IIE Transactions (2008) **40**, 992–1003

ISSN: 0740-817X print / 1545-8830 online
DOI: 10.1080/07408170801961412

Adaptive CUSUM procedures with EWMA-based shift estimators

WEI JIANG[1,*], LIANJIE SHU[2] and DANIEL W. APLEY[3]

[1]*Department of Systems Engineering & Engineering Management, Stevens Institute of Technology, Hoboken, NJ 07030, USA*
E-mail: wjiang@stevens.edu
[2]*Faculty of Business Administration, University of Macau, Taipa, Macau*
[3]*Department of Industrial Engineering & Management Sciences, Northwestern University, Evanston, IL 60208, USA*

Received April 2006 and accepted October 2007

Adaptive Cumulative SUM charts (ACUSUM) have been recently proposed for providing an overall good detection over a range of mean shift sizes. The basic idea of the ACUSUM chart is to first adaptively update the reference value based on an Exponentially Weighted Moving Average (EWMA) estimate and then to assign a weight on it using a certain type of weighting function. A linear weighting function is proposed that is motivated by likelihood ratio testing concepts and that achieves superior detection performance. Moreover, in view of the lower efficiency in tracking relative large mean shifts of the EWMA estimate, a generalized EWMA estimate is proposed as an alternative. A comparison of run length performance of the proposed ACUSUM scheme and other control charts is shown to be favorable to the former.

Keywords: Average run length, change-point detection, EWMA, Markov chain

1. Introduction

Process control charts have been extensively used to track process changes in areas as diverse as quality control, navigation system monitoring, signal segmentation, etc. (Banzal and Papantoni-Kazakos, 1986; Basseville and Nikiforov, 1993; Siegmund, 1985). Suppose that $\{X_t\}$ is an independent sequence following the normal distribution $N(\mu_t, 1)$, where $\mu_t = \delta$ when $t \geq t_0$, and is zero elsewhere. δ and t_0 represent the magnitude and occurrence time of the shift, respectively, both of which are generally unknown. The objective of process monitoring is to detect the mean shift δ as early as possible following its occurrence.

The Cumulative SUM (CUSUM) chart proposed by Page (1954) is one of the most popular algorithms for accomplishing this, in part due to certain Average Run Length (ARL) optimality properties (Lorden, 1971; Pollak, 1985; Moustakides, 1986). It can be viewed as a series of sequential probability ratio tests (Wald, 1947) of whether the process mean has changed to some specific value δ_0. The one-sided upper CUSUM chart for detecting a positive mean shift ($\delta_0 > 0$) is implemented by recursively calculating:

$$C_t^+ = \max(0, C_{t-1}^+ + (X_t - k)), \tag{1}$$

where $k = \delta_0/2$ is often referred to as the reference value of CUSUM schemes. The chart signals as soon as C_t^+ exceeds a decision interval h. The initial value is usually $C_0^+ = 0$, although it may be specified otherwise for a fast initial response (Lucas and Crosier, 1982). Analogously, to detect a negative mean shift ($\delta_0 < 0$), the lower CUSUM can be calculated as

$$C_t^- = \min(0,\ C_{t-1}^- - (X_t - k)), \tag{2}$$

which signals when $C_t^- < -h$.

A drawback of the CUSUM scheme is that it is designed to detect a particular mean shift size and may perform far from optimal for detecting shifts of sizes smaller or larger than δ_0. The Exponentially Weighted Moving Average (EWMA) chart (Roberts, 1959; Lucas and Saccucci, 1990) possesses a similar drawback. To provide good ARL performance over a range of mean shift sizes, various alternatives have been suggested. One alternative is to use multiple control charts simultaneously, each optimized for a different size mean shift (see, for example, Lorden and Eisenberger (1973), Lucas (1982), Sparks (2000) and Zhao *et al.* (2005)). A second alternative is to use a single adaptive control chart, such as the Adaptive CUSUM (ACUSUM) chart and the Adaptive EWMA (AEWMA) chart, see, for example, Sparks (2000), Capizzi and Masarotto (2003) and Wu *et al.* (2008). The adaptive charts are perhaps simpler

*Corresponding author

to implement than the multiple-chart schemes in the sense that they are comprised of only a single chart.

The one-sided upper ACUSUM chart of Sparks (2000) is of the form (the lower part is similarly defined):

$$S_t^+ = \max\left(0, S_{t-1}^+ + \left[X_t - \frac{\hat{\delta}_t}{2}\right] \Big/ h(\hat{\delta}_t)\right), \qquad (3)$$

where $\hat{\delta}_t$ is a suitable estimate of the current mean, and $h(\delta)$ is an operating function that defines the control limit of the CUSUM statistic in Equation (1) for detecting a shift of size δ. Sparks (2000) uses a regression method to obtain a limited number of expressions for $h(\delta)$. Shu and Jiang (2006) develop a two-dimensional Markov chain model for the ACUSUM chart and show that $h(\delta)$ can be closely approximated by

$$h(\delta) \approx \frac{\ln[1 + \delta^2 \times \mathrm{ARL}_0/2 + 1.166\delta]}{\delta} - 1.166, \qquad (4)$$

where ARL_0 represents the in-control ARL for the upper CUSUM chart with $k = \delta/2$. Sparks (2000) shows that the ACUSUM chart can provide an overall good performance for signalling mean shifts over a range of sizes when δ_0 is unknown, compared to the traditional CUSUM chart.

Note that the term $1/h(\hat{\delta}_t)$ in Equation (3) can be viewed as a certain weight assigned to the deviations $\{X_t - \hat{\delta}_t/2\}$. Thus, the ACUSUM chart can be generalized as

$$W_t^+ = \max\left(0, W_{t-1}^+ + w(\hat{\delta}_t)\left[X_t - \frac{\hat{\delta}_t}{2}\right]\right),$$

where $w(\hat{\delta}_t)$ represents a weight function. Other weight functions could be used in the ACUSUM chart. In this paper, we consider using a different weight function $w(\hat{\delta}_t) = \hat{\delta}_t$. This weight function is conceptually appealing because it is motivated by likelihood ratio testing concepts, as shown in Appendix A. Figure 1 compares the two weight functions as a function of δ. It is clear that the former used in Sparks (2000) is curvilinear while the latter is linear. In addition, it will be shown later that using the linear weight function can provide superior performance to the traditional ACUSUM chart using $w(\hat{\delta}_t) = 1/h(\hat{\delta}_t)$.

Moreover, the EWMA scheme has been used in Sparks (2000) to estimate the current process mean level. It is known that the EWMA scheme is inefficient in capturing abrupt process mean changes of moderate and large magnitudes. The best estimating procedures, according to Yashchin (1995), cannot be linear if they are to adapt to changes of large magnitude because the inertia—a long-term measurement of estimation error—increases as the magnitude of the shift increases. To alleviate this limitation, Yashchin (1995) suggests using a non-linear estimator called the EWMA-C estimator, which is a generalization of the EWMA statistic. Thus, it is expected that the ACUSUM chart using an EWMA-C mean estimator improves the same chart using an EWMA estimator.

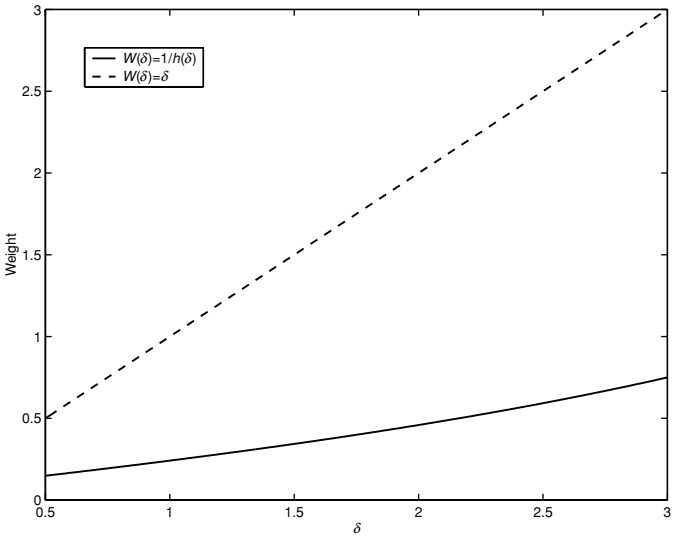

Fig. 1. Sparks' weight function and the linear weight function.

