^{t_{\text{trig}}-\tau} e_{\tau+i} \tilde{f}_i \qquad (6)$$

The GLRT estimate of the mean shift time-of-occurrence, denoted $\hat{\tau}_{\text{GLRT}}$, is the value of τ that maximizes (6), i.e.

$$\hat{\tau}_{\text{GLRT}} = \arg \max_{\hat{\tau} \le \tau \le t_{\text{trig}}} T(\tau) \qquad (7)$$

Table 5 shows the ARLs for the CUSUM-triggered Cuscore chart with a GLRT to estimate the time-of-occurrence of mean shifts in Model 1. Comparing these with the results in Table 3, the use of a GLRT provides a further improvement in the triggered Cuscore performance. It would, however, be slightly more complicated to implement.

5. AN ILLUSTRATIVE EXAMPLE

To compare and illustrate the use of the CUSUM-triggered Cuscore chart, the residual-based CUSUM, and the standard Cuscore chart, data for Model 1 were generated in simulation. The process was in-control for the first ten observations, and a mean shift of magnitude 1.5 was introduced at the 11th observation. The in-control ARL was set at 500, with a reference value of $k = 0.15$, for all three charts. The decision intervals for the CUSUM residual chart, the Cuscore chart, and the CUSUM-triggered Cuscore chart were as in Table 1. Table 6 shows the residuals generated from equation (2), as well as the fault signature and each of the charted statistics.

Figure 2(a) is a plot of the residuals. Figures 2(b), 2(c) and 2(d) show the residual-based CUSUM chart, the standard Cuscore chart, and the CUSUM-triggered Cuscore chart, respectively. Both the residual-based CUSUM and the standard Cuscore charts do not signal until observation 25, for a detection lag of 14 observations. For the CUSUM-triggered Cuscore chart, the trigger CUSUM first signals at observation 13. Tracing the trigger CUSUM backwards, it had last rose above zero at observation 8. Thus, the CUSUM-triggered Cuscore was initialized (retroactively) at observation 8. At observation 17, the CUSUM-triggered Cuscore first signaled an out-of-control condition, for a detection lag of only 6 observations.

The CUSUM-triggered Cuscore with a GLRT to estimate the mean shift time-of-occurrence was also considered. Since the trigger CUSUM first signaled at observation 13 and had last been reset to zero at

Table 6. Simulated data for a mean shift of magnitude 1.5 at observation 11 in Model 1

					CUSUM-triggered Cuscore	
t	e_t	$\tilde{f}_{t-\tau}$	CUSUM	Cuscore	CUSUM	Cusocre
1	−0.339	0	0	0.0000	0	—
2	0.033	0	0	0.0000	0	—
3	1.076	0	0.926	0.3704	0.926	—
4	0.214	0	0.99	0.3896	0.99	—
5	0.097	0	0.937	0.3764	0.937	—
6	−1.718	0	0	0.0000	0	—
7	−0.681	0	0	0.0000	0	—
8	1.622	0	1.472	0.3032	1.472	1.4720
9	−0.407	0	0.915	0.1902	0.915	1.1378
10	−0.166	0	0.599	0.1263	0.559	1.0114
11	1.264	1	1.713	0.3502	1.713	1.3456
12	1.7	0.6	3.263	0.6602	3.263	1.7331
13	1.46	0.4	4.573	0.9222	4.573	2.0279
14	0.028	0.3	4.451	0.8978	—	2.0020
15	1.627	0.25	5.928	1.1932	—	2.3063
16	0.275	0.225	6.053	1.2182	—	2.3316
17	0.676	0.212	6.579	1.3234	—	2.4379
18	0.877	0.206	7.306	1.4688	—	2.5840
19	1.784	0.203	8.94	1.7956	—	2.9108
20	0.906	0.202	9.696	1.9468	—	3.0620
21	−0.388	0.201	9.158	1.8392	—	2.9544
22	−0.941	0.2	8.067	1.6210	—	2.7362
23	0.372	0.2	8.289	1.6654	—	2.7806
24	1.212	0.2	9.351	1.8778	—	2.9930
25	1.593	0.2	10.794	2.1664	—	3.2816