The purpose of this paper is to extend Sparks' ACUSUM chart in the following respects.

1. We recommend using a linear weight function $w(\hat{\delta}_t) = \hat{\delta}_t$ because it achieves superior performance. Moreover, it is conceptually appealing because it is loosely based on the Generalized Likelihood Ratio Test (GLRT).
2. We incorporate the EWMA-C estimator into the traditional ACUSUM chart to further improve its performance at large shifts.
3. We give some guidelines to design the ACUSUM chart.

The remainder of the paper is organized as follows. In Section 2, the ACUSUM procedure based on an EWMA-C estimator is presented. In Section 3, the proposed chart is compared with other control charts, including the CUSUM, Sparks' ACUSUM, combined Shewhart-CUSUM and AEWMA charts. In Section 4, some design guidelines are provided. In Section 5, an example is given to illustrate the implementation of the proposed chart. Section 6 concludes the paper.

2. The ACUSUM procedure based on an EWMA-C mean estimator

For illustration purposes, we focus on the upper ACUSUM chart for detecting positive mean changes. There are many different ways to estimate the current process mean level (see, for example, Barnard (1959), Chernoff and Zack (1964), Yashchin (1995) and Chen and Elsayed (2000)). Sparks (2000) considers using the EWMA statistic:

$$\hat{\delta}_t = (1 - \lambda)\hat{\delta}_{t-1} + \lambda X_t, \qquad (5)$$

to estimate the process mean in the ACUSUM chart due to its simplicity and efficiency, where $0 < \lambda \leq 1$. The initial

value is often set to $\hat{\delta}_0 = 0$. The EWMA scheme with a small value of λ is sensitive to small mean shift but less sensitive to relatively large mean shifts. To improve the estimation efficiency at large mean shifts, Yashchin (1995) suggests the following generalization of the EWMA statistic:

$$\hat{\delta}_t = \hat{\delta}_{t-1} + \phi(e_t), \tag{6}$$

where $e_t = X_t - \hat{\delta}_{t-1}$ is the prediction error, and $\phi(\cdot)$ is a monotone score function. This generalization is Markovian and reduces to the traditional EWMA statistic when $\phi(e) = \lambda e$.

To ensure that the procedure (6) tracks large shifts quickly, Yashchin (1995) proposes using Huber's score function, which is defined as

$$\phi_\gamma(e) = \begin{cases} e + (1-\lambda)\gamma, & e < -\gamma, \\ \lambda e, & |e| \le \gamma, \\ e - (1-\lambda)\gamma, & e > \gamma, \end{cases} \tag{7}$$

where $\gamma \ge 0$ is a constant. The Markovian-type statistic (6) with the Huber function is referred to as an EWMA-C statistic and includes the EWMA statistic as a special case when $\gamma \to \infty$. When $\gamma = 0$ or $\lambda = 1$, Huber's function reduces to $\phi(e) = e$, and the EWMA-C statistic is essentially a Shewhart statistic, X_t. It is interesting to note that the EWMA-C statistic performs similar to an EWMA statistic when prediction errors are small and similar to a Shewhart statistic when prediction errors are large. Although other score functions could be analyzed, we only investigate Huber's score function (7) in this work.

Because the upper ACUSUM chart is intended to detect upward mean shifts, we do not want to substitute negative values of $\hat{\delta}_t$ into Equation (3). Moreover, there is typically a minimum size mean shift that is of practical importance, say $\delta^+_{\min} > 0$. For these reasons, Sparks (2000) suggests using $\delta^+_{\min}$ instead of $\hat{\delta}_t$ whenever $\hat{\delta}_t < \delta^+_{\min}$. This leads to the following upper ACUSUM statistic:

$$Z^+_t = \max(0, Z^+_{t-1} + w(\hat{\delta}^+_t)(X_t - \hat{\delta}^+_t/2)), \tag{8}$$

where

$$\hat{\delta}^+_t = \max(\delta^+_{\min}, \hat{\delta}_t).$$

The initial value Z^+_0 is set at zero, and an upward shift is signaled whenever Z^+_t exceeds a threshold h. Analogously, the lower one-sided ACUSUM chart issues an alarm when:

$$Z^-_t = \min(0, Z^-_{t-1} - w(\hat{\delta}^-_t)(X_t - \hat{\delta}^-_t/2)), \tag{9}$$

is less than $-h$, where $\hat{\delta}^-_t = \min(\delta^-_{\min}, \hat{\delta}_t)$ and $\delta^-_{\min} < 0$.

Sparks' ACUSUM chart corresponds to the ACUSUM statistic in Equation (8) using the weight function $w(\hat{\delta}^+_t) = 1/h(\hat{\delta}^+_t)$ and the EWMA mean estimator in Equation (5). The ACUSUM chart defined in Equation (8) using a linear weight function $w(\hat{\delta}^+_t) = \hat{\delta}^+_t$ and the EWMA-C estimator in Equation (6) will be denoted as an ACUSUM-C chart for simplicity throughout the remainder of this paper.

3. Performance comparisons

In this section, the ACUSUM-C chart is first compared with conventional CUSUM and Sparks's ACUSUM charts. Then, it is compared to other competitive procedures, including the combined Shewhart–CUSUM chart and the AEWMA chart.

The Markov chain approach, integral equation approach and Monte Carlo simulations have been widely used to study the run length performance of control charts. In this section, the run length performance of ACUSUM-C charts is investigated using a Markov chain model. As shown in Appendix B, the ACUSUM-C chart can be represented by the two-dimensional Markov random vector $(Z^+_t, \hat{\delta}_t)'$. This allows one to apply a two-dimensional extension of the univariate Markov chain approach developed by Brook and Evans (1972) to calculate the ARL. From our experience, the Markov chain approximation agrees very well with Monte Carlo simulations.

3.1. *Comparison with CUSUM and Sparks' ACUSUM charts*

In order to maintain the advantages of CUSUM charts for detecting small shifts while improving the sensitivity for detecting large shifts, $\delta^+_{\min}$ usually takes small values in the ACUSUM procedures to guarantee the sensitivity to small shifts (Sparks, 2000). For illustration purposes, Tables 1 and 2 present ARL values of the ACUSUM-C charts with different γ values when; (i) $\delta^+_{\min} = 0.5$, $\lambda = 0.2$; and (ii) $\delta^+_{\min} = 1$, $\lambda = 0.3$, respectively. Both the zero-state ARL and the steady-state ARL are considered. The zero-state ARL refers to the ARL computed assuming a process change occurs at the initial stage of a control chart. The steady-state ARL is the ARL obtained after the control chart statistic has run for some time before a shift takes place. The steady-state analysis allows us to provide more complete performance comparisons.

The same values $\delta^+_{\min}$ and λ are used for Sparks' ACUSUM chart. The two CUSUM charts with $k = \delta^+_{\min}/2$ and $k = 1.5$ are also included, which are designed to detect a small shift of $\delta = \delta^+_{\min}$ and a large shift value of three, respectively. All charts are designed to have a zero-state ARL_0 of 400. The same values of control limits were used for computing both zero-state and steady-state out-of-control ARL (ARL_1) values.

Since the EWMA estimator is a special case of the EWMA-C estimator with $\gamma \to \infty$, the zero-state/steady-state ARL performance of the ACUSUM-C is almost identical for $\gamma \ge 4$, as shown in Tables 1 and 2. This is because when γ takes a large value, the estimation errors are very unlikely to exceed γ in Equation (7), and the EWMA-C estimator is essentially the same as the regular EWMA estimator. Thus, the ACUSUM-C chart with $\gamma = \infty$ in Tables 1 and 2 corresponds to an ACUSUM chart using the EWMA estimator.

Table 1. ARL of ACUSUM-C charts with $\delta^+_{\min} = 0.5$ and $\lambda = 0.2$.

		ACUSUM-C						*Sparks'*	*CUSUM*	
δ	*Type*	γ *=1.5* $h = 6.056$	*2* *5.105*	*2.5* *4.633*	*3* *4.430*	*4* *4.348*	∞ *4.327*	*ACUSUM* *1.152*	k *= 0.25* *6.86*	*1.5* *1.387*
0.00	Zero state	399.68	400.22	399.20	399.33	**399.97**	399.90	400.09	399.69	400.1
	Steady state	386.45	389.60	388.94	388.91	**389.39**	389.38	389.31	388.32	399.8
0.25	Zero state	67.19	67.34	65.51	64.08	**63.34**	63.33	63.86	64.34	168.68
	Steady state	54.06	56.00	54.95	53.84	**53.23**	53.29	53.80	54.13	168.33
0.50	Zero state	**26.73**	**25.66**	**24.72**	**24.17**	**23.91**	**23.87**	**24.03**	**24.27**	**74.87**
	Steady state	18.46	18.15	17.59	17.22	**17.04**	17.03	17.09	17.12	74.43
0.75	Zero state	15.50	14.66	14.13	13.85	**13.72**	13.68	13.82	14.10	35.52
	Steady state	10.24	9.79	9.41	9.20	**9.11**	9.09	9.11	9.13	34.49
1.00	Zero state	10.47	9.91	9.63	9.48	**9.42**	9.39	9.53	9.88	18.31
	Steady state	6.90	6.55	6.31	6.18	**6.12**	6.10	6.14	6.21	17.74
1.50	Zero state	5.90	5.68	5.65	5.66	**5.67**	5.66	5.78	6.20	6.50
	Steady state	4.03	3.88	3.79	3.74	**3.71**	3.70	3.75	3.90	5.97
2.00	Zero state	3.86	3.76	3.84	3.93	**4.01**	4.01	4.12	4.56	3.29
	Steady state	2.78	2.73	2.73	2.72	**2.72**	2.71	2.76	2.94	2.91
2.50	Zero state	2.76	2.70	2.80	2.94	**3.08**	3.10	3.20	3.65	2.13
	Steady state	2.12	2.10	2.14	2.17	**2.19**	2.19	2.23	2.42	1.90
3.00	Zero state	2.10	2.05	2.13	2.29	**2.47**	2.54	2.62	3.06	1.59
	Steady state	1.73	1.71	1.76	1.82	**1.86**	1.87	1.90	2.10	1.47
3.50	Zero state	1.69	1.64	1.69	1.82	**2.04**	2.15	2.24	2.65	1.30
	Steady state	1.48	1.46	1.50	1.56	**1.63**	1.65	1.69	1.89	1.25
4.00	Zero state	1.40	1.36	1.39	1.49	**1.71**	1.88	1.98	2.33	1.14
	Steady state	1.31	1.28	1.31	1.37	**1.46**	1.51	1.54	1.72	1.12
5.00	Zero state	1.10	1.08	1.09	1.12	**1.26**	1.45	1.58	2.01	1.02
	Steady state	1.09	1.07	1.08	1.11	**1.21**	1.31	1.37	1.52	1.02

It is interesting to compare the ACUSUM-C chart with $\gamma = \infty$ and Sparks' ACUSUM chart. Note that both charts use the same estimator, i.e., the EWMA statistic, to estimate the process mean level, but the ACUSUM-C chart with $\gamma = \infty$ uses a linear weight function of $w(\hat{\delta}_t) = \hat{\delta}_t$ whereas Sparks' ACUSUM chart uses a non-linear one, $w(\hat{\delta}_t) = 1/h(\hat{\delta}_t)$. From Tables 1 and 2, the ACUSUM-C chart with $\gamma = \infty$ provides better performance than Sparks' ACUSUM chart, especially when $\delta^+_{\min}$ is small and the shift size is large. This indicates that the linear weight function derived from the GLRT is better than the non-linear one used in Sparks (2000).

Moreover, the benefit of introducing the additional parameter γ for the ACUSUM chart can also be seen from Tables 1 and 2. In particular, the parameter γ can be chosen for the ACUSUM-C chart to achieve a relatively large percentage improvement in the ARL performance at large shifts while only causing slight loss in the efficiency in detecting small shifts. For example, as shown in Table 1, when γ decreases from $\gamma = \infty$ to $\gamma = 1.5$, the relative percentage increase in the ARL of the ACUSUM-C chart at a small shift of, say, $\delta = 0.5$ is 11.9% ((26.73 − 23.87)/23.87) but the relative decrease in the ARL at a large shift of, say, $\delta = 5$, can be as large as 24% ((1.45 − 1.1)/1.45). This demonstrates the benefit of introducing the additional parameter γ for the ACUSUM-C chart. This also agrees with our expectation of the superiority of EWMA-C estimators over the conventional EWMA statistic in tracking moderate and large mean changes.

Although large values of $\delta^+_{\min}$ can also improve the detection of large shifts, they would significantly hurt the detection of small shifts in practice. This can be easily seen from the tables where the ARL increases from 63.86 to 85.75 when $\delta^+_{\min}$ increases from one-half to one. Therefore, we won't recommend larger values of $\delta^+_{\min}$ than one in practice. Instead, small values of γ are suggested to improve the detection of large shifts. Detailed design guidelines will be provided in Section 4.

A very surprising observation from Tables 1 and 2 is that the ACUSUM-C chart may perform better even than the optimal CUSUM chart at the specified mean shift. For example, as can be seen from Table 1, at the mean shift with a value of one-half, the CUSUM chart with $k = 0.25$ takes 24.27 observations on average to detect this shift in the zero state. The ACUSUM-C chart with $\gamma \geq 4$ in Table 1 takes less than 24 steps in the zero state. This observation seems to contradict some of the established results on the ARL optimality of the one-sided CUSUM for detecting step mean shifts in Gaussian processes, in particular Moustakides (1986). However, this can be explained by the worst-case ARL.

Table 2. ARL of ACUSUM-C charts with $\delta^+_{\min} = 1$ and $\lambda = 0.3$.

		ACUSUM-C						*Sparks'*	*CUSUM*	
		$\gamma = 1.5$	2	2.5	3	4	∞	*ACUSUM*	$k = 0.5$	1.5
δ	*Type*	$h = 5.050$	4.730	4.505	4.394	4.337	4.334	1.0355	4.173	1.387
0.00	Zero state	399.70	400.85	400.19	399.29	399.39	399.97	399.37	399.87	400.1
	Steady state	396.17	397.25	396.35	395.27	395.22	395.79	395.13	395.39	399.8
0.25	Zero state	92.82	91.65	88.96	87.02	85.81	85.80	85.75	85.87	168.68
	Steady state	88.28	87.15	84.43	82.47	81.23	81.23	81.15	81.14	168.33
0.50	Zero state	30.52	30.10	29.32	28.79	28.46	28.45	28.45	28.49	74.87
	Steady state	26.14	25.83	25.12	24.63	24.32	24.32	24.30	24.27	74.43
0.75	Zero state	14.70	14.50	14.20	14.00	13.89	13.88	13.88	13.93	35.52
	Steady state	11.28	11.14	10.89	10.73	10.63	10.63	10.62	10.60	34.99
1.00	Zero state	**9.07**	**8.96**	**8.81**	**8.72**	**8.67**	**8.66**	**8.67**	**8.73**	**18.31**
	Steady state	6.57	6.47	6.35	6.27	6.22	6.22	6.21	6.20	17.74
1.50	Zero state	4.89	4.87	4.84	4.83	4.83	4.82	4.84	4.92	6.50
	Steady state	3.52	3.48	3.43	3.39	3.37	3.37	3.37	3.39	5.97
2.00	Zero state	3.23	3.25	3.28	3.31	3.35	3.34	3.36	3.46	3.29
	Steady state	2.44	2.43	2.42	2.41	2.40	2.40	2.41	2.43	2.91
2.50	Zero state	2.36	2.39	2.44	2.49	2.58	2.57	2.59	2.70	2.13
	Steady state	1.90	1.91	1.92	1.92	1.93	1.93	1.93	1.97	1.90
3.00	Zero state	1.84	1.86	1.91	1.97	2.12	2.11	2.12	2.26	1.59
	Steady state	1.58	1.59	1.61	1.63	1.66	1.66	1.66	1.71	1.47
3.50	Zero state	1.50	1.51	1.56	1.62	1.80	1.80	1.80	1.98	1.30
	Steady state	1.37	1.38	1.40	1.43	1.49	1.49	1.49	1.55	1.25
4.00	Zero state	1.28	1.28	1.31	1.36	1.55	1.55	1.55	1.77	1.14
	Steady state	1.22	1.23	1.25	1.28	1.37	1.37	1.37	1.45	1.12
5.00	Zero state	1.05	1.05	1.06	1.08	1.18	1.18	1.18	1.37	1.02
	Steady state	1.05	1.05	1.06	1.08	1.15	1.15	1.15	1.27	1.02