observation 7, the GLR statistics (6) for $\tau = 8$ through 13 must be calculated. They were 10.86, 3.496, 5.70, 7.17, 4.29, and 1.46, respectively. Thus, $\hat{\tau}_{\mathrm{GLRT}} = 8$, which is the same as $\hat{\tau}$ from the trigger-CUSUM. When the GLRT is used to estimate τ, the decision interval $h = 2.6265$ (see Table 5) for the Cuscore is slightly larger than when the GLRT is not used. Thus, from Table 6, the Cuscore statistic will not cross the decision interval until observation 19. Consequently, in this example, the CUSUM-triggered Cuscore chart performs slightly better when the GLRT is not used.

Note that in this example, $\hat{\tau}_{\mathrm{GLRT}} = \hat{\tau} = 8$ differs from the actual $\tau = 11$ (although the second largest GLR statistic, 7.17, is for the correct $\tau = 11$). In spite of this, initializing the Cuscore at *approximately* the mean shift time-of-occurrence seems to result in better detection than if the Cuscore is initialized on the first observation.

6. CONCLUDING REMARKS

As the Cuscore statistic is intended to detect changes in the parameters of a statistical model, its applications are quite broad and include SPC control charting. This paper develops a Cuscore chart to monitor for mean shifts in autocorrelated processes. A primary consideration is that since the fault signature is time-varying, the detector may not exactly match the fault signature in the standard Cuscore chart. To reduce this mismatch, the triggered Cuscore chart was proposed. The triggered Cuscore chart requires an estimate of the fault time-of-occurrence in order to initialize the Cuscore chart at approximately the correct time. This study indicates that the CUSUM chart or the CUSUM with a GLRT to estimate the fault time-of-occurrence are reasonably effective triggering mechanisms. Guidelines for selecting the reference value k and the trigger ARL, as well as examples illustrating the use of the triggered Cuscore chart, have been provided.

Although we have only discussed monitoring step mean shifts in autocorrelated processes, the triggered Cuscore chart can easily be extended to other scenarios. For example, if one is interested in detecting sudden spikes in the data, it is straightforward to extend the triggered Cuscore chart to this scenario. In this case, all that is required is to calculate the fault signature of the sudden spike and use this as