Moustakides (1986) showed that out of the class of all possible stopping rules that are constrained to have a specified in-control ARL, the CUSUM chart of Equation (1) minimizes the "worst-case" out-of-control ARL. The worst-case ARL for the one-sided CUSUM is precisely its zero-state ARL. On the other hand, the worst-case ARL for the ACUSUM-C chart is not necessarily its zero-state ARL with $\hat{\delta}_0 = 0$. The variable $\hat{\delta}_{t_0-1}$ could apparently take on an extreme value that would cause the worst-case out-of-control ARL of the ACUSUM to exceed that of the CUSUM. Note that the increment $\hat{\delta}^+_{t_0}(X_{t_0} - \hat{\delta}^+_{t_0}/2)$ in Equation (8) could be negative when $\hat{\delta}^+_{t_0-1}$ and consequently $\hat{\delta}^+_{t_0}$ are large enough. Because $\delta^+_{\min}$ is used in the ACUSUM whenever $\hat{\delta}_t < \delta^+_{\min}$, the worst-case value for $\hat{\delta}_{t_0-1}$ would most likely be large and positive. Regardless, it is with very small probability that $\hat{\delta}_{t_0-1}$ would drift to a value that causes the ARL to be substantially larger than the zero-state ARL. In fact, Table 1 indicates that the steady-state ARL is typically much shorter than the zero-state ARL. Hence, an ARL substantially larger than the zero-state ARL would most likely be an event of rather small probability.

Even though the zero-state ARL improvement of the ACUSUM is relatively small for the mean shift for which the standard CUSUM has been optimized. This observation has important implications. Many assume that for detecting step mean shifts in independent Gaussian processes, one need not look further than the CUSUM, because of its purported ARL optimality. Our results indicate that it may be possible to find a chart that has better zero-state, or even steady-state, ARL performance than the CUSUM, at the expense of somewhat worse worst-case performance.

3.2. *Comparison with other control charts*

As shown above, the ACUSUM procedure is effective at detecting a broad range of mean shifts. Because the combined CUSUM–Shewhart scheme investigated by Lucas (1982) is also intended to provide good detection of both large and small shifts, it is worthwhile to compare the performances of ACUSUM and combined CUSUM–Shewhart charts. Figure 2 plots the ARL of the ACUSUM-C chart and the combined CUSUM–Shewhart chart for detecting shifts over the range [0.5, 4]. Lucas (1982) suggests the parameters for the CUSUM chart with $k = 0.25$ and $h = 6$, and the Shewhart limit of four. The ACUSUM-C parameters are chosen as $\delta^+_{\min} = 2k = 0.5$, $\lambda = 0.2$, $\gamma = 2.5$ and $h = 3.901$. The in-control zero-state ARLs are maintained as 250. As shown in Fig. 2, the ACUSUM-C chart performs better than the combined Shewhart-CUSUM chart. This is consistent with the observation made in Shu and Jiang (2006).

Table 3. Comparison of zero-state ARL between ACUSUM-C and AEWMA charts ($ARL_0 = 500$).

	Chart parameters			δ										
	δ^+_{min}	λ	γ	*0.25*	*0.50*	*0.75*	*1.00*	*1.5*	*2.0*	*2.5*	*3.0*	*3.5*	*4.0*	*5.0*
Range [0.5, 4]														
AEWMA*	—	0.0398	2.899	115.30	36.71	20.08	13.53	7.77	4.96	3.25	2.19	1.58	1.26	1.04
ACUSUM-C*	0.5	0.0398	2.899	96.34	31.47	17.66	12.18	7.40	5.15	3.75	2.79	2.12	1.67	1.18
AEWMA$^+$	—	0.2	2.5	260.79	79.05	29.31	14.68	6.44	3.87	2.63	1.91	1.47	1.23	1.03
ACUSUM-C$^+$	0.5	0.2	2.5	95.54	31.98	17.76	11.90	6.84	4.58	3.30	2.49	1.95	1.59	1.17
Range [1, 4]														
AEWMA*	—	0.1253	2.7765	168.55	45.02	19.64	11.69	6.17	4.00	2.79	2.03	1.55	1.26	1.04
ACUSUM-C*	1.0	0.1253	2.7765	147.49	39.25	17.42	10.57	5.81	3.99	3.00	2.37	1.91	1.57	1.17
AEWMA$^+$	—	0.3	2.5	276.61	93.6	35.27	16.66	6.52	3.75	2.54	1.87	1.46	1.23	1.03
ACUSUM-C$^+$	1.0	0.3	2.5	153.52	40.58	17.87	10.7	5.72	3.82	2.8	2.17	1.75	1.45	1.11

Note: AEWMA* charts are optimally designed for the specified range according to Capizzi and Masarotto (2003).

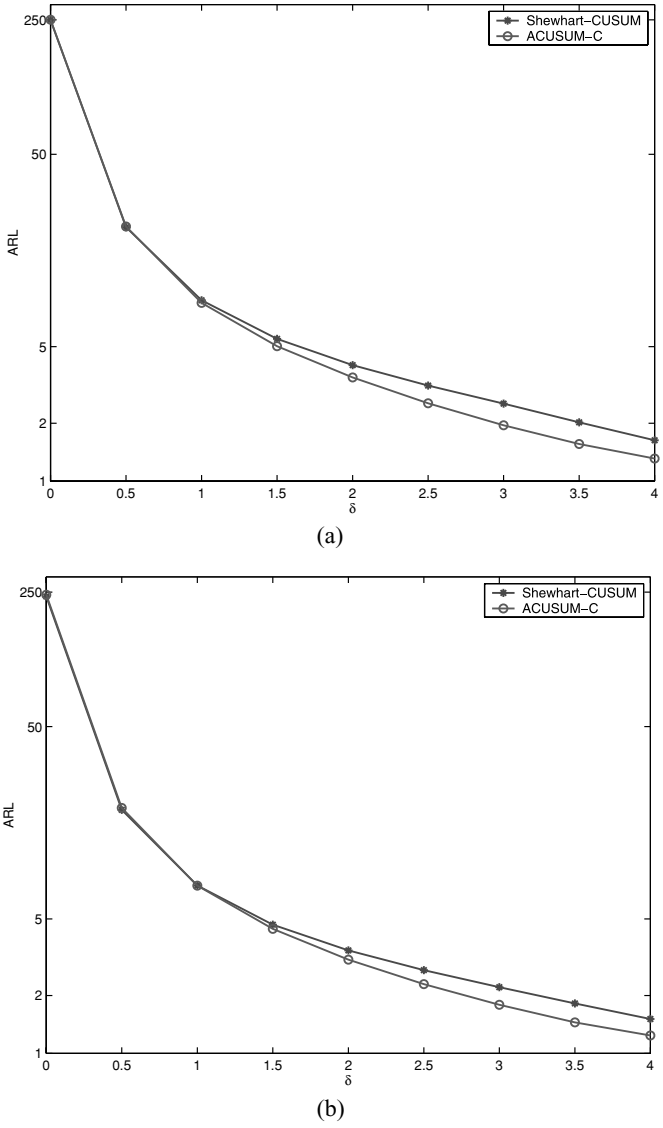

Fig. 2. ARL comparison between the ACUSUM-C and combined Shewhart–CUSUM charts: (a) zero state; and (b) steady state.

Another alternative for detecting a wide range of mean shifts efficiently is running multiple CUSUM charts simultaneously, as discussed in Lorden and Eisenberger (1973), Sparks (2000), Zhao *et al*. (2005), and others. In fact, the combined Shewhart–CUSUM scheme can be viewed as a two-CUSUM procedure, because the Shewhart chart is the limiting case of a CUSUM as the reference value increases. Sparks (2000) shows that monitoring three CUSUM charts simultaneously provides a comparable or slightly better performance than their version of the ACUSUM, but is less robust than their ACUSUM for detecting ramp mean shifts.

In parallel to the ACUSUM chart, the AEWMA chart has also been recently proposed by Capizzi and Masarotto (2003) for providing an overall good detection performance over a range of mean shifts. The AEWMA chart is based on the EWMA-C statistic in Equation (6). Namely, the two-sided AEWMA chart signals when $|\hat{\delta}_t| > h$. Capizzi and Masarotto (2003) show that the AEWMA chart is more effective than traditional EWMA charts for detecting a wide range of mean shifts. It is interesting to compare the detection performance of the AEWMA and ACUSUM-C charts.

Because the AEWMA statistic is inherently two-sided, Table 3 compares ARL values of AEWMA charts and two-sided ACUSUM-C charts. The AEWMA charts were optimized for detecting mean shifts over the ranges [1, 4] and [0.5, 4], respectively, and the ACUSUM-C charts use the corresponding EWMA-C statistics for mean estimation. According to Capizzi and Masarotto (2003), the optimal parameters for the AEWMA charts are $\lambda = 0.1253$ and $\gamma = 2.7765$ for the first range while $\lambda = 0.0398$ and $\gamma = 2.899$ for the second range, when ARL_0 equals 500. Interestingly, the ACUSUM-C chart (ACUSUM* in Table 3) is always more sensitive than its counterpart AEWMA chart (AEWMA*) for detecting small shifts (e.g., less than two) but less sensitive for detecting large shifts.

Interestingly, AEWMA charts appear to be very sensitive to parameter selection, while ACUSUM-C charts are relatively more robust. For example, when both charts are

designed to have $\lambda = 0.2$ for the first range of mean shifts (or 0.3 for the second range) and $\gamma = 2.5$, the AEWMA performance deteriorates severely for detecting small shifts, while the ACUSUM chart has no significant changes (denoted by AEWMA$^+$ and ACUSUM$^+$ respectively in Table 3). This indicates that AEWMA charts are very sensitive to parameter selections and a careful trade-off is thus crucial in order to balance the sensitivities against various levels of mean shifts. ACUSUM-C charts, on the other hand, exhibit a relatively more stable performance than AEWMA charts if different parameters are used. Nonetheless, the ACUSUM-C chart has not been optimized for detecting any ranges of shifts in this comparison.

4. Discussion of parameter selections

In addition to the prespecification of $\delta^+_{\min}$, the design of an ACUSUM-C chart involves the choice of parameters λ, γ and h. In this section, we provide some guidelines for the choice of parameters of ACUSUM-C charts. To provide a complete understanding of the parameter effects on the performance of ACUSUM-C charts, both the individual and overall effects of λ and γ are discussed.

4.1. *Parameter effects on ARL*

It is relatively easy to investigate the individual effect of the parameters λ and γ on the performance of ACUSUM-C charts. Generally, since small values of γ in the EWMA-C estimating procedure tend to improve the estimation efficiency for large shifts (Yashchin, 1995), they are also expected to improve the detection performance of the corresponding ACUSUM-C chart at detecting large shifts. As can be observed from Table 1, the ACUSUM-C chart with $\gamma = 2.5$ always has smaller zero-state ARL_1 values than the chart with $\gamma = 4$ for $\delta \geq 1.5$ and larger ARL_1 values for $\delta \leq 1$. On the other hand, it is well known that small λ values improve the estimation efficiency of EWMA/EWMA-C schemes for small mean shifts while large λ values improve the estimation efficiency for large shifts (Yashchin, 1995).

In summary, small values of λ and/or large values of γ tend to improve the detection performance at small mean shifts while large values of λ and/or small values of γ tend to improve the detection performance of ACUSUM-C charts at large mean shifts. However, the joint effect of λ and γ on the run length performance is more complicated since it depends on both of them.

To better understand the overall effect of the two parameters, Fig. 3 shows contour plots of zero-state ARL_1 values of ACUSUM-C charts as a function of λ and γ when $\delta^+_{\min} = 1.0$. When detecting a small mean shift of $\delta = 1$, as shown in Fig. 3(a), the ARL_1 decreases as λ decreases or γ increases. An extreme case is when $\lambda = 0$ and $\gamma \to \infty$ for which the ACUSUM-C chart reduces to the CUSUM

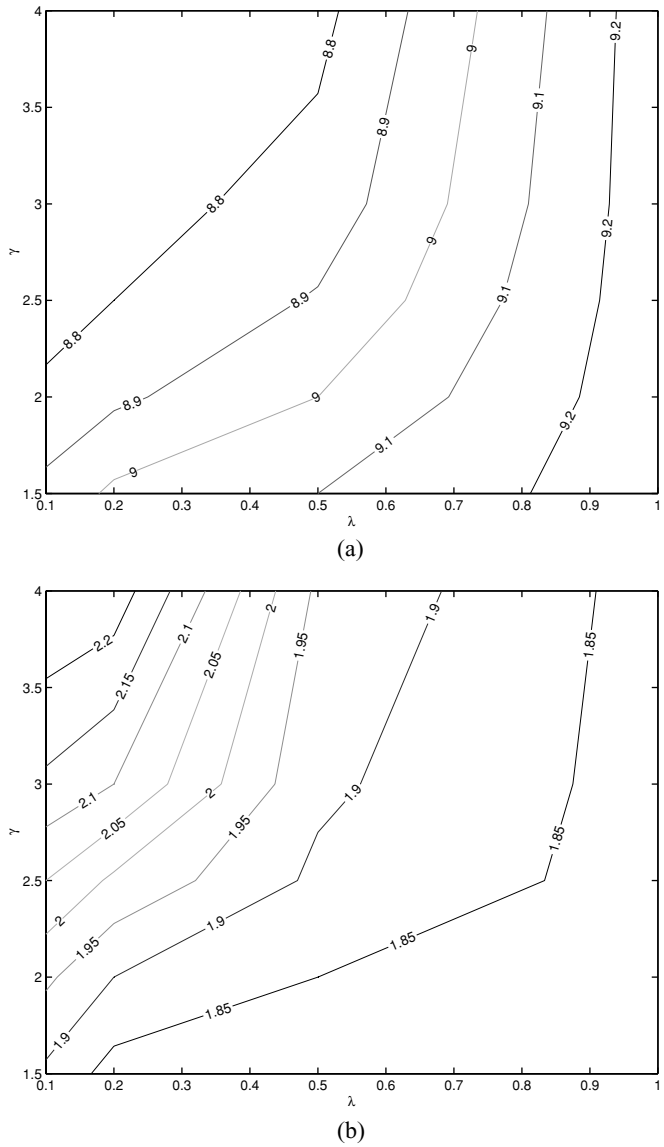

Fig. 3. Contour plot of zero-state ARL for the ACUSUM-C chart with $\delta^+_{\min} = 1.0$: (a) $\delta = 1.0$; and (b) $\delta = 3.0$.

chart with $k = \delta^+_{\min}/2 = 0.5$. When detecting a large shift of $\delta = 3$, as shown in Fig. 3(b), the minimum ARL at a large shift can be approached when $\gamma \to 0$ or $\lambda \to 1$, both of which use the current observation X_t as the mean estimation.

Comparing Figs. 3(a) and 3(b), it can be observed that the percentage change in the ARL_1 values at a large mean shift is much more significant than that at a small mean shift for different values λ and γ considered here. From Fig. 3(a), the ARL_1 values for a small shift of $\delta = 1$ could deteriorate from 8.8 to 9.2, corresponding to an increase of less than 5%. However, as can be seen from Fig. 3(b), the ARL_1 values for a large shift of $\delta = 3$ could drop from 2.23 to 1.85, corresponding to a decrease of 17%. This observation clearly indicates that the parameters λ and γ can

be jointly adjusted to greatly improve the performance of ACUSUM-C charts at larger mean shifts while only causing a minor loss in the detection performance at small mean shifts.