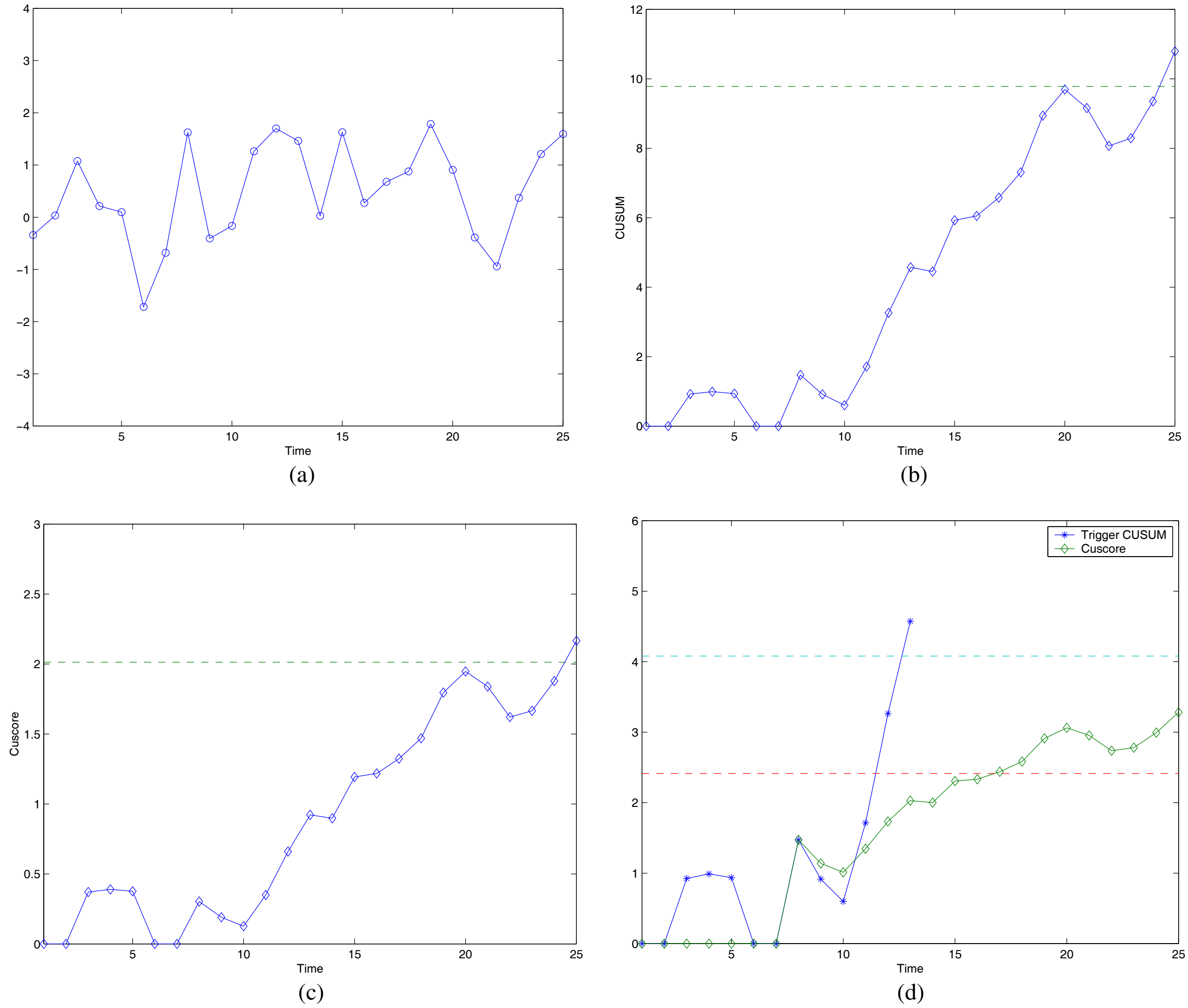

Figure 2. An illustrative example: (a) residuals for a mean shift of magnitude 1.5 occurring at observation 11 in Model 1; (b) residual-based CUSUM chart; (c) standard Cuscore chart; (d) CUSUM-triggered Cuscore chart

the detector. In addition, the Cuscore chart of this paper is 'one-sided' in the sense that it applies to detecting positive mean shifts. To monitor for negative mean shifts, a lower one-sided Cuscore defined via $Q_t^- = \max\{Q_{t-1}^- + r_t(-e_t - k), 0\}$ would be used.

ACKNOWLEDGEMENTS

The authors are grateful to the editor and the referees for their valuable comments. D. Apley was supported in this research by the State of Texas Advanced Technology Program under grant 000512-0287-2001 and the National Science Foundation under grant DMI-0093580. F. Tsung was supported by RGC Competitive Earmarked Research Grants HKUST6073/00E and HKUST6011/01E.

APPENDIX A. THE CUSCORE STATISTIC FOR MONITORING AUTOCORRELATED PROCESSES

The standard Cuscore statistic is a monitoring procedure designed to detect changes in the parameters of a statistical model. Consider a statistical model written in the following form

$$a_t = f(y_t, \theta) \tag{A.1}$$

where y_t is the observation, and θ is some unknown parameter. The standard normal theory models assume that the a_t are i.i.d. normal with zero mean and variance σ^2 when θ is the true value of the unknown parameters. The Cuscore statistic associated with the parameter value $\theta = \theta_0$ is

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

where the a_{i0} values are obtained by setting $\theta = \theta_0$ in equation (A.1) and

$$r_i = -\frac{\partial a_i}{\partial \theta}\bigg|_{\theta=\theta_0}$$

which is usually referred to as the detector (Box and Luceño [18]). Consequently, the Cuscore statistic can be treated as sequentially cumulated products of a null residual a_{i0} and an appropriate detector r_i.