Furthermore, Fig. 3 demonstrates that too large values of λ should be avoided for the ACUSUM-C chart in order to balance the sensitivity to both small and large shifts. As shown in Figs. 3(a) and 3(b), when $\lambda \geq 0.8$, the performance of the ACUSUM-C chart at both small and large shifts is nearly independent of the choice of γ. Therefore, for $\lambda \geq 0.8$, the introduction of the parameter γ cannot improve the chart performance but it increases the complexity if too large λ values are used in the ACUSUM-C chart. This indicates that the correct choice of the parameter λ is of primary importance in the ACUSUM-C charts. Analogously, too small values of γ should be avoided for the design of ACUSUM-C charts. The reason is that when $\gamma \to 0$, the EWMA-C performs closely to the Shewhart statistic, regardless the choices of λ. Therefore, too large λ values and/or too small γ values should be avoided for practical use.

4.2. *Choices of parameters*

One may follow a two-step procedure introduced in Capizzi and Masarotto (2003) to find the optimal or nearly optimal design based on a two-step procedure. However, this approach is too complicated in practice since the design of ACUSUM-C charts involves the joint optimization of the three parameters, $\delta^+_{\min}$, λ and γ. Rather, we provide some guidelines to help practitioners design an ACUSUM-C chart in the following.

The value of $\delta^+_{\min}$ can be specified in advance, which represents the minimum feared shift size. Once $\delta^+_{\min}$ is chosen, λ and γ should be chosen in a way to balance the sensitivity to small and large shifts in order to provide an overall good performance. Extremely large λ values and/or too small γ values should be avoided since the other parameter has no use in adjusting the chart performance in these cases, as discussed above. In practice, for the sake of simplicity, we may follow rules of thumb $0.1 \leq \lambda \leq 0.4$ (Jones *et al.*, 2001; Chen and Elsayed, 2002) and $1.5 \leq \gamma \leq 3$ (Yashchin, 1995). Figures 3(a) and 3(b) show that these choices provide a reasonable trade-off between detecting various levels of mean shifts efficiently.

5. An illustrative example

To illustrate how the ACUSUM-C procedure is implemented and compare it with a conventional CUSUM, we use a numerical example from Montgomery (2005, Table 8-2). Table 4 presents the observations, deviations and corresponding ACUSUM-C statistics. Suppose the shift range

Table 4. An illustrative example.

(Adapted from Montgomery (2005) Table 8-2)

	1σ Shift					*3σ Shift*				
				ACUSUM-C					*ACUSUM-C*	
t	*Obs* X_t	*Dev* $X_t - 10$	*CUSUM* C_t^+	$\hat{\delta}_t$	Z_t^+	*Obs* X_t	*Dev* $X_t - 10$	*CUSUM* C_t^+	$\hat{\delta}_t$	Z_t^+
1	9.45	−0.55	0	−0.17	0.00	9.45	−0.55	0	−0.17	0.00
2	7.99	−2.01	0	−0.72	0.00	7.99	−2.01	0	−0.72	0.00
3	9.29	−0.71	0	−0.72	0.00	9.29	−0.71	0	−0.72	0.00
4	11.66	1.66	0.66	0.00	1.16	11.66	1.66	0.66	0.00	1.16
5	12.16	2.16	1.82	0.65	2.82	12.16	2.16	1.82	0.65	2.82
6	10.18	0.18	1	0.51	2.50	10.18	0.18	1	0.51	2.50
7	8.04	−1.96	0	−0.23	0.04	8.04	−1.96	0	−0.23	0.04
8	11.46	1.46	0.46	0.27	1.00	11.46	1.46	0.46	0.27	1.00
9	9.2	−0.8	0	−0.05	0.00	9.2	−0.8	0	−0.05	0.00
10	10.34	0.34	0	0.07	0.00	10.34	0.34	0	0.07	0.00
11	10.03	0.03	0	0.06	0.00	12.03	2.03	1.03	0.66	1.53
12	12.47	2.47	1.47	0.78	1.97	14.47	4.47	**4.5**	2.37	**9.32**
13	11.51	1.51	1.98	1.00	2.98	13.51	3.51	7.01	2.71	15.16
14	10.4	0.4	1.38	0.82	2.88	12.4	2.4	8.41	2.62	18.01
15	11.08	1.08	1.46	0.90	3.46	13.08	3.08	10.49	2.76	22.70
16	10.37	0.37	0.83	0.74	3.33	12.37	2.37	11.86	2.64	25.48
17	11.62	1.62	1.45	1.00	**4.45**	13.62	3.62	14.48	2.93	31.79
18	11.31	1.31	1.76	1.10	5.29	13.31	3.31	16.79	3.05	37.24
19	9.52	−0.48	0.28	0.62	4.31	11.52	1.52	17.31	2.59	37.82
20	11.84	1.84	1.12	0.99	5.65	13.84	3.84	20.15	2.96	44.81

to be detected of interest is [1, 3]. The parameters of the ACUSUM-C chart is chosen as follows: (i) set $\delta^+_{\min}$ to $\delta^+_{\min} = 1$; (ii) based on the rules of thumb introduced above, a value of $\lambda = 0.3$ is selected within the range [0.1, 0.4] and a value of $r = 3$ is selected within the range [1.5, 3]; and (iii) a threshold of $h = 4.39$ is chosen to provide a desired in-control ARL of 400. The CUSUM chart is designed to optimize the midpoint of the shift interval [1, 3], i.e., $\delta_0 = 2$. The parameters of the CUSUM chart are thus chosen as $k = 1$ and $h = 2.214$ for providing the same in-control ARL as the ACUSUM-C chart.

As described in Montgomery (2005), the in-control mean is ten and the standard deviation is one. Only upward shifts are considered in this example. The process is in control for the first ten observations, and a step mean shift occurs at the 11th observation. For illustration, shifts of 1σ and 3σ are both considered in the example. For the ACUSUM-C procedure, the EWMA-C estimate of the process mean must be calculated. For example, in the tenth observation, the forecast error is

$$e_{10} = (X_{10} - 10) - \hat{\delta}_9 = 0.34 - (-0.05) = 0.39.$$

Since $|e_{10}| < r = 3$, the EWMA-C statistic is computed as

$$\hat{\delta}_{10} = 0.30 \times 0.34 + (1 - 0.3) \times (-0.05) = 0.067.$$

As this estimate is less than $\delta^+_{\min} = 1$, $\hat{\delta}^+_{10} = \delta^+_{\min}$. The ACUSUM-C statistic at the tenth observation is

$$Z^+_{10} = \max(0, Z^+_9 + 0.5(0.34 - 0.5)) = 0.$$

The other ACUSUM-C values are obtained in a similar way.

Figures 4(a) and 4(b) respectively plot the behavior of the two control charts when a small shift of 1σ and a large shift of 3σ is present. As shown in Fig. 4(a), the ACUSUM-C chart signals an out-of-control situation at observation 17 whereas the CUSUM chart does not signal. On the other hand, for detecting a large shift of 3σ, both the ACUSUM-C and CUSUM charts can trigger an out-of-control signal at observation 12. These observations illustrate that the traditional CUSUM chart designed based on a prespecified mean shift may perform poorly when the actual mean shift is different from the specified. Therefore, it cannot provide an overall good detection performance over a range of shifts. In contrast, the ACUSUM-C procedure using a mean estimator to dynamically change its reference value possesses this potential.

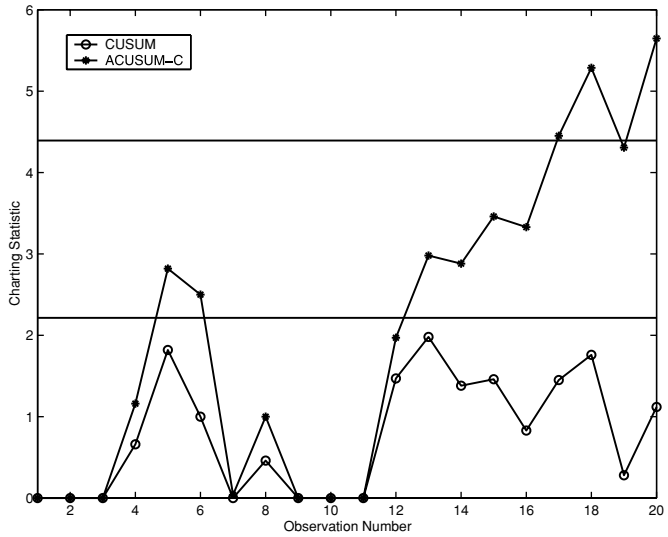

Fig. 4. A numerical example of the ACUSUM-C chart (Montgomery, 2005): (a) a 1σ shift; and (b) a 3σ shift.

6. Concluding remarks

The adaptive CUSUM chart, which adaptively adjusts its parameters as a function of some adaptive estimate of the current process mean, is an appealing idea when one would expect a scheme to be effective at detecting a range of mean shift sizes. The paper extends the work of Sparks' ACUSUM chart from the case based on an EWMA mean estimator to a more general case based on the EWMA-C estimator. Moreover, this paper suggests using a linear weight function that is motivated by the likelihood ratio test.

Due to the attractive property of the EWMA-C statistic in estimating both small and large mean shifts, the comparison results indicate that the ACUSUM-C chart can perform better than the traditional ACUSUM chart at large mean shifts and closely at small mean shifts as expected. The comparison results show that the ACUSUM-C chart performs better than the combined CUSUM-Shewhart chart and the traditional CUSUM chart. Compared to the AEWMA chart, the ACUSUM-C chart has competitive performance but is less sensitive to parameter selection.

It is surprising to observe that the ACUSUM-C chart may even outperform the standard CUSUM for the very

size mean shift for which the CUSUM has been designed. This is surprising because many believe the CUSUM has the smallest zero-state ARL possible for detecting step mean shifts in Gaussian processes, based on the results of Moustakides (1986). We demonstrate that this is not the case.

References

Banzal, R.K. and Papantoni-Kazakos, P. (1986) An algorithm for detecting a change in stochastic process. *IEEE Transactions on Information Theory*, **32**, 227–235.

Barnard, G. (1959) Control charts and stochastic processes. *Journal of the Royal Statistical Society, Ser B*, **21**, 239–257.

Basseville, M. and Nikiforov, I. (1993) *Detection of Abrupt Changes*, Prentice-Hall, Englewood Cliffs, NJ.

Brook, D. and Evans, D.A. (1972) An approach to the probability distribution of CUSUM run length. *Biometrika*, **59**, 539–549.

Capizzi, G. and Masarotto, G. (2003) An adaptive exponentially weighted moving average control chart. *Technometrics*, **45**, 199–207.

Chen, A. and Elsayed, E.A. (2000) An alternative mean estimator for process monitored by SPC charts. *International Journal of Production Research*, **38**, 3093–3109.

Chen, A. and Elsayed, E.A. (2002) Design and performance analysis of the exponentially weighted moving average mean estimate for process subject to random step changes. *Technometrics*, **44**, 379–389.

Chernoff, H. and Zacks, S. (1964) Estimating the current mean of a normal distribution which is subjected to changes in time. *Annals of Mathematical Statistics*, **35**, 999–1018.

Jones, L.A., Champ, C.W. and Rigdon, S.E. (2001) The performance of exponentially weighted moving average charts with estimated parameters. *Technometrics*, **43**, 156–167.

Lorden, G. (1971) Procedures for reacting to a change in distribution. *Annals of Mathematical Statistics*, **42**, 1897–1908.

Lorden, G. and Eisenberger, I. (1973) Detection of failure rate increases. *Technometrics*, **15**, 167–175.

Lucas, J.M. (1982) Combined Shewhart-CUSUM quality control schemes. *Journal of Quality Technology*, **14**, 51–59.

Lucas, J.M. and Crosier, R.B. (1982) Fast initial response for CUSUM quality-control schemes: Give your CUSUM a head start. *Technometrics*, **24**, 199–205.

Lucas, J.M. and Saccucci, M.S. (1990) Exponentially weighted moving average control schemes: properties and enhancements. *Technometrics*, **32**(1), 1–12.

Montgomery, D.C. (2005) *Introduction to Statistical Quality Control*, fifth edition. Wiley, New York, NY.

Moustakides, G.V. (1986), Optimal stopping times for detecting changes in distributions. *Annals of Statistics*, **14**, 1379–1387.

Page, E.S. (1954) Continuous inspection schemes. *Biometrika*, **42**, 243–254.

Pollak, M. (1985) Optimal detection of a change in distribution. *Annals of Statistics*, **13**, 206–227.

Roberts, S.W. (1959) Control chart tests based on geometric moving averages. *Technometrics*, **1**, 239–250.

Shu, L. and Jiang, W. (2006) A Markov chain model for the adaptive CUSUM control chart. *Journal of Quality Technology*, **38**, 135–147.

Siegmund, D. (1985) *Sequential Analysis: Tests and Confidence Intervals*, Spring-Verlag, New York, NY.

Sparks, R.S. (2000) CUSUM charts for signalling varying location shifts. *Journal of Quality Technology*, **32**, 157–171.

Wald, A. (1947) *Sequential Analysis*, Wiley, New York, NY.

Wu, Z., Wang, P.H. and Wang, Q.N. (2008) A loss function-based adaptive control chart for monitoring the process mean and variance. *The International Journal of Advanced Manufacturing Technology*, Online First.

Yashchin, E. (1995) Estimating the current mean of a process subject to abrupt changes. *Technometrics*, **37**, 311–323.

Zhao, Y., Tsung, F. and Wang, Z. (2005) Dual CUSUM control schemes for detecting a range of mean shifts. *IIE Transactions*, **37**, 1047–1057.

Appendices

Appendix A: The likelihood ratio test

Assume the sequence of observations, $\{X_i, i = 1, 2, \ldots, t\}$, is independent and normally distributed. Let H_0 represent the hypothesis that the process has been in control, i.e., $\mu = 0$, and let H_1 be the alternative hypothesis that the process mean has shifted by a size of δ_0 at a certain time t_0 $(t_0 < t)$. Let f be the joint density function of X_i values $i = 1, 2, \ldots, t$. Then, the log-likelihood ratio corresponding to H_0 and H_1 is (Page, 1954):

$$\begin{aligned} LR_{t_0} &= \ln\left(\frac{f(X_1, X_2, \ldots, X_t|H_1)}{f(X_1, X_2, \ldots, X_t|H_0)}\right) \\ &= \ln\left(\frac{\prod_{i=1}^{t_0}\exp\left[-X_i^2/2\right]\times\prod_{i=t_0+1}^{t}\exp[-(X_i-\delta_0)^2/2]}{\prod_{i=1}^{t}\exp\left[-X_i^2/2\right]}\right) \\ &= \ln\left(\prod_{i=t_0+1}^{t}\exp\left[-\frac{1}{2}\left((X_i-\delta_0)^2-X_i^2\right)\right]\right) \\ &= \sum_{i=t_0+1}^{t}\delta_0\left(X_i-\frac{\delta_0}{2}\right). \end{aligned}$$

This implies that if one is interested in detecting a change in the process mean, the CUSUM chart increment at time t should be

$$I_t = \delta_0\left(X_t - \frac{\delta_0}{2}\right).$$

Then the one-sided upper and lower CUSUM statistics for detecting an increase and a decrease in the process mean are given by

$$g_t^+ = \max(0, g_{t-1}^+ + \delta_0(X_t - \delta_0/2)),$$

and

$$g_t^- = \min(0,\ C_{t-1}^- - \delta_0(X_t - \delta_0/2)),$$

respectively. Note that the factor δ_0 pre-multiplying the term $(X_t - \delta_0/2)$ is usually omitted for the sake of simplicity in practice, because a fixed value δ_0 can be absorbed into the decision interval. If so, the g_t^+/g_t^- statistic reduces to the C_t^+/C_t^- statistic. In practice, the future mean shift δ_0 is often unknown and needs to be estimated. In this case, it is intuitive to replace δ_0 by its estimate in the increment I_t. This motivates the use of the linear weight function $w(\hat{\delta}_t) = \hat{\delta}_t$ in the ACUSUM chart.

Appendix B: Markov chain representation of ACUSUM-C charts

Similar to the Markov chain model for ACUSUM charts discussed in Shu and Jiang (2006), the random vector $\mathbf{W}_t = (Z_t^+, \hat{\delta}_t)'$ can be represented by a two-dimensional Markov chain. The region $\{0 \le Z_t^+ \le h, -\infty < \hat{\delta}_t < +\infty\}$ can be divided into subregions to obtain a discretized Markov chain approximation.