The Cuscore developed above focuses on detecting changes in a parameter of interest, but it can also be used for SPC purposes. From equation (2), we have $a_t = e_t - \mu \tilde{f}_{t-\tau}$. When there is no mean shift fault, it is assumed that $\mu = 0$. When there is a step mean shift fault, the residuals have a mean of $\mu \tilde{f}_{t-\tau}$. Detecting the mean shift is equivalent to detecting the change of μ from its zero value, or more precisely, looking for $\tilde{f}_{t-\tau}$ in the residuals. The detector for detecting the change of μ from its zero value is

$$r_i = -\frac{\partial e_i}{\partial \mu}\bigg|_{\mu=0} = \tilde{f}_{i-\tau}$$

The value of a_{i0} is

$$a_{i0} = a_i|_{\mu=0} = e_i$$

Thus, the Cuscore statistic for monitoring the mean of normally-distributed, autocorrelated processes is

$$Q_t = \sum_{i=1}^{t} a_{i0} r_i = \sum_{i=1}^{t} e_i r_i \tag{A.2}$$

where $r_i = \tilde{f}_{i-\tau}$.

For computational purposes, it is convenient to rewrite the Cuscore in (A.2) in the recursive form:

$$Q_t = Q_{t-1} + r_t e_t$$

As pointed out by Box *et al.* [21], however, this Cuscore cannot be used directly. The reason is that most systems begin with a long period of the in-control state, during which Q_t would generally drift further and further away from the threshold. When a fault occurs, Q_t cannot reach the threshold in a timely fashion. They suggested correcting for this by using the one-sided Cuscore

$$Q_t = \max\{Q_{t-1} + r_t e_t, 0\}$$

Furthermore, a reference value can be chosen in the Cuscore to improve the detection performance. The Cuscore chart with reference value k_t at time t is defined as

$$Q_t = \max\{Q_{t-1} + r_t(e_t - k_t), 0\} \tag{A.3}$$

This statistic is referred to as a one-sided upper Cuscore. If Q^+ exceeds the threshold h, it is concluded that a mean shift fault has occurred in the process.

REFERENCES

1. Bagshaw M, Johnson RA. The effect of serial correlation on the performance of CUSUM tests II. *Technometrics* 1975; **17**:73–80.
2. Vasilopoulos AV, Stamboulis AP. Modification of control chart limits in the presence of data correlation. *Journal of Quality Technology* 1978; **10**:20–30.
3. Alwan LC. Effects of autocorrelation on control chart performance. *Communications in Statistics—Theory and Methods* 1992; **21**:1025–1049.
4. Harris TJ, Ross WH. Statistical process control procedures for correlated observations. *Canadian Journal of Chemical Engineering* 1991; **69**:48–57.
5. Montgomery DC, Woodall WH. A discussion on statistically-based process monitoring and control. *Journal of Quality Technology* 1997; **29**:121–162.
6. Apley DW, Tsung F. The autoregressive T-squared chart for monitoring univariate autocorrelated processes. *Journal of Quality Technology* 2002; **34**:80–96.
7. Alwan LC, Roberts HV. Time series modeling for statistical process control. *Journal of Business and Economic Statistics* 1988; **6**:87–95.
8. Montgomery DC, Mastrangelo CM. Some statistical process control methods for autocorrelated data. *Journal of Quality Technology* 1991; **23**:179–193.
9. Lin WS, Adams BM. Combined control charts for forecast-based monitoring schemes. *Journal of Quality Technology* 1996; **28**:289–301.
10. Lu CW, Reynolds MR. EWMA control charts for monitoring the mean of autocorrelated processes. *Journal of Quality Technology* 1999; **31**:166–188.
11. Superville CR, Adams BM. An evaluation of forecast-based quality control schemes. *Communications in Statistics Part B: Simulation and Computation* 1994; **23**(3):645–661.
12. Vander Wiel SA. Monitoring processes that wander using integrated moving average models. *Technometrics* 1996; **38**:139–151.
13. Wardell DG, Moskowitz H, Plante RD. Run-length distributions of special-cause control charts for correlated observations. *Technometrics* 1994; **36**:3–17.
14. Apley DW, Shi JJ. The GLRT for statistical process control of autocorrelated processes. *IIE Transactions* 1999; **31**:1123–1134.
15. Fisher RA. Theory of statistical estimation. *Proceedings of the Cambridge Philosophical Society* 1925; **22**:700–725.
16. Bagshaw M, Johnson RA. Sequential procedures for detecting parameter changes in a time-series model. *Journal of the American Statistical Association* 1977; **72**:593–597.
17. Box GEP, Ramírez JG. Cumulative score charts. *Quality and Reliability Engineering International* 1992; **8**:17–27.
18. Box GEP, Luceño A. *Statistical Control by Monitoring and Feedback Adjustment*. Wiley: New York, 1997.
19. Ramírez JG. Monitoring clean room air using cuscore charts. *Quality and Reliability Engineering International*, 1998; **14**:281–289.