Let m_1 be the number of subintervals along the axis Z^+ over the range $[0, h]$, i.e., the width of each subinterval is $\omega = (2h)/(2m_1 - 1)$, except that the width of the first one is $\omega/2$. These subintervals along the Z^+-axis are labeled as $i = 0, 1, 2, \ldots, (m_1 - 1)$. The infinite interval $(-\infty, +\infty)$ along $\hat{\delta}$-axis is first segmented as $(-\infty, -L)$, $[-L, L]$ and $(L, +\infty)$. Throughout this paper, the value of L is chosen as large as

$$L = 8\sqrt{\frac{\lambda}{2-\lambda}},$$

so that it would be nearly impossible for δ_t to go beyond the interval $[-L, L]$ under the situations we considered. The interval $[-L, L]$ is then further partitioned into an odd number of m_2 subintervals, say, $j = 0, 1, 2, \ldots, (m_2 - 1)$, each of length $\Delta = 2L/m_2$. Also, let the intervals $(-\infty, -L)$ and $(L, +\infty)$ be labeled as $j = -1$ and $j = m_2$, respectively. We approximate Z_t^+ and $\hat{\delta}_t$ by the mid value of each subinterval. It is clear that the discretized Markov chain has $N = m_1 \times (m_2 + 2)$ transient states.

For $n = 1, 2, \ldots, m_1 - 1$ and $l \ne -1, m_2$, the transition probability of $\{Z_t^+, \hat{\delta}_t\}$ from state (i, j) to state (n, l), $P_{(i,j)(n,l)}$ follows:

$$\begin{aligned}
P_{(i,j)(n,l)} &= \Pr\{Z_t^+ \text{ in state } n, \hat{\delta}_t \text{ in state } l \mid Z_{t-1}^+ \text{ in state } i, \\
&\qquad \hat{\delta}_{t-1} \text{ in state } j\} \\
&= \Pr\{(n-0.5)\omega < i\omega + \hat{\delta}_t^+\left(X_t - \frac{\hat{\delta}_t^+}{2}\right) \\
&< (n+0.5)\omega, -L + l\Delta \\
&< -L + (j+0.5)\Delta + \phi(e_t) < -L + (l+1)\Delta \mid Z_{t-1}^+ \\
&= i\omega, \hat{\delta}_{t-1} = -L + (j+0.5)\Delta\} \\
&= \Pr\{(n-i-0.5)\omega/\hat{\delta}_t^+ + \hat{\delta}_t^+/2 < X_t \\
&< (n-i+0.5)\omega/\hat{\delta}_t^+ + \hat{\delta}_t^+/2, \\
&\quad -L + (j+0.5)\Delta + \phi^{-1}[(l-j-0.5)\Delta] < X_t \\
&< -L + (j+0.5)\Delta + \phi^{-1}[(l-j+0.5)\Delta]\} \\
&= \Pr\{a_1 < X_t < a_2,\ b_1 < X_t < b_2\} \\
&= \Pr\{\min[b_2, \max(a_1, b_1)] < X_t < \max[b_1, \min(a_2, b_2)]\},
\end{aligned} \tag{A1}$$

where values of a_1, a_2, b_1 and b_2 are constants given by

$$\begin{aligned}
a_1 &= (n-i-0.5)\omega/\max(\delta_{\min}^+, [-L + (l+0.5)\Delta]) \\
&\quad + \max(\delta_{\min}^+, [-L + (l+0.5)\Delta])/2, \\
a_2 &= a_1 + \omega/\max(\delta_{\min}^+, [-L + (l+0.5)\Delta]), \\
b_1 &= -L + (j+0.5)\Delta + \phi^{-1}[(l-j-0.5)\Delta], \\
b_2 &= -L + (j+0.5)\Delta + \phi^{-1}[(l-j+0.5)\Delta].
\end{aligned}$$

The inverse of $\phi(\cdot)$ is

$$\phi^{-1}(v) = v/\lambda$$

for an EWMA estimator while it is (Capizzi and Masarotto, 2003)

$$\phi_\gamma^{-1}(v) = \begin{cases} v - (1-\lambda)\gamma, & v < -\lambda\gamma, \\ v/\lambda, & |v| \le \lambda\gamma, \\ v + (1-\lambda)\gamma, & v > \lambda\gamma, \end{cases}$$

for an EWMA-C estimator. Note that, we can set $b_1 = -\infty$ or $b_2 = \infty$ when $l = -1$ or m_2, and set $a_1 = -\infty$ when $n = 0$. These settings allow the above calculations to remain valid.

The ARL of this discretized Markov chain can be evaluated by

$$\text{ARL} = \mathbf{p}_{\text{ini}}' \times (\mathbf{I} - \mathbf{R})^{-1} \times \mathbf{1},$$

where $\mathbf{p}_{\text{ini}}$ is the initial probability vector of $(Z_0^+, \hat{\delta}_0)'$. The zero-state ARL is computed based on an initial vector of $Z_0^+ = 0$ and $\hat{\delta}_0 = 0$. To derive the steady-state ARL, Lucas and Saccucci (1990) suggested using the cyclical steady-state probability vector obtained by altering the transition probability matrix so that the control statistic is reset to the initial state whenever it goes into the out-of-control state. That is, the initial probability $\mathbf{p}_{\text{ini}} = \mathbf{p}_{\text{ss}}$ is solved from $\mathbf{p}_{\text{ss}} = \mathbf{P}_1' \times \mathbf{p}_{\text{ss}}$ subject to $\mathbf{1}' \times \mathbf{p}_{\text{ss}} = 1$, where:

$$\mathbf{P}_1 = \begin{pmatrix} \mathbf{R} & (\mathbf{I} - \mathbf{R})\mathbf{1} \\ 0 \cdots 1 \cdots 0 & 0 \end{pmatrix}.$$

Because the above Markov chain model of ACUSUM charts is a bivariate chain, the transition matrix can become quite large as m_1 and m_2 increase. Shu and Jiang (2006) have demonstrated the accuracy of the approximation when $m_1 = m_2 = 39$. Since the interval $[-L, L]$ for $\hat{\delta}_t$ is two-sided while the interval $[0, h]$ for Z_t^+ is one-sided, it is reasonable to choose m_2 larger than m_1. In our numerical analysis, $m_1 = 27$ and $m_2 = 39$ were used, which helped us quickly obtain ARL approximations of ACUSUM-C charts with a reasonable accuracy.

Biographies

Wei Jiang is an Assistant Professor in the Department of Systems Engineering and Engineering Management at Stevens Institute of Technology. He obtained his Ph.D. degree from the Hong Kong University of Science and Technology in 2000 and B.Sc. and M.Sc. degrees from Xi'an Jiaotong University in 1989 and 1992, respectively. He worked at AT&T Labs before joining Stevens. His current research interests include statistical methods for quality improvement, data mining and enterprise intelligence. He is an Associate Editor of the *Journal of Statistical Computation and Simulation* and currently serves as a vice chair for the Data Mining section of the Institute of Operations Research and Management Sciences (INFORMS). He was a recipient of the NSF CAREER Award in 2006.

Lianjie Shu is an Assistant Professor in the Faculty of Business Administration at the University of Macau. He received his B.Sc. degree in Mechanical Engineering and Automation from Xi'an Jiao Tong University, and his Ph.D. in Industrial Engineering and Engineering Management

from the Hong Kong University of Science and Technology (HKUST). His current research interests include quality engineering and management, process control and monitoring, and healthcare surveillance.

Daniel W. Apley is an Associate Professor of Industrial Engineering and Management Sciences at Northwestern University, Evanston, IL, where he serves as the Director of the Manufacturing and Design Engineering Program. He obtained B.S., M.S. and Ph.D. degrees in Mechanical Engineering and an M.S. degree in Electrical Engineering from the University of Michigan. Prior to joining Northwestern University in 2003, he served on the faculty of Texas A&M University for 5 years. His primary research interests are manufacturing variation reduction and quality control, with a focus on processes in which advanced measurement, data collection and automatic control technologies are prevalent. He received the NSF CAREER award for his research and teaching in this area in 2001. He is a past chair of the Quality, Statistics & Reliability Section of INFORMS, and currently serves as an Associate Editor for *Technometrics* and is on the Editorial Board of the *Journal of Quality Technology*.