20. Hu SJ, Roan C. Change patterns in the time series-based control charts. *Journal of Quality Technology* 1996; **28**:302–312.
21. Box GEP, Graves S, Bisgaard S, Gilder JV, Marko K, James J, Seifer M, Poublon M, Fodale F. Detecting malfunctions in dynamic systems. *CQPI Report No.173*, University of Wisconsin, Madison, The Center for Quality and Productivity Improvement, 1999.

Authors' biographies:

Lianjie Shu received his BS degree in Mechanical Engineering and Automation in 1998, from Xi'an Jiao Tong University. He is currently a PhD candidate in the Department of Industrial Engineering and Engineering Management at the Hong Kong University of Science and Technology (HKUST). He has been involved in a number of research projects on quality control. His research interests include statistical process control and applied statistics. He is a member of the Institute for Operations Research and the Management Sciences (INFORMS).

Daniel Apley received his BS and MS degrees in Mechanical Engineering, MS degree in Electrical Engineering, and PhD degree in Mechanical Engineering in 1990, 1992, 1995, and 1997, respectively, all from the University of Michigan. He was a visiting researcher with the Institute for Machine Tools and Factory Management at the Technical University of Berlin, in 1995. From 1997 to 1998 he was a post-doctoral fellow with the Department of Industrial and Operations Engineering at the University of Michigan. Since 1998, he has been with Texas A&M University, where he is currently an Assistant Professor of Industrial Engineering. Dr Apley's research interests are manufacturing variation reduction through in-process measurement and model-based statistical monitoring, diagnosis, and automatic control. His current research activities are sponsored by Ford, Solectron, the National Science Foundation, and the State of Texas Advanced Technology Program. He was an AT&T Bell Laboratories PhD Fellow from 1993 to 1997 and received the NSF CAREER award in 2001. He is a member of IIE, IEEE, ASME, INFORMS, and SME.

Fugee Tsung is an assistant professor of Industrial Engineering and Engineering Management at HKUST. He received both his MS and PhD in Industrial and Operations Engineering from the University of Michigan, Ann Arbor, and his BSc in Mechanical Engineering from the National Taiwan University. He worked for the Ford Motor and Rockwell International and did his post-doctoral research with Chrysler. He has been elected as Secretary/Treasurer of the Quality, Statistics, and Reliability (QSR) Section of the Institute for Operations Research and the Management Sciences (INFORMS). He is also a senior member of IIE and ASQ, and is an ASQ Certified Six Sigma Black Belt. His current research interests include quality engineering and management, process control and monitoring, and Six Sigma implementation.